CESifo WORKING PAPERS

7880
2019

October 2019

Dynamic Games of Common-Property Resource Exploitation When Self-Image Matters

Ngo Van Long

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
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7880*
Category 9: Resource and Environment Economics

Dynamic Games of Common-Property Resource Exploitation When Self-Image Matters

Abstract

The purpose of this paper is to model the influence of Kantian moral scruples in a dynamic environment. Our objectives are two-fold. Firstly, we investigate how a Nash equilibrium among agents who have moral scruples may ensure that the exploitation of a common property renewable resource is Pareto efficient at every point of time. Secondly, we outline a prototype model that shows, in an overlapping generation framework, how a community's sense of morality may evolve over time and identifies conditions under which the community may reach a steady state level of morality in which everyone is perfectly Kantian.

JEL-Codes: C710, D620, D710.

Keywords: tragedy of the commons, dynamic games, Nash equilibrium, self-image, categorical imperative.

Ngo Van Long
Department of Economics
McGill University
Canada - H3A 2T7, Montreal, Québec
ngo.long@mcgill.ca

17 September 2019

1 Introduction

Many economic ills can be attributed to the lack of incentives for agents to cooperate. For example, it is well recognized that contributions to public goods tend to be undersupplied and exploitation of public-owned assets tend to be excessive (Gordon, 1954, Hardin, 1968). A most serious challenge facing the world is the danger of climate change, which is difficult to combat because the quality of global enviromental resources is a public good. The prevailing incentives to free-ride render fruitless the United Nations' efforts of implementing the Kyoto Protocol. (For dynamic games of climate change, see, among others, Wirl (1995, 2011), Wirl and Dockner (1995), Yang (2003), Deissenberg and Leitmann (2004), Grafton et al. (2017); see Long (2010) for a survey.)

However, there are instances where common property resources are properly managed, as well documented by Ostrom (1990). A mechanism which ensures reasonable cooperation by private agents is the enforcement of social norms. Economic models of the working of social norms typically include a group of agents that punish violators (Sethi and Somanathan, 1996; Breton et al., 2010).[1] Recently, there are models in which norms are respected without an explicit punishment mechanism (Brekke et al. 2003; Roemer, 2010, 2015; Wirl, 2011; Long, 2016, 2017). These authors, following the footsteps of Laffont (1975), emphasize the fact that many economic agents, being motivated by moral scruples, feel the need to act in accordance with moral principles such as the categorical imperative (Kant, 1785).[2] The modeling of the influence of morality on economic behavior differs among economists. Following the tradition of Arrow (1973), Sen (1977), and Laffont (1975), the recent papers by Roemer (2010, 2015), Long (2016, 2017), Grafton et al. (2017) rely on the concept of Kantian equilibrium originated from Laffont (1975). This equilibrium concept departs from the Nash equilibrium concept by supposing that that agents do not behave in the Nashian way: they do not take the actions of others as given.[3] Using an alternative approach, the papers by Brekke et al. (2003) and Wirl (2011), following earlier works by Fehr and Schmidt (1999), Bolton and Ockenfelds (2000),

[1] There is a large literature on social norms in a market environment. For some recent contributions, see Deissenberg and Peguin-Feissole (2006), Dasgupta et al. (2016), Ulph and Ulph (2017).

[2] Kant (1875) wrote that "There is only one categorical imperative and it is this: Act only on the maxim by which you can at the same time will that it should become a universal law." Translated by Hill &Zweig (2002, p. 222).

[3] Binmore (2005) argued against the Kantian approach. A counter-argument was offered in Grafton et al. (2017).

and Charness and Rabin (2002), keep the Nashian framework but endow agents with a sense of morality, such that deviations from the Kantian ideal imposes a quadratic loss of one's self-respect.[4]

Most of the Kant-based models mentioned in the preceding paragraph (with the exception of Wirl, 2011) restrict attention to a static framework, i.e., there is no stock dynamics. The purpose of this paper is to model explicitly the influence of Kantian moral scruples in a dynamic environment. Our objectives are two-fold. Firstly, we investigate how a Nash equilibrium among agents who have moral scruples may ensure that the exploitation of a common property renewable resource is Pareto efficient at every point of time. Secondly, we outline a prototype model that shows, in an overlapping generation framework, how a community's sense of morality may evolve over time and identifies conditions under which the community may reach a steady state level of morality in which everyone is perfectly Kantian.[5]

2 Related literature

Many generations of economics students have been told that a central result of economic theory is that if all agents are self-interested maximizers of their own material wellbeing, the outcome of a competitive equilibrium is Pareto efficient. This result is of course subject to a number of qualifications, but these are quite often relegated to footnotes. Many authors have attributed to Adam Smith the vision of a miraculous achievement of the price mechanism, ignoring the fact that Smith himself held a much more nuanced view. In fact, in *The Wealth of Nations*, Smith (1776) pointed out that there are cases where the pursuit of self-interest ought to be severely restrained.[6] Moreover, Adam Smith

[4] Wirl (2011) assumes the co-existence of green consumers and brown consumers, who behave in a Nashian fashion in a dynamic game of global warming.

[5] For an alternative approach without overlapping generations, see Alger and Weibull (2016).

[6] On banking regulation, Smith (1776, p. 308) wrote that "Such regulations may, no doubt, be considered as in some respect a violation of natural liberty. But those exertions of natural liberty of a few individuals, which might endanger the security of the whole society are, and ought to be, restrained by the laws of all governments." On moral hazard, he noted that the interest of agents are not aligned with that of the principals: "The directors of such companies, however, being the managers rather of other's money than of their own, it cannot be well expected, that they should watch over it with the same anxious vigilance with which the partners in a private copartnery frequently watch over their own...Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such a company." (Smith, 1776, Book 5, Chapter 1, p. 700.)

never said that economic agents are solely interested in personal gains. In *The Theory of Moral Sentiments*, Smith (1790) emphasized the crucial importance of the respect for social norms and moral duties. He wrote:

> "*Upon the tolerable observance of these duties, depends the very existence of human society, which would crumble into nothing if mankind were not generally impressed with a reverence for those important rules of conduct.*"[7]

Smith (1790) discussed at length the role of natural sympathies in human activities and the human urge to be accepted as a respectable moral being. According to Smith, humans desire to merit the approval of other members of their community: we judge our actions as we think others would judge them. Through interaction with those around us, we learn "*general rules concerning what is fit and proper either to be done or to be avoided.*" [8] Moreover, humans desire not only to be praised, but to be truly deserving of praise. They feel happiness by acting in a way which merits the self-approval which comes from knowing that they have acted according to the standard of "*the impartial and well-informed spectator... within the breast.*"[9]

In the last few decades, Smith's views have been vindicated by research in experimental economics; see e.g. Dawes and Thaler (1988), Bolle and Ockenfels (1990), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Camerer (2003), Camerer and Fehr (2006), Andreoni et al. (2008). Referring to Adam Smith's *The Theory of Moral Sentiments,* Vernon Smith (2003, p. 466) elaborates on the important message of the 18th century Scottish philosophers such Smith and Hume:

"Research in economic psychology has prominently reported examples where "fairness" considerations are said to contradict the rationality assumptions of the standard socioeconomic science model (SSSM). But experimental economics have reported mixed results on rationality: people are often better (e.g., in two-person anonymous interactions), in agreement with (e.g., in flow supply and demand markets), or worse (e.g., in asset trading), in achieving gains for themselves and others than is predicted by rational analysis. Patterns of these contradictions and confirmations provide important clues to

[7]Smith, *The Theory of Moral Sentiments*,1790, Part III, Chapter V, p. 190.

[8]The Theory of Moral Sentiments, edited by A.L. Macfie and D.D. Raphael (1976), The Glasgow Edition of the Works and Correspondence of Adam Smith, Oxford University Press. Book III, Chapter 4, part 7, page 159.

[9]The Theory of Moral Sentiments, Book III, Ch2, p. 130.

the implicit rules or norms that people may follow, and can motivate new theoretical hypotheses for examination in both the field and the laboratory. The pattern of results greatly modifies the prevailing, and I believe misguided, rational SSSM, and richly modernizes the unadulterated message of the Scottish philosophers."

The importance of self-image has been emphasized in the economic literature. Recent contributions to this stream of literature include Brekke et al. (2003), Akerlof and Kranton (2005), and Elster (2017). Outside of economics, self-image has been a key theme in moral philosophy and in psychology. Indeed, Rabbi Hillel, a first century sage, posed the following questions:

> "*If I am not for myself, then who is for me? And if I am not for others, then who am I? If not now, when?*"[10]

While the concern for self-image can be a source of good, the failure of not being seen as having lived up to one's ideal can be a source of misery. In Jean-Paul Sartre's 1947 play, titled Huis Clos, the main character, Garcin, finally reached a devastating awareness:

> "*Tous ces regards qui me mangent...Alors, c'est ça l'enfer. Je n'aurais jamais cru... Vous vous rappelez: le soufre, le bûcher, le gril...Ah! quelle plaisanterie. Pas besoin de gril: l'enfer, c'est les Autres.*"[11]

However, the self-image (as reflected in the eyes of others) that Garcin was obsessed with should be only a first rung in the moral ladder. According to Adam Smith, a higher rung is reached when the eyes of others no longer matter. One then applies the standard of "*the impartial and well-informed spectator... within the breast.*" Smith's view echoes Confucius' doctrine of shame as a guiding principle for moral behaviour, as recorded in the Analects:

> "*Guide them with government orders, regulate them with penalties, and the people will seek to evade the law and be without shame. Guide them with virtue, regulate them with ritual, and they will have a sense of shame and become upright.*"[12]

[10] Cited in Arrow, K. J. (1974), *The Limits of Organization.* New York: W.W. Norton.

[11] "All those looks that eat me ... So that is hell. I never thought ... You remember: the sulfur, the stake, the grill ... Ah! what a joke. No need for a grill: Hell is the Others." Scene 5, Huis Clos, by Sartre (1947).

[12] Cited in Bowles (2016, p. 11).

Complementing the growing literature on the need to modify the standard model of economic behaviour to account for humans' concern for morality, this paper constructs a model of a dynamic game of common property resource exploitation in which agents care not only about their material wellbeing, but also about their self-image. I show that, despite the well-known incentives to free ride when agents exploit a common asset, a social optimum may be within reach provided that agents have a precise idea of what actions would be prescribed by Kantian ethics, and they feel bad if their own actions do not match the moral ideal.

3 Modelling individual tradeoff between self-image and material wellbeing

For exposition, this section restricts attention to a static framework. We assume that individuals care about their material wellbeing, denoted by M_i, while at the same time, they attach a value v_i to their self-image. Their self-image suffers if they under-contribute to a public good, or if they over-exploit a common property asset.

3.1 Specification of the self-image function and the material wellbeing function

In the case of exploitation of a common property resource, such as a pasture, the economic literature typically supposes that individuals have a tendency to over-exploit, i.e., their demands are excessive. Let $e_i \geq 0$ denote the individual's actual level of exploitation, and e_i^K the level of exploitation that the Kantian social norms would dictate. Then $e_i - e_i^K$ is the individual's extent of excessive demand (excessive exploitation). We assume that exploitation in excess of the social norms causes a loss of self-image equal to $\theta_i \times \left(e_i - e_i^K\right) \times \sigma$, where $\sigma > 0$ is a scale parameter that reflects the (objective) severity of the effect of the over-exploitation, and $\theta_i > 0$ is the individual's coefficient of the (subjective) loss of self-esteem associated with excessive exploitation.

For tractability, we assume that an individual's self-image function, denoted by v_i, takes the following simple form

$$v_i = A_i - \theta_i^u \times \max\left\{0, (e - e_i^K)\sigma\right\} \tag{1}$$

where A_i is a positive constant.

Turning to the material payoff M_i of an individual i we assume that it consists his "harvest" q_i from the common-property resource, net of the effort cost of harvesting $g_i(e_i)$.

The size of his harvest may depend not only on his exploitation level e_i but also on the aggregate level of exploitation, because of over-crowding externalities. We write

$$q_i = f_i(e_i, E),$$

with $\partial f_i/\partial e_i > 0$ and $\partial f_i/\partial E < 0$, where

$$E \equiv \sum_{i=1}^{n} e_i.$$

Let us define

$$E_{-i} = E - e_i.$$

The material wellbeing of individual i is

$$M_i = f_i(e_i, E_{-i} + e_i) - g_i(e_i). \tag{2}$$

Individual i chooses $e_i \geq 0$ to maximize his payoff, defined as the sum of his material wellbeing and his self-image:

$$U_i = M_i + v_i. \tag{3}$$

In this maximization problem, he takes E_{-i} as given. In other words, here we use the concept of Nash equilibrium.

3.2 A digression: specification of the individual-specific Kantian ideals

If all individuals have identical charactersistics and circumstances, as is assumed in the model formulated in Laffont (1975), one may suppose that $e_i^K = e^K$ for all i, and that

e^K is the value of e that would maximize the material wellbeing of a representative individual. In the case of homogeneous individuals, clearly there are no differences between the Kantian levels and the optimum that a Benthamite utilitarian social planner would want to achieve. Let us turn now to the case where individuals are heterogeneous. What would be a plausible specification of individual-specific duties?

Due to space limitation, it is not possible to offer here a detailed discussion of this important issue. Let me simply mention two important approaches that have been proposed to address this subject. The first approach is that of Bilodeau and Gravel (2004). They argue that "to treat everyone similarly, a maxim must prescribe to everyone actions that are in some sense equivalent" (p. 647). They propose the concept of morally equivalent actions by introducing a system of universalization, i.e., a binary relation that compares any two actions (possibly undertaken by two persons with different characteristics) and determines whether they are morally equivalent. They insist that a Kantian maxim, if obeyed by all, must "yield everyone's most preferred outcome if everyone else is constrained to play a morally equivalent strategy" (p. 647). Bilodeau and Gravel (2004) show that, in the setting of voluntary contributions to a public good, if a system of universalization satisfies certain axioms, any Kantian maxim that is consistent with it is necessarily Pareto efficient.[13]

The second approach is more operational and is due to Roemer (2010, 2015). Roemer (2010) defines a Kantian equilibrium for a class of games where each individual can only take a single action, for which he can contemplate alternative outcomes that would result from scaling his action level up or down. We can shed light on Roemer's approach by considering the following example.

Consider a game of exploitation of a common property resource (such as a common pasture). Consider a small community in which there are n households. Let e_i be the the number of goats that household i keeps. Assume that the final output, say goat milk, is obtained by letting the goats (an input) graze on the common pasture (a second input). The community's agregate output of milk is $Q = \xi F(E)$, where $E = \sum e_i$, and $\xi > 0$ is the quality of the pasture. Assume that $F(0) = 0$, $F'(0) > 0$, and $F''(E) < 0$. The output of milk per goat is Q/E, and therefore the quantity of milk collected by household i is $e_i Q/E$. Assume that, due to different levels of skills among households, the effort cost

[13] Technically, the axioms involve two requirements on a system of universalization: tightness and differentiability (p. 648).

incurred by household i in keeping e_i goats is given by

$$g_i(e_i) = \beta_i g(e_i),$$

where $g(.)$ is a strictly convex and increasing function defined for all $e_i \geq 0$, with $g(0) = 0 = g'(0)$. Without loss of generality, assume $1 = \beta_1 \leq \beta_2 \leq \beta_3 ... \leq \beta_n$. What is the Kantian number of goats that household i should keep? Following Roemer (2010), let us define a Kantian allocation of input levels as a strictly positive vector $(e_1^K, e_2^K, e_3^K, ..., e_n^K)$ such that for each household i, if it were to modify e_i^K by applying a scaling factor $\lambda > 0$ (so that its exploitation would be changed to λe_i^K), it would find that, for all λ such that $0 < \lambda \neq 1$, its material wellbeing would fall, on the assumption that all other households would change their e_j^K by the same factor λ. This thought experiment reflects the Kantian dictum that when one contemplates doing something, one should ask oneself: how would I like it if everyone else behaved in the same way?

Formally then, in our common pasture example, an allocation $(e_1^K, e_2^K, e_3^K, ..., e_n^K)$ is a Kantian equilibrium (in thought) if and only if

$$1 = \arg\max_{\lambda > 0} \frac{\lambda e_i^K \xi F(\lambda e_i^K + \lambda E_{-i}^K)}{\lambda e_i^K + \lambda E_{-i}^K} - \beta_i g(\lambda e_i^K).$$

Let the material payoff of household i be denote by M_i. Let

$$M_i(\lambda) \equiv \frac{\lambda e_i^K \xi F(\lambda e_i^K + \lambda E_{-i}^K)}{\lambda e_i^K + \lambda E_{-i}^K} - \beta_i g(\lambda e_i^K).$$

Differentiating M_i with respect to λ, we get the first order equation

$$\frac{e_i^K}{E^K} \xi F'(\lambda E^K) E^K - \beta_i g'(\lambda e_i^K) e_i^K = 0 \text{ for } i = 1, 2, ..., n.$$

Evaluated at $\lambda = 1$, we get the condition that characterizes the Kantian allocation:

$$\xi F'(E^K) = \beta_i g'(e_i^K). \tag{4}$$

Remark: Equation (4) implies that the Kantian input allocation is efficient: the marginal social product of the total input is equated to the marginal cost for each agent. Condition

(4) that characterizes the Kantian equilibrium allocation in this model (where utility is *linear* in consumption) is also the condition that characterizes the optimal allocation under the standard utilitarian objective of maximizing the *non-weighted* sum of individuals' utilities. (However, this is not always the case; as shown in the Appendix, the Kantian equilibrium allocation in a public good model (where utility is *non-linear* in the public good) can be obtained only by maximizing a *weighted* sum of individuals' utilities.)

We can next compute e_i^K and E^K as follows

$$e_i^K = g'^{-1}\left(\frac{\xi F'(E^K)}{\beta_i}\right). \tag{5}$$

Summing (5) over $i = 1, 2, ..., n$, we get

$$E^K = \sum_i g'^{-1}\left(\frac{\xi F'(E^K)}{\beta_i}\right). \tag{6}$$

Since F' is decreasing and g'^{-1} is increasing, the right-hand side of equation (6) is decreasing E. The left-hand side is linear and increasing in E. Therefore there exists a unique $E^K > 0$. Next, we can calculate e_i^K using (5). It can be shown that at the Kantian equilibrium, weaker households (those with a high value β_i) keep fewer goats than stronger households and enjoy a lower level of material wellbeing.

4 Renewable resource exploitation by image-conscious agents

In this section, we show how the tragedy of the commons can be avoided if agents are endowed with a sufficiently strong desire to maintain a good self-image. For simplicity, let us assume that individuals are homogeneous. To fix ideas, we use a model of common access fishery. The "common access fishery model" has been interpreted more broadly to mean a model of rivalrous exploitation of any kind of renewable resource.

Let $R(t)$ denote the resource stock, and $x_i(t)$ denote agent i's rate of exploitation. Assume that

$$\dot{R}(t) = G(R_t) - \sum_{i=1}^{n} x_i(t),$$

where $G(R)$ is the natural growth function, with $G(0) = 0$, $G'(0) > 0$, and $G''(R) \leq 0$.

Let us assume that agent i's instantaneous material wellbeing is simply an increasing and concave function of his rate of exploitation. We denote this function by $M_i(x_i(t))$. Agents live for ever and discount their future utility at the rate $\rho > 0$. The life-time payoff of agent i, starting from any time $\tau \geq 0$ is

$$P_i(\tau) = \int_\tau^\infty e^{-\rho(t-\tau)} M_i(x_i(t))dt.$$

4.1 Cooperative solution when individuals are homogeneous

When individuals are homogeneous, the cooperative solution is straightforward. It is as if there were a social planner who would maximize the life-time utility of an infinitety-lived representative individual. (One can think of this agent as a family line). The planner solves the following optimal control problem: choose the extraction rate per capita to maximize the discounted life-time material wellbeing of the representative agent:

$$\max_{x(t)\geq 0} \int_0^\infty e^{-\rho t} M(x(t))dt,$$

subject to

$$\dot{R}(t) = G(R(t)) - nx(t),$$

with

$$\lim_{t\to\infty} R(t) \geq 0.$$

The above optimal control problem can also be solved using the Hamilton-Jacobi-Bellman equation. Let $V_P(.)$ denote the value function of the planner (here, the subscript P denote the planner). The Hamilton-Jacobi-Bellman (HJB) equation is

$$\rho V_P(R_t) = \max_{x\geq 0} \left[M(x_t) + (G(R_t) - nx_t)V_P'(R_t) \right]. \tag{7}$$

The first order condition is

$$M'(x_t) - nV_P'(R_t) = 0.$$

This yields x_t as a function of R_t

$$x_t = \phi\left(nV_P'(R_t)\right),$$

where

$$\phi(.) = (M')^{-1}.$$

Thus we obtain the following first order differential equation[14]

$$\rho V_P(R) = M\left(\phi\left(nV_P'(R)\right)\right) + \left[G(R) - n\phi\left(nV_P'(R)\right)\right] V_P'(R).$$

Define

$$x^K(R) \equiv \phi\left(nV_P'(R)\right). \tag{8}$$

Then we obtain a first order differential equation relating V_P to V_p' :

$$\rho V_P(R) = M(x^K(R)) + \left[G(R) - nx^K(R)\right] V_P'(R).$$

Using the usual transversality condition, this equation can be solved to yield the value function and hence the optimal harvest rule.

Example 1:

Assume that the growth function of the biomass is

$$G(R) = R^\gamma - \delta R,\ 0 < \gamma < 1,$$

and the material wellbeing function is unbounded above

$$M(x) = \frac{x^{1-\gamma}}{1-\gamma}, \text{where } \gamma \in (0,1).$$

Denote by $V_P(R)$ the social planner's value function. The HJB equation is

$$\rho V_P(R) = \max_{x \geq 0}\left\{\frac{x^{1-\gamma}}{1-\gamma} + V_P'(R)\left[R^\gamma - \delta R - nx\right]\right\}.$$

[14] We seek a solution $V_P(R)$ such that an appropriate transversality condition is met, e.g. $\lim_{t\to\infty} e^{-\rho t} V_P(R(t)) = 0$. See Dockner et al. (2000).

The first oder condition is

$$x^{-\gamma} = nV_P'(R).$$

Try the value function

$$V_P(R) = A + B\frac{R^{1-\gamma}}{1-\gamma},$$

where A and B are to be determined. Then

$$V_P' = BR^{-\gamma}.$$

The first order condition then gives

$$x^{-\gamma} = nBR^{-\gamma},$$

i.e., the harvesting rule is linear:

$$x = (nB)^{-1/\gamma}R.$$

Substituting this into the HJB equation to get

$$\rho A + \rho B\frac{R^{1-\gamma}}{1-\gamma} = \frac{(nB)^{\frac{(\gamma-1)}{\gamma}}R^{1-\gamma}}{1-\gamma} + B - \delta R^{1-\gamma} - (nB)^{\frac{(\gamma-1)}{\gamma}}R^{1-\gamma}.$$

Then, since the above equation must hold for all $R > 0$, the coefficients of the terms involving $R^{1-\gamma}$ must add up to zero, i.e.,

$$(nB)^{-1/\gamma} = \frac{\rho + \delta(1-\gamma)}{n\gamma} > 0.$$

Thus the Kantian rate of exploitation is

$$x^K(R) \equiv \frac{\rho + \delta(1-\gamma)}{n\gamma}R.$$

Example 2

Let

$$G(R) = \kappa R - R^\eta \text{ where } \eta > 1 \text{ and } \kappa > 0,$$

and assume the utility function is bounded above:

$$M(x) = \frac{x^{1-\eta}}{1-\eta} \text{ where } \eta > 1.$$

The planner's HJB equation is

$$\rho V_P(R) = \max_{x \geq 0} \left\{ \frac{x^{1-\eta}}{1-\eta} + V_P'(R) \left[\kappa R - R^\eta - nx \right] \right\}.$$

Assume that $\rho > \kappa - 1$. The first order condition is

$$x^{-\eta} = nV_P'(R).$$

We conjecture the following value function

$$V_P(R) = A + \frac{DR^{1-\eta}}{1-\eta},$$

where A and D are to be determined. Then

$$V_P'(R) = DR^{-\eta},$$

$$x = (nD)^{-\frac{1}{\eta}} R.$$

Plugging this exploitation rule into the HJB equation, we obtain

$$\rho A + \rho \frac{DR^{1-\eta}}{1-\eta} = \frac{(nD)^{(\eta-1)/\eta} R^{1-\eta}}{1-\eta} + \kappa DR^{1-\eta} - D - (nD)^{(\eta-1)/\eta} R^{1-\eta},$$

which yields

$$(nD)^{-1/\eta} = \frac{\rho + \kappa(\eta - 1)}{n\eta} > 0.$$

Thus the Kantian rate of exploitation is

$$x^K = \frac{\rho + \kappa(\eta - 1)}{n\eta} R.$$

4.2 Non-cooperative exploitation by agents with moral scruples

Does the central planner's solution co-incide with Nash behaviour by agents who have concerns for self-image? We assume that self-image is related to the difference between one's action level and the Kantian action, $x^K(R)$, as specified by equation (8) above. Assume that an individual's utility function is the sum of two functions: (i) the material wellbeing function, $M_i(x)$, and (ii) the self-esteem function, $v_i(R, x_i, x^K(R))$ defined by

$$v_i(R, x_i, x^K(R)) = A_i - \theta_i \max\left[0, \sigma_i(R)(x_i - x^K(R))\right],$$

where A_i is a constant (let us call A_i "agent i's intrinsic level of self-esteem"), $\theta_i \in [0, 1]$ is a parameter called agent i's "degree of moral scruple", $\sigma_i(R)$ is agent i's perception of the harm that he would inflict on other individuals if he were to overexploit the resource stock, and $x_i - x^K(R)$ is a measure of his deviation from the Kantian ideal action. This formulation says that if $x_i > x^K(R)$, then agent i feels bad because he overextracts, violating the Kantian norm. Note that if $x_i < x^K(R)$, then his self-esteem is not affected.

Each individual chooses $x_i(t)$ to maximize

$$W_i = \int_0^\infty e^{-\rho t}\left\{M(x_{it}) + A_i - \theta_i \max\left[0, \sigma_i(R_t)(x_{it} - x^K(R_t))\right]\right\} dt,$$

subject to

$$\dot{R}_t = G(R_t) - x_{it} - \sum_{j \neq i} x_{jt},$$

and $\lim_{t\to\infty} R(t) \geq 0$.

We now state and prove Proposition 1.

Proposition 1: *Suppose that the agent's "perception of harm function" $\sigma_i(R)$ is equal to $(n-1)V_P'(R)$, where $V_P(R)$ is the value function of the social planner's problem, as defined in equation (7). If agent i expects that all other agents use the extraction strategy $x_j = x^K(R)$ as given by (8) then, provided that $\theta_i = 1$, he will himself use the same extraction strategy, $x_i = x^K(R)$, resulting in a equilibrium that is socially optimal at every point of time. At the Markov-perfect Nash equilibrium, the value function of agent i turns out to be equal to the social planner's value function, $V_P(R)$, plus the constant term A_i/ρ.*

$$W_i(R) = \frac{A_i}{\rho} + V_P(R). \tag{9}$$

Proof:

We only need to verify that the candidate value function $W_i(R)$ as specified by equation (9) does indeed satisfy agent i's HJB equation and leads to the exploitation strategy $x_i = x^K(R)$. Given that $\sigma_i(R) = (n-1)V_P'(R)$, the HJB equation for agent i is

$$\begin{aligned}\rho W_i(R) &= \max_{x_i} \left\{ M(x_i) + A_i - \theta_i \max\left[0, (n-1)V_P'(R)(x_i - x^K(R))\right] + \right. \\ &\left. \left[G(R) - (n-1)x^K(R) - x_i\right] W_i'(R) \right\}.\end{aligned}$$

Using our candiadate value function, the first order condition is

$$M_i'(x_i) - [\theta_i(n-1)+1] V_P'(R) = 0.$$

With $\theta_i = 1$, we get

$$x_i = M'^{-1}\left(nV_P'(R)\right) \equiv x^K(R).$$

Substituting this into the HJB equation of agent i, we get

$$\rho W_i(R) = M(x^K(R)) + A_i - 0 + \left[G(R) - x^K(R)\right] V_P'(R).$$

By plugging (9) to the left-hand side of the above equation, we can verify that the claim that (9) is agent i's value function is indeed valid.■

5 A discrete-time model of renewable resource exploitation by image-conscious agents

Let us see how our result for the continuous-time model can be adapted for the case of discrete time. Again we first solve the social planner's problem. After that, we show how the socially optimal outcome can be implemented as a Nash equilibrium among agents with a sufficiently strong concern for self-image. As expected, the basic result of the continuous-time model carries over to the discrete -time model, provided that the "perception of harm" function $\sigma_i(R)$ is suitably modified, as discussed after the statement of Proposition 2 below. This shows the robustness of our conclusion concerning achieving the social optimum by means of Nash behaviour of agents who have a sufficiently strong concern for self-image.

5.1 The social planner's problem in discrete time

Let X_t be the agregate harvest, i.e.,

$$X_t = \sum_{i=1}^{n} x_{it}.$$

We assume that the law governing the dynamics of the stock is

$$R_{t+t} = g(R_t, X_t),$$

where $g_R > 0$ and $g_X < 0$.

Let β be the discount factor, where $0 < \beta < 1$. The social planner's Bellman equation is

$$\begin{aligned} V_P(R_t) &= \max_{x_t \geq 0} \{M(x_t) + \beta V_P(R_{t+1})\} \\ &= \max_{x_t \geq 0} \{M(x_t) + \beta V_P(g(R_t, nx_t)\} . \end{aligned}$$

The first order condition is

$$M'(x_t) + \beta V_P'(g(R_t, nx_t))ng_X(R_t, nx_t) = 0.$$

(Note that $V_P' > 0$ and $g_X < 0$). From the first order condition, we obtain x_t as a function of R_t. We denote this solution by

$$x_t = x^K(R_t, V_P'). \tag{10}$$

Then, substituting (10) into the Bellman equation, we get a first order differential equation that relate V_P to V_P' :

$$V_P(R_t) = M(x^K(R_t, V_P')) + \beta V_P(g\left[R_t, nx^K(R_t)\right])ng_X\left[R_t, nx^K(R_t)\right].$$

Imposing the transversality condition, this first order differential equation in V_P can be solved to yield the value function V_P and hence the Kantian level of exploitation.

Example 3

This example is drawn from the fish war model of Levhari and Mirman (1980). Assume

$$G(R, X) = (R - X)^{\alpha} \text{ where } 0 < \alpha < 1,$$

and

$$M(x_i) = \ln x_i.$$

Then the Bellman equation is

$$V(R) = \max_{x} \{\ln x + \beta V((R - X)^{\alpha})\}.$$

The first order condition is

$$\frac{1}{x} = \alpha n(R - nx)^{\alpha - 1} \beta V'((R - nx)^{\alpha}).$$

Try the value function

$$V(R) = D + B \ln R.$$

where B and D are to be determined. Then

$$V'(R_{t+1}) = \frac{B}{R_{t+1}} = \frac{B}{(R_t - nx_t)^{\alpha}}.$$

This allows us to solve for the optimal harvesting rule:

$$x_t = \frac{R_t}{n(1 + \alpha\beta B)}$$

By the standard method, we find that

$$B = \frac{1}{(1 - \alpha\beta)} > 0.$$

5.2 The individual's optimization problem in discrete time

Assume that agent i has a utility function that is the sum of two functions: (i) the material wellbeing function, $M(x)$, and (ii) the self-esteem function, $v_i(R, x_i, x^K(R))$ defined by

$$v_i(R, x_i, x^K(R)) = A_i - \theta_i \max\left[0, \sigma_i(R)(x_i - x^K(R))\right].$$

where A_i is a constant and $\theta_i \in [0, 1]$. We may think of A_i as agent i's intrinsic level of self-esteem.

Proposition 2: *Suppose the agent's "perception of harm function" $\sigma_i(R)$ is equal to $-(n-1)\beta V_P'(g\left[R_t, nx^K(R_t)\right])g_{X_t} > 0$, where g_X is evaluated at $X_t = nx^K(R_t)$. If agent i expects that all other agents use the extraction strategy $x_j = x^K(R)$, then, provided that $\theta_i = 1$, he will himself use the same extraction strategy, $x_i = x^K(R)$, resulting in a equilibrium that is socially optimal. At the Markov-perfect Nash equilibrium, the value function of agent i turns out to be equal to the social planner's value function, $V_P(R)$, plus the constant term A_i/ρ,*

$$W_i(R) = \frac{A_i}{\rho} + V_P(R), \tag{11}$$

where

$$\frac{1}{1+\rho} \equiv \beta.$$

Discussion: Comparing Proposition 2 (for the discrete time model) with Proposition 1 (for the continuous time model), we see the "perception of harm" function $\sigma_i(R)$ must be suitably modified to get the desired result. In the continuous time case, we required that $\sigma_i(R_t) = (n-1)V_P'(R_t)$, where $V_P(R)$ is the value function of the social planner's problem, and thus $V_P'(R_t)$ is the marginal value of the *concurrent* stock of resource. In the discrete time case, we required that $\sigma_i(R_t) = -(n-1)\beta V_P'((gR_t, nx^K(R_t)))g_{X_t}$, i.e., that

$$\sigma_i(R_t) = (n-1)\beta V_P'(g(R_t, nx^K(R_t)))|g_{X_t}|,$$

i.e.,

$$\sigma_i(R_t) = (n-1)\beta V_P'(R_{t+1})|g_{X_t}|. \tag{12}$$

where, of course, $R_{t+1} = g(R_t, nx^K(R_t))$. Here, we note two differences between the "perception of harm" functions for the discrete time case and for the continuous time case. First, in the equation (12), V_P' is evalued at R_{t+1}, not at R_t: it is the shadow price of the next period's stock, not of the concurrent stock, that matters. Second, the discount factor β appears in the equation (12) because an agent's extraction at date t reduces the stock at a later date, $t+1$.

Proof of Proposition 2

We only need to verify that the value function $W_i(R)$ specified by eq (11) satisfies agent i's Bellman equation and leads to the exploitation strategy $x_i = x^K(R)$. The Bellman equation is

$$\begin{aligned} W_i(R_t) &= \max_{x_i} \left\{ M(x_{it}) + A_i - \theta_i \max\left[0, -(n-1)\beta V_P'(g\left[R_t, nx^K(R_t)\right])g_X(x_{it} - x^K(R_t))\right] + \right. \\ & \left. \beta W_i'(g\left[R_t, (n-1)x^K(R_t) + x_{it}\right]) \right\}. \end{aligned}$$

The first order condition is

$$\begin{aligned} & M'(x_{it}) + (n-1)\beta V_P'(g\left[R_t, nx^K(R_t)\right])g_X \\ = & -\beta W_i'(g\left[R_t, (n-1)x^K(R_t) + x_{it}\right])g_X. \end{aligned}$$

Given that all agents $j \neq i$ use the strategy $x_j = x^K(R)$, the above first order condition is identical to the social planner's first order equation,

$$M'(x_{it}) + \beta V_P'(g\left[R_t, x^K(R_t)\right])g_X = 0.$$

This completes the proof. ■

6 A model of the evolution of the concern for self-image

In the preceding model, the parameter θ_i may be called the degree of pro-socialness of agent i. So far, we assume that θ_i is time-independent. Now, we open a new window, and ask: what if agent i actually is a sequence of overlapping generations? How would θ_i change from one generation to the next?

Let us consider a simple model that addreses this issue. For simplicity, we abstract from the dynamics of the resource stock. To compensate for this over-simplification, we add a feature that reflects overcrowding externalities.

Think of a village populated by n famillies. Each family consists of a parent and a child. Time is discrete. In period t, the parent works to feed the family and contributes a fraction of her income to the village's education of the young generation. We assume that moral attitude is formed in an individual when he is a child. Once the child becomes an adult in period t, he cannot change his θ_{it} (which was formed in period $t-1$).

Assume that in period t, each parent i chooses the number of goats e_{it} to maximize his utility function, which is the sum of the material payoff and of his self-image. His material payoff is

$$M_{it} = M(e_{it,}E_{-it}) = \frac{e_{it}\xi F(e_{it} + E_{-it})}{e_{it} + E_{-it}} - \beta_i g(e_{it}),$$

where ξ is the productivity parameter of the pasture. His self-image function is

$$v_{it} = A - \theta_{it} \max \left\{0, (e_{it} - e^K)\sigma\right\},$$

where σ is an objective measure of the degree of damage that his overexploitation inflicts on other members of the community. The individual takes θ_{it} as given. We assume that e^K is the exploitation level that a social planner would ask each agent to carry out, assuming that the social planner's objective is to maximize Ω_t, defined as the sum of the material payoffs:

$$\Omega_t = \sum_{i=1}^{n} M_{it}.$$

Consider the case where all members of generation t are homogeneous, in the sense that $\theta_{it} = \theta_t$ and $\beta_i = \beta_j = \beta$. We can then solve for the Kantian level e^K (which is of course independent of σ and θ_t) and Nash equilibrium e_t^N of this game. Let $s \equiv 1/n$. Clearly $e^K = sE^K$, where E^K is the solution of $\xi F'(E^K) = \beta g'(E^K/n)$.

Let $E_t^N = ne_t^N$. The symmetric Nash equilibrium can be shown to satisfy the Kuhn-Tucker condition

$$\left[(1-s)\frac{\xi F(E_t^N)}{E_t^N} + s\xi F'(E_t^N) - \beta g'\left(\frac{E_t^N}{n}\right)\right] - \theta_t \sigma \leq 0,$$

with equality holding iff $e_t^N = e^K$. We can state the following result:

Lemma 1: *There is a threshold level $\widetilde{\theta}$ such that if $\theta_t \geq \widetilde{\theta}$ then $e_t^N = e^K$. The* *threshold $\widetilde{\theta}$ is given by*

$$\widetilde{\theta} \equiv \frac{1-s}{\sigma}\left[\frac{\xi F(E^K)}{E^K} - \beta g'(sE^K)\right] > 0.$$

Proof: This follows from the above Kuhn-Tucker condition.

Corollary 1: *If the agents perceive that σ is equal to σ^*, where*

$$\sigma^* \equiv (n-1)\left[\left(\frac{1}{n}\right)\left(\frac{\xi F(E^K)}{E^K} - \xi F'(E^K)\right)\right] > 0,$$

then $\widetilde{\theta} = 1$. Under these conditions, as long as $\theta_t < 1$, the Nash equilibrium exploitation E_t^N will exceed E^K.

Proof: This follows immediately from Lemma 1 and from the fact that $\xi F'(E^K) = \beta g'(sE^K)$.

Remark: The value σ^* as defined in Corollary 1 has an intuitive economic interpretation. The term inside the square brakets is the excess of average product over marginal product, divided by the number of agents in the community. It is therefore an indicator of the marginal loss imposed on the representative agent if an agent deviates by increasing e_{it} above the Kantian level e^K. When this term is multiplied by $n-1$, the result is a measure of harm that a deviating agent inflicts on the other $n-1$ agents. If $\sigma = \sigma^*$ then when $\theta_t = 1$, each agent's concern for self-image fully internalizes the cost that his deviation would impose on others. The resulting Nash equilibrium is then Pareto efficient.

In what follows, we assume $\sigma = \sigma^*$ and consider the realistic scenario where $\theta_t \leq 1$.

Proposition 3: *Assume $\theta_t < 1$. Then the Nash equilibrium exploitation E_t^N is a function of θ_t and of ξ. An increase in θ_t will reduce E_t^N, and a increase in ξ will increase E_t^N.*

Proof: Apply the implicit function theorem to the equation

$$\left[(1-s)\frac{\xi F(E_t^N)}{E_t^N} + s\xi F'(E_t^N) - \beta g'\left(\frac{E_t^N}{n}\right)\right] - \theta_t \sigma^* = 0.$$

Example 4:

Assume that

$$\beta g(e) = \gamma e, \tag{13}$$

where $\gamma > 0$ and

$$\xi F(E) = \xi E - \frac{\xi E^2}{2} \text{where } \xi > \gamma \tag{14}$$

Then

$$E^K = 1 - (\gamma/\xi) > 0.$$

In this case, $\sigma^* = (1-s)(\xi - \gamma)/2$. Then, for all $\theta_t \in (0,1)$,

$$E_t^N = \frac{(2 - \theta_t(1-s))E^K}{1+s} < E^K.$$

And the Nash equilibrium material wellbeing of the representative adult in period t is

$$\widehat{M}(\theta_t) = \frac{1}{n}\left[(\xi - \gamma)E_t^N(\theta_t) - \frac{\xi\left(E_t^N(\theta_t)\right)^2}{2}\right]. \tag{15}$$

Proposition 4: *For all $\theta_t \in (0,1)$, a marginal increase in θ_t leads to an improvement in the community's material wellbeing in period t.*

Proof: The Nash equilibrium material wellbeing of the community in period t is

$$\frac{e_t^N(\theta_t)\xi F(E_t^N(\theta_t))}{E_t^N(\theta_t)} - \beta g(e_t^N(\theta_t)) = \frac{1}{n}\xi F(E_t^N(\theta_t)) - \beta g\left(\frac{E_t^N(\theta_t)}{n}\right) \equiv \widehat{M}_t(\theta_t).$$

Then

$$\frac{d\widehat{M}_t}{d\theta_t} = \frac{dM_t}{dE_t}\frac{dE_t^N(\theta_t)}{d\theta_t} = \left[\frac{1}{n}\xi F'(E_t^N) - \frac{1}{n}g'\left(\frac{E_t^N}{n}\right)\right]\frac{dE_t^N}{d\theta_t} > 0.$$

This completes the proof. ■

We assume that parents care about the future material wellbeing of their children when they reach their adulthood. Parents in period t know that if every member of the future generation has a higher value θ_{it+1}, then everyone will be have a higher level of material wellbeing. For this reason, they collectively have an incentive to provide a moral education for their children. Let us consider a simple model of the cost of providing moral education and show how θ evolves over time.

Let $\kappa > 0$ be the discount factor. The representative adult in period t wants to choose

e_{it} and aggregate education expenditure Z_t to maximize

$$W_{it} \equiv \left[\frac{e_{it}\xi F(e_{it}+E_{-it})}{e_{it}+E_{-it}} - \beta_i g(e_{it}) - \frac{1}{n}Z_t\right] + A - \theta_{it}\max\left\{0, (e_{it}-e^K)\sigma^*\right\} + \kappa\widehat{M}_{t+1}(\theta_{t+1}), \tag{16}$$

where $\kappa\widehat{M}_{t+1}(\theta_{t+1})$ is the value that the parent attaches to the material wellbeing of the child in the latter's adult phase. In this formulation, each parent pays (e.g., through taxation) a fraction $1/n$ of the aggregate education expenditure Z_t.

While the parent chooses e_{it} non-cooperatively, taking E_{-it} as given, we assume that all parents make a collective choice (e.g., by viting) when it comes to choosing the common level Z_t. Thus Z_t is determined as an outcome of a collective deliberation on the community's educational budget. Once Z_t has been voted on, everyone has to pay his share, Z_t/n.

We must model how θ_{t+1} is influenced by Z_t.

Let $I_t \geq 0$ denote the gross investment in the stock θ_t, such that

$$\theta_{t+1} = (1-\delta)\theta_t + I_t,$$

where $\delta \geq 0$ is the rate of depreciation of θ_t. We assume that for any target I_t, the required expenditure in terms of the numeraire good is

$$Z_t = \eta I_t + \frac{1}{2}I_t^2,$$

where η is a positive constant.

The community chooses $I_t \geq 0$ that maximizes

$$\kappa\widehat{M}_{t+1}(\theta_{t+1}) - \frac{1}{n}\left(\eta I_t + \frac{1}{2}I_t^2\right), \tag{17}$$

subject to $\theta_{t+1} = (1-\delta)\theta_t + I_t$.

Proposition 5: *Assume (13) and (14). Let*

$$\omega \equiv \frac{\kappa(1-s)^2(\xi-\gamma)(1-\gamma/\xi)}{(1+s)^2}.$$

Moreover, assume $\omega > \eta$. Then problem (17) gives rise to a dynamic path of θ_t that *converges to a positive steady-state θ^* given by*

$$\theta^* = \frac{\omega - \eta}{\omega + \delta} \leq 1.$$

If both $\eta = 0$ and $\delta = 0$, then $\theta^* = 1$, which implies that at the steady state, all agents will achieve the Kantian level of exploitation, i.e., $e_i^* = e^K$.

Proof: Omitted.

7 Conclusion

We have shown that the problem of excessive exploitation of the commons can be avoided if agents who choose their exploitation level non-cooperatively in the manner described by Nash are at the same time sufficiently concerned about their self-image as a person imbued with Kantian morality.[15] Moreover, we argue that in each generation, parents have an interest in the collective provision of moral education for their children. This can give rise to an evolution of pro-social attitude in the population. Darwin himself has written on the evolution of moral qualities. In *The Descent of Man*, Darwin (1874) wrote that "Selfish and contentious people will not cohere, and without coherence, nothing can be affected. A tribe possessing a greater number of courageous, sympathetic and faithful members, who were always ready to warn each other of danger, to aid and to defend each other would spread and be victorious over other tribes. Thus, the social and moral qualities would tend slowly to advance and be diffused throughout the world."(Darwin, 1874, Chapter 5, p. 134-5.)[16]

While Darwin did not explicitly mention moral education as a factor that reinforces the cultural selection process, it should be obvious that tribal leaders do provide moral education to children in the form of morality tales, so that they would grow up as cooperative adults and benefit from the material gains brought about by social cooperation. The transmission of pro-social values across generations is in fact a co-evolutionary process, both by conscious decisions and by natural selection.[17]

[15] As pointed out by a reviewer, if there are both "green" and "brown" agents, as in Wirl (2011), the effect of "green" agents is weakened because the incentive to free ride increases for the "browns".

[16] Darwin's argument was the basis for the theory evolution employing group selection. Admittedly, this theory is not without its critics. Whether group selection is a good hypothesis or not is a matter of debate. For interesting discussions of these issues, see Stephen Jay Gould (1980, 1993).

[17] For a discussion of co-evolution, see e.g. Binmore (2005).

Acknowledgement: I thank two anonymous reviewers for their helpful comments.

APPENDIX: Kantian equilibrium with heterogeneous contributors to a public good.

In Section 3.2, we found that the condition characterizing the Kantian equilibrium allocation in the common-property resource model (where utility is *linear* in consumption) is also the condition that characterizes the optimal allocation under the standard utilitarian objective of maximizing the *non-weighted* sum of individuals' utilities. This appendix shows that this equivalence between the Kantian equilibrium (with heterogeneous agents) and the Benthamite utilitarian maximization does not carry over to a public good model (where utility is *non-linear* in the public good). Indeed, we prove below that the Kantian equilibrium in a public good model with heterogeneous consumers is equivalent to maximizing a *weighted* sum of individuals' utilities.

Consider the following simple model of private contributions to a public good. Let s_i denote the contribution of agent i. Assume that the benefit that each agent derives from the public good S is $B(S)$ where $B(S)$ is increasing and strictly concave, with $\lim_{S\to\infty} B'(S) = 0$. The cost to agent i is

$$\psi_i(s_i) = \frac{1}{\alpha_i}c(s_i),$$

where $1 = \alpha_1 \leq \alpha_2 \leq \alpha_3 \leq .... \leq \alpha_n$, and $c(s)$ is strictly convex and increasing function, with $c'(0) = 0 = c(0)$. Agent 1 is the highest cost agent. Define a Kantian equilibrium of contributions as a strictly positive vector $(s_1^K, s_2^K, s_3^K, ..., s_n^K)$ such that for each household i, if it changes s_i^K to λs_i^K, it will find that, for all λ such that $0 < \lambda \neq 1$, its material wellbeing will fall, assuming that all other households would change their s_j^K by the same factor λ.

Formally a vector $(s_1^K, s_2^K, s_3^K, ..., s_n^K)$ is a Kantian equilibrium (in thought) if and only if

$$1 = \arg\max_{\lambda>0} B(\lambda S^K) - \frac{1}{\alpha_i}c(\lambda s_i^K).$$

Again, let M_i denote the material payoff of household i:

$$M_i(\lambda) = B(\lambda S) - \frac{1}{\alpha_i}c(\lambda s_i^K).$$

Differentiating M_i with respect to λ, we get the first order equation

$$B'(\lambda S)S^K - \frac{1}{\alpha_i} c'(\lambda s_i^K) s_i^K = 0 \text{ for } i = 1, 2, ..., n.$$

Evaluated at $\lambda = 1$, we get

$$B'(S^K)S^K = \frac{1}{\alpha_i} c'(s_i^K) s_i^K = \frac{1}{\alpha_1} c'(s_1^K) s_1^K. \tag{18}$$

Take the special case where

$$c(s) = \frac{s^{1+\varepsilon}}{1+\varepsilon} \text{ with } \varepsilon > 0.$$

Then

$$B'(S^K)S^K = \frac{1}{\alpha_i} \left(s_i^K\right)^{1+\varepsilon},$$

and

$$\frac{s_j^K}{s_1^K} = \left(\frac{\alpha_j}{\alpha_1}\right)^{\frac{1}{1+\varepsilon}} \equiv \gamma_j \geq 1.$$

It follows that

$$S^K = s_1^K \sum_{j=1}^{n} \gamma_j \equiv s_1^K \Gamma,$$

and

$$s_1^K = \frac{S^K}{\Gamma} \text{ and } s_j^K = \gamma_j s_1^K = \frac{\gamma_j}{\Gamma} S^K.$$

Then

$$B'(S^K)S^K = \left(s_1^K\right)^{1+\varepsilon} = \left(\frac{S^K}{\Gamma}\right)^{1+\varepsilon},$$

and

$$B'(S^K) = \frac{1}{\Gamma^{1+\varepsilon}} \left(S^K\right)^{\varepsilon}.$$

Since the left-hand side is decreasing in S and the right-hand side is increasing in S, there exists a unique $S^K > 0$, given that we have assumed that $\lim_{S\to\infty} B'(S) = 0$. Thus we can compute s_1^K and $s_j^K = \gamma_k s_1^K$, for all $j = 2, 3, ..., n$.

It is easy to see that the Kantian solution $(s_1^K, s_2^K, s_3^K, ..., s_n^K)$ maximizes M, a weighted sum of material payoffs,

$$M \equiv \sum_{i=1}^{n} \omega_i M_i,$$

where the weights ω_i are given by

$$\omega_i \equiv \frac{\gamma_i}{\Gamma}.$$

It can also be verified that

$$\frac{s_i^K}{S^K} = \frac{\gamma_i}{\Gamma} = \omega_i.$$

The Kantian solution is Pareto efficient. Indeed, the Samuelsonian efficiency condition is satisfied: the sum of individuals' marginal rate of substitution (MRS) of the private good for the public good is equal to the marginal rate of transformation (MRT) between the private good and the public good. At the Kantian allocation, the Lindhal price for individual i is

$$P_i = \frac{B'(S^K)}{c'(s_i^K)/\alpha_i} = \frac{\frac{s_i^K}{S^K} B'(S^K)}{\frac{s_1^K}{S^K} c'(s_1^K)/\alpha_1}.$$

Thus the sum of these Lindhal prices are equal to 1 (using eq. (18)):

$$\sum P_i = \frac{B'(S^K)}{\frac{s_1^K}{S^K} c'(s_1)/\alpha_1} = 1,$$

i.e., the sum of MRS equals MRT.

References

[1] Akerlof, G. A., and R.E. Kranton (2005), Identity and the Economics of Organizations, *Journal of Economic Perspectives,* Vol. 19 pp. 9-32

[2] Alger, I. and Weibull, J. W. (2016), "Evolution and Kantian Morality." *Games and Economic Behavior* 98:56-67.

[3] Andreoni, J., Harbaugh, W.T. and Vesterlund, L. (2008), "Altruism in Experiments." In S.N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary in Economics*, 2nd Edition. New York: Palgrave Macmillan.

[4] Arrow, K. J. (1974), *The Limits of Organization*. New York: W.W. Norton.

[5] Arrow, K. J. (1973), "Social Responsibity and Economic Efficiency," *Public Policy* 21: 303-317.

[6] Binmore, K. (2005). Natural Justice. Oxford University Press.

[7] Breton, M., L. Sbragia, and G. Zaccour (2010), "A Dynamic Model of International Environmental Agreements," *Environmental and Resource Economics* 45: 25-48.

[8] Bilodeau, M. and N. Gravel (2004). Voluntary Provision of a Public Good and Individual Morality. *Journal of Public Economics* 88, pp. 645-666.

[9] Bolle, F. and and A.Ockenfels (1990), "Prisoners' Dilemma as a Game with Incomplete Information." *Journal of Economic Psychology* 11, 69-84.

[10] Bolton, G. E. and A.Ockenfels (2000), "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90(1): 166-192.

[11] Bowles, Samuel (2016), *The Moral Economy: Why Good Incentives are No Substitute for Good Citizens* Yale University Press.

[12] Brekke, K. A., Kverndokk, S. and K. Nyborg (2003), An Economic Model of Moral Motivation, *Journal of Public Economics,* Vol. 87, pp. 1967-83.

[13] Camerer, C. (2003), *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press.

[14] Camerer, C. and E. Fehr (2006), "When Does "Economic Man" Dominate Social Behavior?" *Science*, Vol. 331, Issue 5757: 47-52.

[15] Charness, G. and M. Rabin (2002), "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117: 817-869.

[16] Darwin, C. (1874). *The Descent of Man.* New York: Caldwell.

[17] Dasgupta, P., D. Southerton, A. Ulph, and D. Ulph, (2016). "Consumer Behavior with Environmental and Social Externalities: Implications for Analysis and Policy," *Environmental and Resource Economics*, 65(1), 191-226.

[18] Dawes, R., and R. Thaler (1988), "Anomalies: Cooperation." *Journal of Economic Perspectives* 2, 187-197.

[19] Deissenberg, C. and G. Leitmann (2004), An Algorithmic Solution to the Interval Kyoto Game. *Computing in Economics and Finance 2004-312.* Society for Computational Economics.

[20] Deissenberg, C. and A. Peguin-Feissole (2006), Beliefs, Norms,and Markets. Editorial. European Journal of Economic and Social Systems. Lavoisier, Vol. 19(2), 167-669.

[21] Dockner, E., S. Jorgensen, N. V. Long and G. Sorger (2000), *Differential Games in Economics and Management Science.* Cambridge University Press, Cambridge, U.K.

[22] Elster, J. (2017). On Seeing and Being Seen. *Social Choice and Welfare*, Vol 49(3/4): 721-734.

[23] Fehr, E. and Schmidt, K. M. (1999), "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114(3): 817-868.

[24] Gordon, S. (1954), "The Economic Theory of a Common Property Resource: The Fishery," *Journal of Political Economy* 62(2):124-142.

[25] Gould, Stephen Jay, 1993, Eight Little Piggies: Reflections on Natural History, W.W. Norton, N.Y.

[26] Gould, Stephen Jay, 1980, The Panda's Thumb: More Reflections on Natural History, W.W. Norton, N.Y.

[27] Grafton, R.Q., T. Kompas, and N. V. Long (2017). "A Brave New World? Kantian-Nashian Interactions and The Dynamics of Global Climate Change Mitigation," *European Economic Review* 99: 31-42.

[28] Hardin, G. (1968), "The Tragedy of the Commons," *Science* 162: 1243-8.

[29] Hill, T. E. Jr. and A. Zweig (2002), *Groundwork for the Metaphysics of Morals.* Oxford: Oxford University Press.

[30] Kant, I. (1785), *Grundlegung zur Metaphysik der Sitten.* Translated by Mary J. Gregor: *Groundwork of the Metaphysics of Morals*, Cambridge University Press, Cambridge, U.K. (1998).

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

[32] Levhari, D. and L. Mirman (1980), "The Great Fish Wars: An Example of Using a Dynamic Nash-Cournot Solution," *Bell Journal of Economics* 11:322-34.

[33] Long, N. V. (2010), *A Survey of Dynamic Games in Economics.* World Scientific, Singapore and London.

[34] Long, N. V. (2016), "Kant-Nash Equilibrium in a Quantity-Setting Oligopoly," in P. von Mouche and F. Quartieri (eds.), *Equilibrium Theory for Cournot Oligopolies and Related Games: Essays in Honour of Koji Okuguchi,* Springer, New York, pp 179-202.

[35] Long, N. V. (2017), "Mixed-Strategy Kant-Nash Equilibrium in Games of Voluntary Contributions to a Public Good," in W. Buchholz and D. Ruebbelke (eds.), *The Theory of Externalities and Public Goods.* Springer, New York, pp. 107-126.

[36] Muller, Jerry Z. (1993). *Adam Smith in His Time and Ours: Designing the Decent Society.* The Free Press, New York, N.Y.

[37] Ostrom, E. (1990), *Governing the Commons: the Evolution of Institutions for Collective Actions.* Cambridge University Press, Cambridge.

[38] Roemer, J. E. (2010), "Kantian Equilibrium,"*Scandinavian Journal of Economics* 112(1): 1-24.

[39] Roemer, J. E. (2015), "Kantian Optimization: A Microfoundation for Cooperation," *Journal of Public Economics* 127: 45-57.

[40] Sartre, J.-P. (1947). *Huis Clos.* Editions Gallimard, Paris.

[41] Sen, A. (1977), "Rational Fools: A Critique of the Behavioral Foundation of Utility," *Philosophy and Public Affairs* 6(4): 317-344.

[42] Sethi, R. and E. Somanathan (1996), "The Evolution of Social Norms in Common Property Resource Use," *American Economic Review* 86(4): 766-88.

[43] Smith, A. (1790 [2002]) *The Theory of Moral Sentiments* (revised edition, 1790). T. Creech and J. Bell &co., Edinburg, 1790. Republished in the series Cambridge Texts in the History of Philosophy, edited by Knud Haakonssen, Cambridge University Press, 2002.

[44] Smith, V. L. (2003), Constructivist and Ecological Rationality in Economics. *American Economic Review* 93(3), 456-508.

[45] Ulph, A. and D. Ulph (2017), A Theory of Consumption Norms and Implications for Environmental Policy. Typescript, University of Manchester.

[46] Wirl, F. (1995), "The Exploitation of Fossil Fuels under the Threat of Global Warming and Carbon Taxes: A Dynamic Game Approach," *Environmental and Resource Economics* 5: 333-352.

[47] Wirl, F. (2011), "Global Warming and Heterogeneous (Green and Brown) Consumers." *Scandinavian Journal of Economics* 113, 866-84.

[48] Wirl, F. and Dockner, E. (1995)," Leviathan Governments and Carbon taxes: Costs and Potential Benefits," *European Economic Review* 39: 1215-1236.

[49] Yang, Z. (2003), "Reevaluation and Renegotiation of Climate Change Coalitions: A Sequential Closed-loop Game Approach," *Journal of Economic Dynamics and Control* 27: 1563-1594.