

University of St Andrews

School of Economics & Finance Discussion Papers

Moral Behaviour, Altruism and Environmental Policy

Marc Daube and David Ulph

School of Economics and Finance Discussion Paper No. 1409
3 Feb 2014
JEL Classification: Q5; Q58; D110
Keywords: Altruism; Climate Change; Environmental Economics; Environmental Tax; Externalities; Moral Behaviour; Pro-Social Behaviour; Public Goods

School of Economics & Finance
Online Discussion Paper Series
issn 2055-303X
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk

Moral Behaviour, Altruism and Environmental Policy

Marc Daube[12]
University of St Andrews

David Ulph[3]
University of St Andrews

Abstract

Free-riding is often associated with self-interested behaviour. However if there is a global mixed pollutant, free-riding will arise if individuals calculate that their emissions are negligible relative to the total, so total emissions and hence any damage that they and others suffer will be unaffected by whatever consumption choice they make. In this context consumer behaviour and the optimal environmental tax are independent of the degree of altruism. For behaviour to change, individuals need to make their decisions in a different way. We propose a new theory of moral behaviour whereby individuals recognise that they will be worse off by not acting in their own self-interest, and balance this cost off against the *hypothetical* moral value of adopting a Kantian form of behaviour, that is by calculating the consequences of their action by asking what would happen if everyone else acted in the same way as they did. We show that: (a) if individuals behave this way, then altruism matters and the greater the degree of altruism the more individuals cut back their consumption of a 'dirty' good; (b) nevertheless the optimal environmental tax is exactly the same as that emerging from classical analysis where individuals act in self-interested fashion.

Key Words: Altruism; Climate Change; Environmental Economics; Environmental Tax; Externalities; Moral Behaviour; Pro-Social Behaviour; Public Goods

JEL Codes: Q5; Q58; D110

[1] Financial support from the Economic and Social Research Council is gratefully acknowledged.

[2] School of Economics & Finance, University of St Andrews, St Andrews, Fife KY16 9AL, Scotland, UK, E-mail: mpl@st-andrews.ac.uk

[3] School of Economics & Finance, University of St Andrews, St Andrews, Fife KY16 9AL, Scotland, UK, E-mail: du1@st-andrews.ac.uk

1 Introduction

In a recent paper assessing the challenges that policy-makers and society face in addressing climate change, Galarraga and Markandya (2009) point out the key ethical and welfare considerations that need to be taken into account – in particular the intra and inter-generational impact of the damage that climate change may bring.

While economists frequently consider these ethical considerations in terms of the formulation of the appropriate welfare objective for policy-makers to pursue, an equally important issue is how far individuals themselves take these factors into account when deciding what actions to take and what policies they are willing to support. This raises the important question of the extent to which policy action might be needed if individuals themselves are willing to alter their behaviour as they recognise the potential harm that their actions might cause.

As both Stern (2007) and Galarraga and Markandya (2009) point out, climate change represents one of the largest externalities that society has had to face, and, like all externalities, the fundamental need for policy intervention arises from free-riding behaviour – in this case not just by individuals and companies but also by governments.

In the classical analysis of externalities, free-riding arises because individuals are purely self-interested, and so perceive that, while they bear all the costs of changing their consumption behaviour, they may get only a very small gain in terms of the reduced damage that they themselves will suffer. In the extreme case, individuals may calculate that their emissions are so insignificant relative to the total, that total emissions and hence any damage that they (and others) might suffer will be unaffected by whatever consumption choices they make.[4] In these circumstances individuals make their consumption choices ignoring any effect these choices have on climate change and the damage it will cause.[5] The classic prescription is the introduction of a Pigovian tax (equal to social marginal damage) so that individuals face the full economic cost of their consumption decisions.

[4] We will refer to this inability to influence the total level of the externality as the 'atomless' consumer.

[5] Echoes of such free-riding behaviour can be found in a paper by Longo et al. (2012) reporting on a study conducted in the Basque Country on people's Willingness to Pay for the Ancillary Benefits of Climate Change mitigation. They note that, as in many such studies, while many people reveal a positive willingness to pay, there is group of "protestors" who say they do not want to pay for these benefits, either because they do not think the proposed policy actions will be effective, or because they feel that others should contribute.

It is sometimes thought that if individuals are not self-interested, but instead are altruistic and so take account of the effect of their actions on others, this may overcome free-riding behaviour. However, as we will show, as long as individuals continue to believe that total emissions are completely unaffected by whatever they do, then however much they care about others, their behaviour will not change, and the optimal policy is to impose exactly the same tax as if individuals were self-interested.

For individual behaviour to change, individuals need to make their consumption decisions in a different way. While there have been a number of theories of pro-social behaviour, these typically involve individuals obtaining some kind of utility gain from behaving *morally*. In this paper we propose a new theory of moral behaviour whereby individuals recognise that they will be worse off by not acting in their own self-interest, and balance this cost off against a *hypothetical* moral value of adopting a Kantian form of behaviour, that is by calculating the consequences of their action by asking what would happen if everyone else acted in the same way as they did. We show that

(i) Individuals behaving this way will adjust their behaviour to take account of the impact of their decision on themselves and others.

(ii) If individuals behave in this way, then altruism matters and the greater the degree of altruism the more individuals cut back their consumption of a "dirty" good.

(iii) Nevertheless the optimal environmental tax is exactly the same as that emerging from the classical analysis where individuals purely self-interested.

The paper will proceed as follows. In Section 2 we conduct a brief discussion of the literature on pro-social behaviour and altruism. Section 3 then develops the model, starting with standard theory excluding any form of altruism in Section 3.1. This serves as counterfactual to the remaining analysis. We then proceed in Section 3.2 to layer on a type of Pure Altruism that will be defined carefully in what follows. Section 3.3 then turns to the central element of this paper by developing the theory of moral behaviour. Finally, Section 4 will present a brief discussion of the model and the results.

2 Review of Literature

As described in the introduction, pro-social (and pro-environmental) behaviour is usually at odds with standard economic analysis, which predicts that in a non-cooperative setting, individuals only make negligible contributions to public goods. For example, Andreoni (1988) shows that in large populations

the share of individuals making contributions to a public good tends to zero as the free-riding effect dominates.[6] However, when contributing to the public good also yields some utility benefit to the individual, voluntary contributions can be consistent with standard economic models. Andreoni (1989, 1990) models the individual's utility not just as a function of the consumption of the private and public goods, but also of the individual's contribution to the public good itself. This is commonly referred to as the 'warm-glow' effect and describes a form of Impure Altruism.[7] 'Warm-glow' can also be interpreted as a self-image gain from contributing to the public good.[8] While Andreoni makes no assumptions regarding the psychological cause of this 'warm-glow', various other authors have developed more sophisticated models with regard to the underlying motivation. These models usually work on the premise that individuals derive intrinsic value from a self-image desire or social norms. Bénabou and Tirole (2006), for example, model a "reputational payoff" (p. 1656) from contribution to a public good, which is also a function of the belief others have regarding the type of consumer this individual is. In a model by Ellingsen and Johannesson (2008) the level of social approval depends on whether the individual himself approves of the person who approves him. Nyborg et al. (2006) also construct a model where individuals are motivated by a concern for self-image. However, this self-image is a function of the total benefit a 'green' good yields to the population, as well as their perception of what share of the population is choosing to consume the 'green' option.[9] This means the consumer's intrinsic incentive to

6 Also see Bergstrom et al. (1986).

7 Andreoni's development of warm-glow giving is based on the analysis of the provision of impure public goods by Cornes and Sandler (1984). Impure public goods are public goods that also have some private characteristics, and therefore, contributing to the public good has some private benefit to the individual. The concept of Impure Altruism can be linked to the concept of an impure public good, though the two concepts are not identical. The analysis of impure public goods is of particular importance in the area of environmental policy. For example, Markandya and Rübbelke (2004) analyse ancillary benefits of climate policy and argue that policymakers should consider these more thoroughly. Further, Markandya and Rübbelke (2012) look at the under-provision of impure public technologies in an environmental context while Altemeyer-Bartscher et al. (2013) take into account that climate policy is an impure public good in their analysis of tax-transfer schemes in relation to international public goods.

8 Also note that Andreoni refers to the case of an individual who only cares about the total supply of the public good as Pure Altruism. However, although the total supply of a public good also affects others, we, as well as other literature on altruism, use the term Pure Altruism to refer to an individual's direct concern for the utility of others. We will provide a more detailed description of how we model Pure Altruism later in this section.

9 To some extent this also captures the idea of a social norm or peer pressure.

be pro-social increases as the share of the population acting that way increases.[10] On the other hand, Brekke et al. (2003) develop a model where individuals are able to make a more sophisticated calculation of the "morally ideal effort" (p.1971). This is achieved by evaluating the socially optimal contribution to a public good if they and everybody else were to make the same choice. The individual then derives self-image value depending on how close their contribution is to that socially optimal level, while trading off the self-image gain against the utility benefit from consumption.

Essentially most of the contributions discussed above capture a form of Impure Altruism where, to some extent, the contribution itself matters to an individual's utility. However, altruism is instead often viewed as a concern for the welfare of others.[11] The two main types of altruism that capture an individual's concern for others' welfare are Pure Altruism and Paternalistic Altruism. Pure Altruism uses the idea that an individual's utility may to some degree be a function of others' utility, but not just a specific component of it. Applications of this type of altruism are often used with smaller settings, such as the family where one might care about the welfare of specific individuals (for example Becker (1974, 1981). In large-scale contexts it can also be assumed that an individual cares for the total or average welfare of all other individuals in the population (see Hammond (1987), Johansson (1997) for examples).[12] Paternalistic Altruism differs from Pure Altruism in that it does not assume that an individual's utility is a function of others' utility as such, but a specific component of that utility (see Archibald and Donaldson (1976)). In an environmental context, this component may be the damage experienced by others from the environmentally harmful good. While Impure Altruism only takes into account the individual's contribution to the externality, Paternalistic Altruism means that the individual is affected by others' experience of the externality, regardless of the individual's contribution. All these types of altruism still assume that individuals maximise their utility when acting pro-socially. Genuine Altruism as defined by Kennett (1980), on the other hand, requires that individuals' behaviour is driven by some function other than maximising their utility. For example, the individual may make the

[10] Because what matters in their model is the individual's perception of what others do, Nyborg et al. (2006) further argue that policy makers may be able to influence this perception, for example through advertising. Also, a temporary tax on the environmentally harmful good could move the population to a permanent 'green' equilibrium even when the tax is later removed.

[11] The literature on altruism is very rich and we will only make a superficial survey as relevant to this paper. For more detail, Fontaine (2008) provides a summary of the history of the concept of altruism in economics.

[12] Also see Nyborg and Rege (2003) for basic modelling approaches to public good provision including Pure and Impure Altruism.

consumption choice by maximising a function consisting of their own utility and everybody else's utility, but by doing so will incur a loss in utility compared to standard economic behaviour.[13] Since this implies a deviation from 'rational' behaviour driven by self-interest economists usually assume it is the most drastic form of altruism.

A key component of this paper is the evaluation of what the existence of altruistic and 'moral' behaviour mean for the socially optimal tax on an environmentally harmful good. Johansson (1997) previously modelled the socially optimal tax on an externality for all four types of altruism described above and compares it to the standard Pigovian tax level.[14] His results show that, depending on the type of altruism analysed, the socially optimal tax on the externality can be higher, lower, or equal to the Pigovian tax.[15] In his analysis, Johansson (1997) recognises that Genuine Altruism is the most controversial of all types of altruism as it contradicts mainstream economic theory. Yet the traditional model has already for a long time drawn fundamental criticism. Sen (1977) describes why the economist prefers to assume the existence of a selfish being: "It is possible to define a person's interests in such a way that no matter what he does he can be seen to be furthering his own interests in every isolated act of choice." (p. 322). Indeed, it means that every action can be explained simply through the concept of revealed preferences. This saves the economist from having to take a closer look at what constitutes

[13] See Johansson (1997).

[14] Johansson (1997) uses a model of discrete, identical consumers, who choose between a 'clean' good and a 'dirty' good, which in turn causes the externality. The government imposes a tax on the dirty good, but tax revenues are distributed back to the consumer.

[15] Under Pure Altruism, the optimal level of tax is equal to the Pigovian level in large populations, as the size of altruistic concern has to be small relative to the intrinsic consumption utility in order to maintain realistic proportionality of the utility function. For Paternalistic Altruism, the tax is higher than the Pigovian tax because, as individuals take into account others' experience of the externality, the socially optimally level of the externality is reduced and therefore the optimal tax level increases. For the case of Impure Altruism, the optimal tax is exactly equal to the Pigovian tax. Although the socially optimal level of consumption is lower than the standard level, this shift is exactly the amount that occurs due to the 'warm glow' from the individual's contribution to the externality. Therefore, the optimal tax level does not need to be higher in order to achieve the socially optimal level of consumption. Finally, with a form of Genuine Altruism where individuals maximise a function of the weighted sum of their own utility and everybody else's utility, Johansson (1997) finds that the optimal tax is lower than the Pigovian tax. The socially optimal level of consumption is unchanged from the standard level as this type of altruism does not affect it, but the individual will demand this lower level of consumption due to the function maximised and therefore the requirement on the tax level is reduced. If the weight in the maximisation were equal between the individual's utility and all others' utility, the tax rate would drop to zero.

preferences and utility, or as Sen (1977) puts it: "…a robust piece of evasion." (p. 323). Sugden (1982) further argues that assuming utility-maximising behaviour in the traditional sense may be too constricting and different people might indeed maximise very different functions in order to determine their consumption choice.

The criticism presented here does not seek to suggest that the utility-maximising model is without value, and the critical literature usually acknowledges the usefulness of the traditional model, especially in market-exchange situations. But in the context of this paper, it has to be recognised that environmental issues are often subject to strong moral views. As Frey (1999) summarises, contrary to the utilitarian perspective, a moralist views the environment with unique value and believes that it should be protected purely for ethical reasons and with little or no regard to any trade-offs. Furthermore, Arrow (1986) argues that rationality as the economist defines it often only occurs through market exchange, rather than through the intrinsic motivation of an individual. However, as Shogren and Taylor (2008) point out, environmental resources are frequently not subject to the market exchange required for consistent choices. Furthermore, we put forward that when individuals realise that the government is not able to set the incentives correctly, they may also recognise that the result will be far from optimal for themselves and everybody else without a change in behaviour. This may, in turn, lead to moral motivation that requires a different type of model to explain. Along the lines of Harsanyi (1955), Sen (1977) develops the idea that people may have two preference relations working at the same time. [16]Utility-maximising behaviour with some consideration for the utility of others, he labels as 'sympathy', while actions driven only by their moral value with no consideration of utility, he calls 'commitment'.[17] The latter is very much in line with Kant's concept of 'duty'.[18]

[16] The model developed in this paper can also be thought of as capturing the idea of two different preference sets to some extent. However, combining two preference sets into a conventional model usually assumes that one supersedes the other or individuals follow certain rules (see Thaler and Shefrin (1981) for an example). The model we present in this paper assumes that individuals trade-off moral preferences against utility preferences depending on their propensity to act morally.

[17] Sen (1977) also points out that the resulting behaviour may well be the same as the utility-maximising one, but that this may not have been the individual's motivation.

[18] White (2003) provides a good summary of the essentials of Kant's philosophy as relevant to an economist and discusses linking it to the traditional model of economic behaviour.

Laffont (1975) develops the idea of a Kantian approach to behaviour where individuals consider what would be optimal if they and everybody else were to make the same choice.[19] However, Laffont (1975) assumes that individuals are identical, with all individuals applying this rule equally. In his model this implies that individuals will correctly anticipate that everybody else will indeed behave the same way, meaning that ultimately individuals still maximise their personal utility. Ulph (2006) also presents a model of consumption choice where some people behave in a Kantian sense. A key difference to Laffont (1975) is that individuals have no expectations regarding the behaviour of others and assume that all others continue to behave in the utility-maximising way. Individuals therefore recognise that acting morally means they will suffer a loss of personal utility. The central aspect of Ulph's (2006) model is that individuals assess the intrinsic 'rightness' of their action by comparing the utility they actually get to the utility they would get if everyone else were to make the same choice as this individual.[20] Individuals then trade off the 'rightness' of the action against the associated utility cost when making their consumption choice.

The model we present in this paper is a generalisation and extension of this initial approach. A key difference to Ulph (2006) is that we consider that individuals may also have altruistic concern for others' utility (Pure Altruism). In line with Hammond's (1987) definition, we model that people may exhibit altruistic concern for the utility of all other individuals as a collective. Specifically, our model assumes that an individual's utility function is the sum of their direct utility from consumption and the total utility of all other individuals weighted by the degree of altruistic concern.[21] While Johansson (1997) also models this form of Pure Altruism, we differ by explicitly modelling a continuum of individuals in order to capture the atomless nature of individuals in the context of large-scale environmental problems such as climate change. Such an individual would recognise their atomless nature and hence, when making their consumption decision, takes average and total levels of consumption, welfare and damage from the externality as given. It is noteworthy that both the discrete and continuous approaches are widely

[19] Such an approach (a version of which is also used in our model) is generally inspired by Kant's Formula of Universal Law (a version of the categorical imperative) which states: "Act only according to that maxim whereby you can at the same time will that it become a universal law." (Kant (1875), p.421).

[20] This type of 'Kantian' calculation is similar to the one used by Brekke et al. (2003), although they link the moral value of this Kantian behaviour to a concern for self-image instead.

[21] Although for simplicity reasons we assume that the degree of altruistic concern is the same for all individuals in the population, it is possible to generalise the result to allow for varying degrees of altruism for different types of individuals without affecting the main conclusions from our analysis.

used in the literature depending on the issue to be addressed. Finally, by modelling a combination of two different types of altruism, 'Pure Altruism' (the concern for others' utility) and 'Genuine Altruism' (the propensity to act morally through non-utility-maximising behaviour), we capture the idea that people may be concerned about the welfare of others but may also have some propensity to act morally despite their inability to influence the impact of the environmentally harmful good.

Of course a distinction between Kant's categorical imperative and our model is that Kant sees moral duties as absolute that completely supersede any considerations of utility. Our model, on the other hand, does not use such a binary approach, but allows the individual's propensity to act morally to determine in how far they are driven by the moral value of the action relative to the associated loss in utility. Only individuals with full propensity to act morally will completely disregard their own utility and act in a fully Kantian fashion. At the same time, individuals with no propensity to act morally will not take into account the moral value at all and thus act in a standard utility-maximising way. The propensity to act morally is assumed to be an exogenous parameter that is distributed across different types of individuals in the economy. Since we make no assumption regarding what this distribution may look like, we do not argue that people necessarily have a certain propensity to act morally but we simply evaluate the consequences for the socially optimal tax on an environmentally harmful good if some people do act this way.

3 The Model

3.1 Standard Theory

We start with a population consisting of a continuum of potentially different types of individuals. These types are indexed by $k \in [0,1],\ 0 \leq k \leq 1$. The distribution is given by the density function

$$f(k) > 0, \quad k \in [0,1], \quad \text{where } \int_0^1 f(k)dk = 1.$$

Each individual has the same initial endowment of income $y > 0$ and chooses between a 'clean' good x and a 'dirty' good z. The clean good is a numeraire good with a price of 1 and therefore represents the expenditure on all other goods but z. Its consumption is assumed to generate no externalities. On the other hand, the dirty good generates one unit of emissions per unit consumed, which is a negative externality to all individuals. Further, let $\zeta(.)$ denote the function that assigns to an individual of type k their chosen consumption of the dirty good, $\zeta(k) \geq 0$, and let $\bar{z}$ denote the average consumption

of the dirty good. The size of the population is measured by $M > 0$ and therefore total emissions E are:

$$E = M\bar{z}, \quad where \quad \bar{z} = \int_0^1 \zeta(k) f(k) dk. \tag{1}$$

Each individual derives utility from the personal consumption of the two goods. For simplicity we assume that preferences over the two goods are the same across all types and as such all individuals have identical utility functions.[22] Additionally, utility is assumed to be linear in consumption of the clean good in order to avoid issues of income distribution. This means as long as consumers always consume a positive amount of the clean good, the marginal utility of income is constant and welfare losses that may arise are not due to inequality but inefficiencies.

Utility derived from the consumption of the two good takes the following form:

$$u(x, z, E) = x + \phi(z) - D(E) \tag{2}$$

$$where \ \ \phi'(z) > 0, \ \phi''(z) < 0; \ and \ \ \forall E > 0, \ D'(E) > 0, \ D''(E) \geq 0.$$

The damage from emissions experienced by individuals is captured by the damage function $D(E)$. It is a strictly increasing and convex function for all positive levels of emissions. Furthermore, $\phi(z)$ describes the private benefit gained from consumption of the dirty good and is strictly increasing and strictly concave in z.

The dirty good is produced by a perfectly competitive industry operating with constant unit costs $c > 0$. At the same time, the government imposes an emissions tax $t \geq 0$ on the consumption of z. The tax revenue is redistributed to the individuals through a lump-sum transfer σ that is identical for all individuals. Therefore the government budget constraint is defined by

$$\sigma = t\bar{z}. \tag{3}$$

Using the definition of emissions provided in (1) and the government budget constraint, we can now derive the utility from personal consumption and emissions as a function that only depends on the individual's consumption level of the dirty good, and the average consumption of the dirty good:

[22] Although individuals have identical utility functions, different types of individuals will later be defined by their propensity to act morally.

$$u(z;\bar{z},t) = (y + t\bar{z}) - (c + t)z + \phi(z) - D(M\bar{z}). \quad (4)$$

In line with the idea of the atomless consumer, in all stages of the analysis it is assumed that when making the choice over consumption of the dirty good, the individual always takes the consumption by everyone else as given. This means that the individual treats average emissions as independent of z because their choice of consumption of the dirty good has no influence on average and total emissions. Indeed, based on the definition of total emissions as shown in (1), even if every other individual of the same type as this one were to change their consumption simultaneously, it would still not impact the level of average or total emissions.

Using this very basic setup, the individual chooses z to maximise their direct personal utility as given by (4). Under standard utility-maximising behaviour the chosen level of consumption of the dirty good can therefore be characterised by

$$\phi'[\tilde{z}(t)] = c + t. \quad (5)$$

As we would expect, the left hand side of (5) describes the private marginal benefit of consumption of the dirty good and the right hand side describes the private marginal cost of consumption. Further, as we would expect with the atomless nature of individuals in this model, the damage incurred from the dirty good is not a factor in determining the chosen consumption level under standard theory.

Given this we can determine that for any tax rate t and any dirty good assignment function $\zeta(.)$, social utility across all individuals in the population is

$$S(\zeta,t) = M\int_0^1 u[\zeta(k),\bar{z},t]f(k)dk, \qquad where \quad \bar{z} = \int_0^1 \zeta(k)f(k)dk. \quad (6)$$

Since individuals have identical, and strictly concave utility functions, the socially optimal consumption level requires that everyone consume the same amount of the dirty good. This common level, $\hat{z}$, is defined as

$$\hat{z} = \underset{z}{ArgMax}\, S(\zeta,t) \;=\; \underset{z}{ArgMax}\,[y - cz + \phi(z) - D(Mz)]. \quad (7)$$

As we can see, the social planner takes account of the link between the taxes paid on the dirty good and the lump-sum transfer received by consumers through the government budget constraint. This means that the socially optimal level of consumption is independent of the tax rate t. In addition, the social planner also takes into account the connection between the consumption of the dirty good and total

emissions. As we would expect, this means that the social planner can fully internalise the externality. The socially optimal level of consumption of the dirty good $\hat{z}$ is then implicitly defined by

$$\phi'(\hat{z}) = c + MD'(M\hat{z}), \quad \text{where } \hat{z} < \tilde{z}(0). \tag{8}$$

This shows that, from a social planner's perspective, the social marginal cost of consumption consists of the direct marginal cost of production of the dirty good, as well as the social marginal damage created by the emissions externality $MD'(M\hat{z})$, but not the tax on the dirty good since the revenues are entirely redistributed to the individuals.

From (5) and (8) it is also straightforward to see that the government can achieve the socially optimal consumption level by setting the tax on the dirty good equal to the social marginal damage of consumption (the standard Pigovian tax). Denoting this optimal tax rate by $\hat{t}$, it follows that

$$\hat{t} = MD'(M\hat{z}). \tag{9}$$

So far we have only described the basic model setup and laid out the results under standard theory. These will serve as a counterfactual for the remaining analysis. The next step is now to layer on the concept that individuals may exhibit altruistic concern for the utility of all other individuals in the population.

3.2 Pure Altruism

As discussed in the introduction, Pure Altruism in this model means that individuals may put some weight on the total utility of all other individuals in the population. Individuals will therefore not only maximise their own private utility from consumption, but the (weighted) sum of their own private utility and the total utility of all others. For simplicity we will assume that the degree of altruistic concern does not vary by different types of individuals.[23] The weight given to the total utility of others is denoted

[23] This assumption makes it more straightforward to isolate the key drivers of behaviour when we develop the model of moral behaviour. It is, however, possible to allow the degree of altruism to vary across different types without altering the key results of this paper.

$\alpha \geq 0$ for all k. Then the total welfare of an individual who consumes an amount z of the dirty good is:[24]

$$w(z;\zeta,t,\alpha) = U(z;\bar{z},t) + \alpha S(\zeta,t), \qquad where \quad \bar{z} = \int_0^1 \zeta(k)f(k)dk. \tag{10}$$

Given the atomless nature of the consumer, the total utility of all other individuals in the population is simply equal to the social utility as defined in (6). Furthermore, individuals still treat the consumption of all other individuals as given when making their consumption choice over the dirty good. This means individuals treat the assignment function $\zeta(k)$, and hence the average consumption level $\bar{z}$, as well as the level of social utility $S(.)$, as given. Indeed this means that individuals do not just take as given the damage they themselves suffer from the externality, but also the damage that everyone else suffers.

It follows that, with atomless consumers, if the individual chooses the level of consumption of the dirty good that maximises total individual welfare, then, independent of the level of altruistic concern, consumption will simply be the same as the level that maximises their utility as characterised by (5).

Since we assumed a constant level of α for all types of individuals, we still have identical personal welfare functions for all individuals. Using (10) it is then straightforward to derive that, for any given tax rate t and assignment function $\zeta(.)$, the total welfare of all individuals, or social welfare, is given by

$$W(\zeta,t,\alpha) = (1+\alpha M)S(\zeta,t). \tag{11}$$

This shows that the level of altruistic concern across the population only scales up the total level of welfare, but the socially optimal level of consumption is still one in which everyone consumes the same amount of the dirty good as determined by standard theory. Consequently the level of $\hat{z}$ is still defined by (8) and the optimal tax inducing everyone to consume the socially optimal level is still described by (9). This result leads us to the first proposition.

Proposition 1 *In a world of atomless consumers, with every individual choosing their consumption by maximising their utility, altruistic concern for the utility of others has no impact on the optimal level of consumption for the individual, the socially optimal level of consumption or the socially optimal tax level.*

[24] We will refer to the utility that includes the effect of altruistic concern as 'welfare'. This is simply done to separate the direct type of utility from the altruistic type.

Proof: *By maximising individual welfare as shown in (10), it is straightforward to derive that* $\phi'(z) = c + t$*, which is the same as derived without altruism in (5). Similarly, maximising the social welfare function as shown in (11), we derive that* $\phi'(\hat{z}) = c + MD'(M\hat{z})$*, the same level as under standard theory in (8). Combining these two findings it is also evident that* $\hat{t} = MD'(M\hat{z})$ *induces everyone to consume the same socially optimal amount, the same as shown under standard theory without altruism.*

So far we have established the results with and without altruistic concern for the utility of others in the case of standard utility-maximising behaviour. We will now turn to the key part of this paper and develop the theory of moral behaviour.

3.3 Moral Behaviour

We start by assuming that initially everyone consumes $\tilde{z}(t)$, which is the chosen consumption level consistent with standard theory as derived in Section 3.1. And since we want to investigate whether moral behaviour can make up for a shortfall in government policy, we further begin with the assumption that the tax on the dirty good is below the socially optimal level, i.e. $t < \hat{t}$, and therefore we also have $\tilde{z}(t) > \hat{z}$.

The model of moral behaviour developed here assumes that individuals recognise that if they deviate from the consumption level $\tilde{z}(t)$, they will incur a loss in personal welfare. This means that individuals still know that the rational choice for them and everyone else would be to choose $\tilde{z}(t)$ (given their atomless nature). However, they may also be driven by moral concerns for the environment and recognise that the overall outcome will be far from optimal for themselves and everyone else as long as the government continues to set the wrong tax. This may induce some individuals to act differently. Such an individual may measure the *hypothetical* moral value of deviating from the standard consumption level without expecting any personal benefit from this moral action.[25]

Our model measures the hypothetical moral value of the choice through a type of 'Kantian' calculation as discussed in Section 2. This means that an individual with some propensity to act morally will assess how much better off they and everyone else would be if they and everyone else were to choose the

[25] It is crucial to emphasize that moral value does not mean the individual derives any benefit (or utility) from the moral action.

same level of consumption, z. As already mentioned, the individual recognises that deviating will entail a cost to personal welfare, but may balance off this cost against the moral value, depending on their propensity to act morally. We assume that different types of individuals may differ in their propensity to act morally. Some individuals may only be concerned with the moral value of their action and not take any account of the loss of personal welfare they will incur. Yet other individuals may give no weight to the moral value but fully take account of the personal welfare cost associated with deviating from $\tilde{z}(t)$. In addition to the propensity to act morally, the model recognises that individuals may still exhibit altruistic concern for others' utility as modelled in the previous section. The previous results have shown that altruistic concern does not have any influence on the consumption choice under utility-maximising behaviour so the question arises whether this changes when individuals have some propensity to act morally.

The first step is to quantify the level of welfare loss associated with deviating from the initial consumption level. Let $\zeta(t)$ be the assignment function that assigns everyone the initial level of consumption $\tilde{z}(t)$. Therefore, the average consumption of the dirty good is also $\tilde{z}(t)$. For an individual of type k who gives altruistic weight $\alpha \geq 0$ to the utility of other individuals, this gives an initial welfare level of

$$w(\tilde{z}(t); \zeta(t), t, \alpha) = (1 + \alpha M)\big[y - c\tilde{z}(t) + \phi\big(\tilde{z}(t)\big) - D\big(M\tilde{z}(t)\big)\big]. \tag{12}$$

However, if the individual chooses a different level of consumption such that $\hat{z} \leq z \leq \tilde{z}(t)$, but everyone else carries on consuming $\tilde{z}(t)$, this generates the following level of welfare for the individual:

$$w(z; \zeta(t), t, \alpha) = U(z; \zeta(t), t) + \alpha M\big[y - c\tilde{z}(t) + \phi\big(\tilde{z}(t)\big) - D\big(M\tilde{z}(t)\big)\big]. \tag{13}$$

Combining (12) and (13) leads us to the welfare cost of choosing a different level of consumption of the dirty good, z, as opposed to the initial level, $\tilde{z}(t)$:

$$\begin{aligned} C(z; \tilde{z}(t), t) &= w(\tilde{z}(t); \zeta(t), t, \alpha) - w(z; \zeta(t), t, \alpha) \\ &= \big[\phi\big(\tilde{z}(t)\big) - (c + t)\tilde{z}(t)\big] - [\phi(z) - (c + t)z]. \end{aligned} \tag{14}$$

It is noteworthy that in the calculation of the welfare cost everything that depends on what the other individuals do, and therefore anything associated with the level of altruistic concern, has cancelled out. In addition, since an individual's choice has no impact on the total level of emissions, the damage function has also cancelled out. The welfare cost is purely driven by the difference in the priavte benefit

and cost elements of the dirty good. Were the individual to minimise this cost they would choose $\phi'(z) = c + t$ and then we would be back at the standard level of consumption where $z = \tilde{z}(t)$. Intuitively, in order to minimise the cost of deviating from the initial consumption level, one would not deviate at all.

The next step is to derive a measurement of the hypothetical moral value of choosing a different level of consumption. As already mentioned, this moral value is assessed by evaluating the welfare the individual would obtain if they and everyone else were to choose that same level of consumption. For an individual who places an altruistic weight $\alpha \geq 0$ on the welfare of all other individuals this would yield the following level of welfare

$$w(z, \zeta^K(z); t, \alpha) = (1 + \alpha M)[y - cz + \phi(z) - D(Mz)], \tag{15}$$

where $\zeta^K(z)$ is the 'Kantian' assignment function that assigns everyone else the same level of consumption of the dirty good as that chosen by the individual.

There are two factors the individual will incorporate when making this calculation. First, the individual sees that if they and everyone else choose the same level of z, then the lump-sum transfer to the individual from the government tax revenues will equal the tax paid on the dirty good, rendering the tax rate irrelevant for the level of welfare obtained. Second, the individual takes account of the impact the choice of z has on total emissions and therefore the damage associated with it. Since it is a specific calculation of the level of welfare if everyone were to choose the same level of the dirty good, the individual effectively makes the same calculation a social planner would make. This means that the individual does not just evaluate the moral value for themselves, but also the benefit to everyone else in the population. Therefore, the moral value of choosing a level of consumption, z, as opposed to the initial level, $\tilde{z}(t)$, is

$$\begin{aligned} B(z; \tilde{z}(t), t, \alpha) &= w(z, \zeta^K(z); t, \alpha) - w(\tilde{z}(t); \zeta(t), t, \alpha) \\ &= (1 + \alpha M)\big\{[y - cz + \phi(z) - D(Mz)] - \big[y - c\tilde{z}(t) + \phi\big(\tilde{z}(t)\big) - D\big(M\tilde{z}(t)\big)\big]\big\}. \end{aligned} \tag{16}$$

From here it is easy to see that were an individual to simply maximise the hypothetical moral value, the individual would choose the socially optimum level of consumption $\hat{z}$ regardless of their level of altruistic concern for the utility of others.

We can now put together the moral value and welfare cost of deviating from the initial level by assuming an individual will make the consumption choice putting some weight on the moral value of

deviating and balancing this off against the associated loss in personal welfare. Let this propensity to act morally be measured by μ. For simplicity we assume that individuals will choose their consumption by maximising the weighted difference between the moral value and the associated personal welfare cost:

$$\mu B - (1-\mu)C, \qquad 0 \leq \mu \leq 1. \tag{17}$$

Substituting (14) and (16) into (17), we see that for an individual with $0 \leq \mu \leq 1$, consumption of the dirty good will be chosen to

$$\underset{z}{MAX} \quad \mu(1+\alpha M)[y - cz + \phi(z) - D(Mz)] + (1-\mu)[\phi(z) - (c+t)z]. \tag{18}$$

At this point it is noteworthy that, due to the linear nature of the choice function, the initial level of consumption $\tilde{z}(t)$ is irrelevant to the consumption choice and could indeed be at any level.

Now, if we let

$$k = \frac{\mu(1+\alpha M)}{1+\mu\alpha M}, \qquad 0 \leq k \leq 1, \tag{19}$$

then it is straightforward to derive from (18) that the individual's consumption choice can be characterised by

$$\phi'(z) = c + [kMD'(Mz) + (1-k)t]. \tag{20}$$

The parameter k is a combined parameter capturing both the individual's propensity to act morally μ, and the level of altruistic concern for others' utility α. This parameter can be thought of as the overall level of 'virtue' individuals may exhibit, where $0 \leq k \leq 1$. Examining (20), we see that if $k = 1$ the individual will choose a level of consumption equal to the socially optimal level $z = \hat{z} = c + MD'(Mz)$. On the other hand, if $k = 0$, the individual will choose a consumption level of the dirty good equal to the conventional choice, $z = \tilde{z}(t) = c + t$. Of course the level of k can be anywhere between 0 and 1 and the level of z will vary accordingly between $\tilde{z}(t)$ and $\bar{z}$. Figure 1 further illustrates the different levels of consumption for a level of k such that $0 < k < 1$.

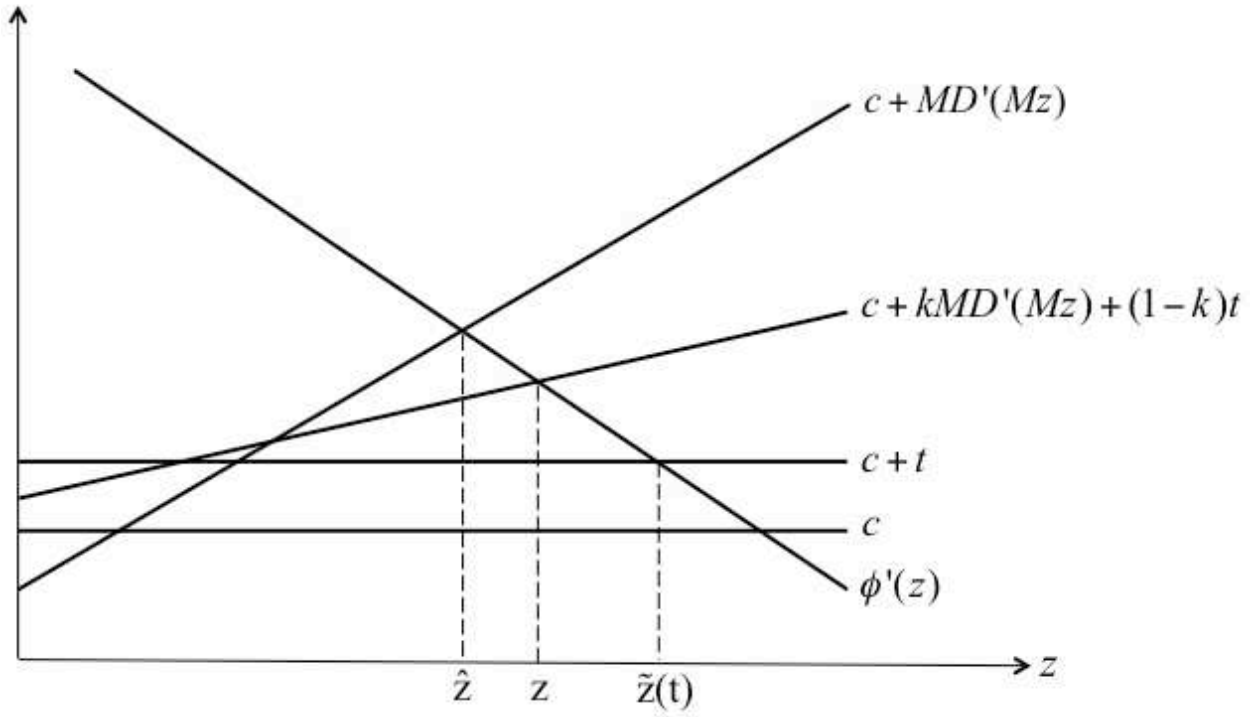


Figure 1

Next, let us take a closer look at the determinants of k. Suppose the individual has no altruistic concern for others' utility ($\alpha = 0$). Then it follows that $k = \mu$. This means that for an individual with no altruistic concern, behaviour only depends on the their propensity to act morally. However, if $\mu = 1$, the individual will still cut down consumption of the dirty good to $\hat{z}$ regardless of the level of α. It follows that altruistic concern for others' utility is not a necessary component of this type of behaviour.

Now suppose that the individual has no propensity to act morally ($\mu = 0$). Then, it is straightforward to see that $k = 0$ regardless of the value of α. Effectively, if $\mu = 0$, we are back to the case where people choose consumption of dirty good by maximising their welfare function which means that individuals will choose $\tilde{z}(t)$ regardless of their level of altruistic concern. Using these insights, we can set out the next proposition.

Proposition 2 *Altruistic concern for others' utility is neither necessary nor sufficient to induce people to cut back consumption of the dirty good, but a propensity to act morally is both necessary and sufficient.*

Proof: *From (20) we know that z is a decreasing function of k for any $t < \hat{t}$. From (19) we can further determine that if $\mu = 1 \rightarrow k = 1$, if $\mu = 0 \rightarrow k = 0, and\ \frac{\partial k}{\partial \mu} = \frac{1+\alpha M}{(1+\mu\alpha M)^2} > 0$ for any value*

of α. Therefore α is neither necessary nor sufficient, but μ is both necessary and sufficient to induce a decrease in z.

However, let us now look at the case where individuals have some, but not full propensity to act morally ($0 < \mu < 1$). Then k is an increasing function of α. This means that the level of altruistic concern can affect the consumption choice even though people recognise that their consumption of the dirty good has no effect on total emissions and anybody else's welfare. This is because, in calculating the hypothetical moral value of the action, people take into account their level of altruistic concern for others' utility and therefore the impact their choice will have on everybody else's utility, but the associated utility cost of the consumption choice is independent of the level of altruistic concern. Yet we should also note that, even if $\alpha = 1$, as long as $\mu < 1$, an individual will never cut down consumption of the dirty good all the way to $\hat{z}$. This leads us to the next proposition:

Proposition 3 *If individuals have some, but not full propensity to act morally, an individual with a higher level of altruistic concern for others will reduce consumption of the dirty good if the government sets the tax on the dirty good too low, but never all the way to the socially optimal level.*

Proof: *From (19) we can derive that $\frac{\partial k}{\partial \alpha} = \frac{\mu M(1-\mu)}{(1+\mu\alpha M)^2} > 0$ for any $0 < \mu < 1$. Therefore k is an increasing function of α. Furthermore, we know from (20) that z is a decreasing function of k for any $t < \hat{t}$. Therefore an increase in α will lead to a reduction in the consumption of the dirty good for any $0 < \mu < 1$. However, to achieve $z = \hat{z}$, we require $k = 1$. As per the definition of k in (19), this is only the case when $\mu = 1$ and therefore $z = \hat{z}$ can never be achieved for any $0 < \mu < 1$ regardless of the value of α.*

Next, suppose that the individual has some level of altruistic concern and some propensity to act morally (i.e. $\alpha > 0$, and $\mu > 0$). Then we can see that k is also an increasing function of the size of the population, M. It shows that the larger the population of people affected, the more people will cut back their consumption of the dirty good. The size of the population does not just influence the level of k though. Taking a closer look at (20) it is easy to see that for individuals with some propensity to act morally (i.e. $k > 0$), there is another channel through which an increase in M will cause individuals to cut back their consumption of the dirty good. An increase in M increases the social marginal damage of

consumption and therefore reduces the socially optimal level of consumption $\hat{z}$.[26] The population size M impacts the social marginal damage because, (a) it increases the number of people affected, and (b) it increases the total level of emissions which leads to increasing marginal damage since damage is convex in z (ie $D'' > 0$). We can now state the following general proposition.

Proposition 4 *If the government fails to set the tax that would be optimal given the usual economic behaviour and the usual conception as to what constitutes welfare, then, to the extent that individuals have some propensity to act morally, private action will to some extent compensate for the lack of government action and drive consumption of the dirty good down towards the optimal level.*

Proof: *For any* $0 < \mu \leq 1$ *we have* $0 < k \leq 1$ *as per the definition of* k *in (19). Using this and comparing (20) with (5) and (8), it is straightforward to see that we must have* $\phi'(\hat{z}) \geq \phi'(z) > \phi'[\tilde{z}(t)]$ *and therefore we also know that* $\hat{z} \leq z < \tilde{z}(t)$.

We further know that although people may behave according to a different set of rules, individuals' welfare is still determined by the altruistic welfare function as defined in (10). The government recognises this and therefore still uses the same social welfare function as given in (11) to determine the socially optimal allocation of welfare. Similarly, the social utility function $S(.)$ is also still the same as described in (6). It is then straightforward to see that the socially optimal tax level for the government is still the standard Pigovian tax rate $\hat{t}$, the same that would be optimal if the population had no propensity to act morally (i.e. $\mu = 0$ for all types). Indeed, by examining (20) while plugging in the Pigovian tax as shown in (9), we can easily determine that each individual will choose consumption level $\hat{z}$ regardless of the level of k. This means with a tax rate equal to $\hat{t}$ we can still achieve the first-best solution where everyone consumes the same amount of the dirty good and the damage is fully internalised. This leads us to the final proposition.

Proposition 5 *Although individuals may to some extent compensate for the government's failure to set the optimal tax, this does not imply that the government should not set the optimal tax.*

Proof: *From (20) we can derive that by increasing* t *to* $\hat{t} = MD'(M\hat{z})$ *everyone will consume* $\hat{z}$ *as characterised by (8). This raises social welfare because (a) aggregate consumption of the dirty good is socially optimal and (b) consumption is equalised across consumers.*

[26] This impact of M on the socially optimal level of consumption is of course not an exclusive feature of moral behaviour as modelled here, but can also be observed in the characterisation of the socially optimal level of consumption as shown in (8).

Initially we assumed that the government set the tax on the dirty good below the optimal level ($t < \hat{t}$). Now suppose that the government actually sets the tax rate too high, so $t > \hat{t}$ and therefore $\tilde{z}(t) < \hat{z}$. Because the initial level of consumption is now lower than the socially optimal level, an individual with some propensity to act morally will actually increase their consumption of the dirty good in order to bring it closer to the socially optimal level. The benchmark for an individual with some propensity to act morally is always the socially optimal level of consumption and depending on the weight they give to the moral value, they will choose a consumption level of the dirty good that moves closer to $\hat{z}$. This point illustrates that the propensity to act morally does not necessarily imply a reduction in the consumption of an environmentally harmful good. Rather, an individual with some propensity to act morally will change consumption in order to get closer to the social optimum in the absence of the correct tax on the dirty good.

4 Discussion

The aim of this paper was to develop an alternative model of moral behaviour and altruism in an environmental context and evaluate what implications such behaviour may have for environmental policy. We have shown that when people have some propensity to act morally they will cut back consumption of an environmentally harmful good. At the same time, altruistic concern for the utility of others can contribute to the amount that individuals cut consumption, but only if they also have some propensity to act morally. The optimal tax on the dirty good remains the same as under standard theory and the first-best solution can be achieved. A key reason for this is that, even though some individuals may cut consumption towards to socially optimal level for moral reasons, we are still only dealing with market failure, which the Pigovian tax can correct for all individuals. However, this result hinges on individuals correctly making the (sophisticated) calculation of the moral value when acting in this genuinely altruistic way. We mentioned earlier that the individual recognises both the link to the government budget constraint as well as the impact on total emissions when making this calculation. Suppose instead that the individual fails to take account of the connection to the government budget constraint. It is then straightforward to show that an individual with full propensity to act morally ($\mu = 1 \rightarrow k = 1$) will actually overcompensate and consume less than the socially optimal level. At the same time, a tax at the Pigovian level means that every individual with some propensity to act morally ($k > 0$) will consume less than the socially optimal level. Indeed, we are no longer able to find a tax rate that induces the first-best solution where everyone consumes the same amount of the dirty good. This is

because we are now dealing with a behavioural failure as well as a market failure, which cannot both be corrected with a single instrument. We would need a separate mechanism to correct the behavioural failure.[27]

Another key assumption in our model is that the propensity to act morally for each type of individual is a given and static parameter. There are two questions that develop from this observation. First, where does the propensity to act morally come from? And second, can the propensity to act morally be influenced by extrinsic incentives on the dirty good? The latter addresses whether the propensity to act morally can be crowded-in or crowded-out by extrinsic incentives. Crowding of intrinsic motivation has received a lot of attention in the literature on pro-social behaviour due to its potential policy relevance.[28] Extensions to our paper could investigate how the model can be adapted to give further insights on this issue.

References

Altemeyer-Bartscher M, Markandya A, Rübbelke DTG (2013) International Side-Payments to Improve Global Public Good Provision when Transfers are Refinanced through a Tax on Local and Global Externalities. Int Econ J

Andreoni J (1988) Privately Provided Public Goods in a Large Economy: The Limits of Altruism. J Public Econ 35:57–73

Andreoni J (1989) Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. J Polit Econ 97:1447–1458

Andreoni J (1990) Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. Econ J 100:464–477

Archibald GC, Donaldson D (1976) Non-Paternalism and the Basic Theorems of Welfare Economics. Can J Econ 9:492–507

Arrow KJ (1986) Rationality of Self and Others in an Economic System. J Bus 59:385–399

[27] See Shogren and Taylor (2008) for a discussion on the existence of both market and behavioural failures in an environmental economics context.

[28] For example Frey (1997), Frey (1999), Bénabou and Tirole (2006), Nyborg and Rege (2003), Nyborg et al. (2006) and Brekke et al. (2003) among others.

Becker GS (1974) A Theory of Social Interactions. J Polit Econ 82:1063–1093

Becker GS (1981) Altruism in the Family and Selfishness in the Market Place. Economica 48:1–15

Bénabou R, Tirole J (2006) Incentives and Prosocial Behavior. Am Econ Rev Econ Rev 96:1652–1678

Bergstrom T, Blume L, Varian H (1986) On the Private Provision of Public Goods. J Public Econ 29:25–49

Brekke KA, Kverndokk S, Nyborg K (2003) An Economic Model of Moral Motivation. J Public Econ 87:1967–1983

Cornes R, Sandler T (1984) The Theory of Public Goods: Non-Nash Behaviour. J Public Econ 23:367–379

Ellingsen T, Johannesson M (2008) Pride and Prejudice: The Human Side of Incentive Theory. Am Econ Rev 98:990–1008

Fontaine P (2008) Altruism, History of the Concept. In: Durlauf SN, Blume LE (eds) The New Palgrave Dictionary of Economics, 2nd edn. Palgrave Macmillan, London

Frey BS (1997) Not Just for The Money : An Economic Theory of Personal Motivation. Edward Elgar, Cheltenham

Frey BS (1999) Morality and Rationality in Environmental Policy. J Consum Policy 22:395–417

Galarraga I, Markandya A (2009) Climate Change and Its Socioeconomic Importance. Basque Centre for Climate Change

Hammond PJ (1987) Altruism. In: Eatwell J, Milgate M, Newman P (eds) The New Palgrave: A Dictionary of Economics, 1st edn. Palgrave Macmillan, London

Harsanyi JC (1955) Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. J Polit Econ 63:309–321

Johansson O (1997) Optimal Pigovian Taxes under Altruism. Land Econ 73:297–308

Kant I (1875) Grounding for the Metaphysics of Morals. 3rd edn. Hackett, Indianapolis

Kennett DA (1980) Altruism and Economic Behavior, I: Developments in the Theory of Public and Private Redistribution. Am J Econ Sociol 39:183–198

Laffont J-J (1975) Macroeconomic Constraints, Economic Efficiency and Ethics: An Introduction to Kantian Economics. Economica 42:430–437

Longo A, Hoyos D, Markandya A (2012) Willingness to Pay for Ancillary Benefits of Climate Change Mitigation. Environ Resour Econ 51:119–140

Markandya A, Rübbelke DTG (2004) Ancillary Benefits of Climate Policy. J Econ Stat 224:488–503

Markandya A, Rübbelke DTG (2012) Impure Public Technologies and Environmental Policy. J Econ Stud 39:128–143

Nyborg K, Howarth RB, Brekke KA (2006) Green Consumers and Public Policy: On Socially Contingent Moral Motivation. Resour Energy Econ 28:351–366

Nyborg K, Rege M (2003) Does Public Policy Crowd out Private Contributions to Public Goods. Public Choice 115:397–418

Sen AK (1977) Rational Fools: A Critique of the Behavioral Foundations of Economic Theory. Philos Public Aff 6:317–344

Shogren JF, Taylor LO (2008) On Behavioral-Environmental Economics. Rev Environ Econ Policy 2:26–44

Stern N (2007) Stern Review of the Economics of Climate Change. HM Treasury, London

Sugden R (1982) On the Economics of Philanthropy. Econ J 92:341–350

Thaler R, Shefrin HM (1981) An Economic Theory of Self-Control. J Polit Econ 89:392–406

Ulph D (2006) A Theory of Green Consumerism. (mimeo), School of Economics & Finance, University of St Andrews

White MD (2003) Can Homo Economicus follow Kant's Categorical Imperative? J Socio Econ 33:89–106