Discussion Papers In Economics And Business

Kantian Equilibrium, Income Inequality, and Global Public Goods

Yukihiro Nishimura

Discussion Paper 25-04-Rev.2

July 2025

Graduate School of Economics
The University of Osaka, Toyonaka, Osaka 560-0043, JAPAN

Kantian Equilibrium, Income Inequality, and Global Public Goods

Yukihiro Nishimura*

Abstract

This paper develops a Kantian equilibrium framework that extends the global pollution model with private ownership, where agents condition their contributions on a universalizable moral imperative reflecting both income and preference heterogeneity. For both the Lindahl outcome and other proposed mechanisms, we identify specific proportionality conditions under which Kantian reasoning replicates these solutions as equilibrium behavior. We further show that the provision of public good does not necessarily increase with income inequality, and that some solutions exhibit invariance to inequality. Finally, we demonstrate that tradable permits may fail to achieve sufficient international redistribution to Southern countries to generate Pareto improvements over the voluntary contribution (disagreement) equilibrium. Grandfathering involves a form of proportionality between income and permit endowments, we show that this structure is better motivated by altruism. Our analysis contributes to a reinterpretation of morally grounded mechanisms for global public good provision, bridging normative ethics with economic design.

Keywords: Global externalities, Kantian equilibrium, Income inequality, International emissions trading

JEL classification number: H41, D63, Q54

*Osaka University and CESifo; E-mail: ynishimu@econ.osaka-u.ac.jp

1 Introduction

The provision of global public goods — particularly climate stability — poses persistent challenges at the intersection of equity, efficiency, and international cooperation. Classical approaches such as Lindahl (1919) pricing, voluntary contribution mechanisms (Warr (1982), Warr (1983), Bergstrom et al. (1986), "disagreement equilibrium" in the pollution game), and Pigouvian compensation have each provided partial solutions. Yet the deeper ethical and strategic dimensions of global public good provision remain under-explored, particularly in light of international income inequality and heterogeneity in preferences.

Uzawa (2003) introduced the Lindahl solution in interpreting the allocation of pollution permits where its Lindahlian determination results in unanimous agreement. Nishimura (2008) reformulated this outcome as an equilibrium of a game of international negotiations on the level of global pollution. Nishimura (2008) also clarifies that Uzawa's determination corresponds to the Pigouvian compensation as well as a matching grant familiar in the public economics literature.

This paper contributes by embedding the Kantian equilibrium framework by Bordignon (1990), Roemer (2010) and Roemer and Silvestre (2023), where universalizable moral imperative is introduced into strategic behavior by positing that agents choose their contributions. As such, our framework generalizes standard public good models by nesting the Lindahl equilibrium within a Kantian motive. We further evaluate our formulation in the Global Public Good Purchasing (GPGP) framework proposed by Bradford (2007) and Guesnerie (2007), and we link the features of international transfers to international environmental agreements to be ethically valid and politically feasible.

As a natural extension of the Warlasian equilibrium in the private-good economy, we design the global public-good mechanisms based on following candidate properties for the resource allocation:

- Benefit principle: compensation against bilateral externalities (Nishimura's (2008) interpretation of Uzawa (2003); Chander and Tulkens' (1997) ratio equilibrium)
- Initial endowment (entitlement) on your country's production (Uzawa (2003)) or the disagreement level of individual production (Chander and Tulkens (1997)).

Adding "no inter-country transfer" to the second principle would imply the no-transfer-efficient equilibrium (Shiell (2003)).

We show that, as an application to the climate-stability model and also as a clarifying explanation of Roemer and Silvestre (2023)'s mathematical derivation, Kantian implementation of Lindahl equilibrium requires agents to adjust their moral expectations based on both income differences and preference heterogeneity — the latter is inequality sensitive even when agents have identical utility functions. This links ethical reasoning to economic structure and establishes a micro-foundation for cooperative behavior absent centralized enforcement.

However, the Kantian imperative is, in Edgeworth's (1897) terminology, according to post-contribution outcome. To reconcile this with Lindahl pricing, the clean air benefit must be treated as a virtual income (Proposition 2). Our derivation highlights that the ways to reallocate initial endowments critically affect the Kantian logic of proportionality.

Moreover, through the equivalence with Kantian is observed, we show counterintuitive implications: due to the strategic structure of proportional contributions, the provision of public good increases with income inequality — a result that does not have an immediate intuition from Kant (Proposition 1). However, in the ratio equilibrium and the no-transfer allocation, Warr's (1982, 1983) type of neutrality holds, in contrast to Buchholz and Rübbelke (2023) and rather consistent with Baumgärtner et al. (2017)).

Additionally we demonstrate several key findings. First, contrary to Uzawa (1999), his Lindahl allocation does not Pareto dominate the disagreement equilibrium, echoing results in Nishimura (2008). Second, we confirm that even Pareto-efficient allocations without monetary transfers may leave some countries worse off compared to disagreement, suggesting fundamental limits to decentralized solutions based on tradable permits (Propositions 3 and 4). Applying a notion of the membership Kantian by Eichner and Pethig (2024) would give justification to the ratio equilibrium, but not to the Lindahl equilibrium where members of a coalition can be better off through the Kantian logic applied only to the members.

Bordignon (1990) did not aim for achieving Pareto efficiency of the Kantian behavior. Consistent with this motivation, we also discuss if Pareto inefficient outcomes can be compatible with Kantian imperatives. An example is a grandfathering scheme, where Pareto domination of the disagreement equilibrium is possible but Pareto efficient outcome is not possible (Tadenuma (2005)). Although the implied allocation rule involves a form of proportionality between income and permit endowments, we show that this structure resists a Kantian justification by Roemer and Silvestre (2023). Altruism is a better motivation for such outcomes.

2 Model

Consider a simple economy consisting of two groups of countries, —South (group 1) and North (group 2)— where the North is both wealthier and places a higher valuation on environmental quality as a global public good. Intrinsically, a representative citizen in each group of the countries has identical utility functions, but the preference towards the global public good is contingent on the circumstance — which is shaped by income level. The total number of the countries is $n > 1$ with n_1 northern countries and $n_2 = n - n_1$ southern countries.

Suppose that, with respect to private good x_i and public good G, the utility function of a citizen in country i is $u_i = u(x_i, G)$ which is increasing in both arguments and quasi-concave. The countries involve in production defined as a concave production function $f_i(y_i)$, $f_i'(y_i) \geq 0$ and $f_i''(y_i) \leq 0$, where air pollution y_i is an input. The initial level of the public good is $\widetilde{G}$, say, the level of clean air before industrialization, and $G = \widetilde{G} - n_1 y_1 - n_2 y_2$. The resource constraint is:

$$n_1 x_1 + n_2 x_2 = n_1 f_1(y_1) + n_2 f_2(y_2). \tag{1}$$

The Samuelson condition is:

$$n_1 \frac{\partial u_1/\partial G}{\partial u_1/\partial x_1} + n_2 \frac{\partial u_2/\partial G}{\partial u_2/\partial x_2} = f_1'(y_1^*) = f_2'(y_2^*), \tag{2}$$

where we denote the notation y_i^* when the marginal product of the pollution becomes equalized at the allocation under consideration.

In contrast, the disagreement equilibrium is determined by a Nash equilibrium where each agent maximizes $u_i(f_i(y_i), \widetilde{G} - Y_{-i} - y_i)$ given Y_{-i} and $Y = Y_{-i} + y_i$. Let the corresponding Nash equilibrium level of pollution be $\hat{y}_i$ by each agent in Northern or Southern country.

2.1 Remark on the ownership and cost structure

A notable difference from the public-good studies by Roemer and Silvestre (2023) and Buchholz and Rübbelke (2023) is that individual country (or countries that form a coalition defined below) do not have an access to the minimized cost function in (4). In the disagreement equilibrium, each country accesses its own production technology according to the agent's preference, so that cost minimization is not necessarily achieved. In Table 1, we show that some widely-studied allocations cannot be represented by a balanced linear cost sharing mechanism. In some of the proposed solutions in the present paper, η_i^2 in (5) is not a constant.

3 Global public good

We first introduce Lindahl equilibrium by Uzawa (2003): $x_i = f_i(y_i^*) + p(\theta_i Y - y_i^*)$, where p is the price of the right to pollute, and θ_i is a variable such that $n_1\theta_1 + n_2\theta_2 = 1$ which is an ownership of the polluting right Y. Optimization in production results in $f_i'(y_i^*) = p$ and the market-clearing price p is uniquely determined by $Y = n_1 y_1^* + n_2 y_2^*$.

Definition 1. Uzawa's *Lindahl equilibrium with tradable permits (TP)*: (θ_1, θ_2) constitutes a TP-Lindahl equilibrium if the representative citizens maximize utility subject to $x_i = f_i(y_i^*) + p(\theta_i Y - y_i^*)$ that generates the utility-maximizing level of the pollution $Y_1^L = Y_2^L$.

Uzawa gave the meaning of the share of the permits θ_i as the willingness to pay for the pollution reduction. The Lindahlian optimization, $\frac{\partial u_i/\partial G}{\partial u_i/\partial x_i} = \theta_i p$, combined with $n_1\theta_1 + n_2\theta_2 = 1$, generates the unanimity on the level of desired pollution constitutes the Lindahl equilibrium which satisfies the Samuelson condition. Uzawa's introduction of this definition is rather mechanical, but Nishimura (2008) gave a meaning as a bilateral price that each pollution firm should pay at the margin to agent i as a victim of multilateral pollution.[1]

As such, this Lindahl allocation generates the transfer $p(\theta_i Y^L - y_i^*)$ by permit trading. In the global climate model, Bradford (2007) and Guesnerie (2007) investigate mechanisms for the provision and financing of global public goods (GPGs), such as climate stability. Their proposal on Global Public Good Purchasing (GPGP) lays foundational insights that closely relate to our model and evaluates the Lindahl solution. Their discussion surrounds the following two points:

1. Redistribution for Participation and Equity:

   Their central theme is the need for redistribution to render participation in international environmental agreements politically and economically acceptable for all countries. Refinement of this argument by Murty (2007) is particularly relevant. She distinguishes two components: Pigouvian pricing (to internalize externalities) and equity (income redistribution).

2. Voluntary Cooperation Under Political Constraints:

   For international climate policy, emphasis is laid on its potential to facilitate cooperation. The concept of core that the allocation being immune to coalitional deviations has been placed

[1] For the case of $n_1 = 1 = n_2$ where $x_i = f_i(y_i^*) + p(\zeta_i s_i + \zeta_j s_i - y_i^*)$, and $\zeta_1 + \zeta_2 = 1$ and $s_1 + s_2 = y_1^* + y_2^*$. This is the public-bad version of the matching contribution system by Guttman (1978) and Danziger and Schnytzer (1991) where citizens announce the tolerable pollution and the other citizen compensates. For $n > 2$, Nishimura (2008) also showed that the application of Walker (1981) mechanism works where, renumbering agents with $k = 1, 2, ..., n$, $x_i = f_i(y_i^*) + p(\zeta_i \sum_{i=1}^n s_i - y_i^*)$ $\zeta_i = (1/n) + s_{i+1} - s_{i+2}$ with the conventions of $n+1 \to 1$ and $n+2 \to 2$.

an important role in the literature (Chander and Tulkens (1997) and Uzawa (1999)). As a necessary condition, the South's free-riding benefit needs to be rectified in order to realize cooperative solution. As such, Pareto domination against the disagreement equilibrium is a minimum requirement for the proposed allocation (Shiell (2003)). It may be worthwhile to mention that Uzawa's (2003) Lindahl mechanism does *not*, contrary to Uzawa (1999), Pareto dominate the disagreement equilibrium (Proposition 4 below).

Independently, Chander and Tulkens (1997) and Eyckmans (1997) showed that there exist core allocations immune to coalitional deviations which we introduce in the next section. Also, Shiell (2003) paid attention to an efficient allocation under the constraint that no net monetary transfers occur between countries. Their proposals relate to the two points of GPGP mentioned above, which we elaborate in the following.

3.1 Ownership and Pigouvian cost share

An allocation proposed by Chander and Tulkens (1997) and Eyckmans (1997) is as follows. Let the disagreement equilibrium level of pollution be $\hat{y}_i$ by each agent in Northern or Southern country. Each country adjusts y_i to y_i^*, and x_i is determined with $MRS_i \equiv \frac{\partial u_i/\partial G}{\partial u_i/\partial x_i}$ and $r_i \equiv \frac{MRS_i}{n_1 MRS_1 + n_2 MRS_2}$ as follows:

$$x_i = f_i(\hat{y}_i) - r_j \sum_{j=1,2} n_j(f_j(\hat{y}_j) - f_j(y_j^*)) \tag{3}$$

The allocation of y_i^* and $G^* = \widetilde{G} - n_1 y_1^* - n_2 y_2^*$ where each agent maximizes the utility in (3) resembles the classic ratio equilibrium by Kaneko (1977). Also, such allocation satisfies the following *production efficiency*: for $\widetilde{G} - G = n_1 y_1 + n_2 y_2$,

$$(y_1^*, y_2^*) \text{ minimizes } \sum_{i=1,2} n_i(f_i(\hat{y}_i) - f_i(y_i)) \; s.t. \sum_{i=1,2} n_i y_i = \widetilde{G} - G \tag{4}$$

so Eyckmans (1997) defines the allocation by, denoting $y_i(G)$ as the production-efficient y_i given Y, $C(G) \equiv \sum_{i=1,2} n_i(f_i(\hat{y}_i) - f_i(y_i(G)))$. Indeed, plugging the minimized cost in (4) into (3) with $C'(G) = f_i'(y_i^*) \equiv p$ through the envelope theorem, the ratio-equilibrium type utility maximization under (3) given r_i yields $MRS_i = r_i p$ for all i so the Samuelson condition is satisfied. Also, in the terminology of Roemer and Silvestre (2021), this ratio equilibrium is a special form of the following "generalized Lindahl-Foley equilibrium": consider a class of allocations where agents maximize utilities from the budget constraint of

$$x_i + p_i G = f_i(\widetilde{y}_i) + \eta_i^1(pG - C(G)) - \eta_i^2, \quad \widetilde{y}_i = \hat{y}_i \text{ or } \widetilde{y}_i = y_i^* \tag{5}$$

to select x_i and preferred Y consistent with unanimity, where η_i^1 is the Arrow-Debreu shareholding of the centralized firm and $p = n_1 p_1 + n_2 p_2$. η_i^2 is assumed to depend on y_i or G. The Foley (1970)-type firm has the profit $pG - C(G) = [p\widetilde{G} - \sum_{j=1,2} n_j f_j(\hat{y}_j)] + \sum_{j=1,2} n_j (f_j(y_j^*) - p y_j^*)$, where the first bracket consists of a portion of Lindahl taxes $\sum n_i p_i \widetilde{G}$ and the transfer $\sum_{j=1,2} n_j f_j(\hat{y}_j)$, and the second term is the maximized profit subject to the pollution fee p.

If the profit share is $\eta_i^1 = r_i$ and $f_i(\widetilde{y}_i) - \eta_i^2 = f_i(\hat{y}_i)$, then this cost-share equilibrium yields the equilibrium individual price $p_i = r_i p$ and the profit-maximizing production $p = C'(G)$ and each private firm in country j pays $r_i f_j(y_j^*)$ to agent i out of the production (Pigouvian compensation).

The ratio allocation satisfies both Pareto efficiency and Pareto domination of the disagreement equilibrium.[2] Nishimura (2008), however, showed that $\hat{y}_i - y_i^*$ may be negative for some country, even though the pollution is reduced in Pareto efficient allocation.

3.2 Transfers for voluntary cooperation

Shiell's (2003) allocation of no transfer can be obtained with the profit share $\eta_i^1 = r_i$ and $f_i(\widetilde{y}_i) - \eta_i^2 = f_i(y_i^*) + r_i C(G)$ where the second term is the compensation from the firms.[3] Y is determined by $MRS_i = p_i$ for all i. A less widely noted feature of Shiell's allocation is that it does *not* Pareto dominate the disagreement (voluntary contribution) equilibrium, thereby violating this principle of the GPGP framework — namely, the feasibility and desirability of cooperative outcomes (Nishimura (2008)).[4] But in the next section, we show that this conclusion is even stronger: *tradable permits* may not provide sufficient international redistribution of income. Chander and Tulkens (1997), in contrast, as discussed above, accompanies the redistribution of initial incomes which Pareto dominates the disagreement equilibrium. However, with tradable permit, by $f_i(\hat{y}_i) - r_i \sum_i n_i (f_i(\hat{y}_i) - f_i(y_i^*)) = f_i(y_i^*) + p(\theta_i^{CT} Y - y_i^*)$, then no natural restriction allows θ_i^{CT} to be a normal range of $0 < \theta_1^{CT} < 1$. Though Chander and Tulkens (2011) mentions that the tradable permits can be assigned

[2]The choice of $Y = \hat{Y}$ and $x_i = f_i(\hat{y}_i)$ is feasible under the proposed budget constraint (5) with $\eta_i^2 = 0$.

[3]The reason why $-\eta_i^2 = r_i C(G)$ appears as a compensation in the RHS of (5) of the no-transfer solution, in contrast with the fact that $-\eta_i^1 C(G)$ appears in the RHS of (3) and (5) of the the ratio allocation is easy to understand. The former is paid by the profit-maximizing firm, whereas in the latter, $-\eta_i^1 C(G)$ is used to calculate the marginal cost for agent i to announce a desirable level of pollution.

[4]"In the absence of regulation, the countries will be located below the [Pareto] frontier, ... [T]here is one Pareto-efficient allocation associated with zero transfers ... This point is located to the north-east of [the disagreement equilibrium], since both countries benefit from some level of pollution control, even given the existing imbalance in the distribution of income." (Shiell (2003, pp. 42-43)).

compatible with the proposed allocation, we show in Section 5 that it is not possible to assign an interior value of initial allocation of tradable permits compatible with the ratio equilibrium.

3.3 Incentive compatibility, fairness and summary

Bradford (2007) and Guesnerie (2007) also proposed the third principle of implementation, inspired from the Clarke-Groves-Vickrey framework. We have discussed above that a natural extension of the standard mechanism-design theory works for Uzawa's (1997) allocation. Eyckmans (1997) showed that the ratio equilibrium is implementable in the Nash equilibrium.

allocations	TP-Linhahl	ratio equilibrium	no transfers
If $x_i = f_i(y_i^*) + p(\theta_i Y - y_i^*)$	$0 < \theta_i < 1$	θ_i may not be in $(0,\ 1)$	$\theta_i = \frac{y_i^*}{n_1 y_1^* + n_2 y_2^*} \in (0, 1)$
η_i^1	0	r_i	r_i
$f_i(\widetilde{y}_i) - \eta_i^2$	$f_i(y_i^*) - py_i + p_i\widetilde{G}$	$f_i(\hat{y}_i)$	$f_i(y_i^*) + r_i C(G)$
Nash implementation	Yes (Nishimura (2008))	Yes (Eyckmans (1997))	n.a.
inequality of resources	increases G^*	increases G^*	neutral*
Membership-Kantian deviation	Vulnerable**	Immune	Vulnerable**

Table 1: Lindahl equilibrium and other solutions.
Uzawa (2003, Definition 1) for the TP-Lindahl, and the ratio equilibrium is associated with consumption level by (3).
$*$: Proposition 2
$**$: This subsumes the question if the allocation Pareto dominates disagreement equilibrium or not. Uzawa (1999) and Shiell (2003) respectively said "Yes", but Propositions 3 and 4 conclude "No".

Notice that the notion of private ownership in the TP-Lindahl is distinct from the other. In the latter, the share η_i^1 such that $\sum_i \eta_i^1 = 1$ determines the shareholding of firms with $f_j(y_j^*) - py_j^*$ in addition to the adjusting taxes whose sum is $p\widetilde{G} - \sum_{j=1,2} n_j f_j(\hat{y}_j)$. In the former, agent i owns $f_i(y_i) - py_i$.

4 Inequality and Kantian concept

In this section, we consider the following nested model, where the correspondence with the standard Lindahl-ratio equilibrium is easier to understand. Suppose that, with respect to private good x_i and public good G, the utility function of a citizen in country i is $u_i = \min\{(1-\alpha_1)\ln x_i + \alpha_1 \ln G,\ (1-\alpha_2)\ln x_i + \alpha_2 \ln G\}$ with $0 < \alpha_1 < 0.5 < \alpha_2 < 1$. In other words, $u_i = (1-\alpha_1)\ln x_i + \alpha_1 \ln G$ when

$x_i < G$, whereas $u_i = (1 - \alpha_2) \ln x_i + \alpha_2 \ln G$ with greater marginal rate of substitution between G and x_i if $G < x_i$. We assume that

$$n\alpha_1 < 1 - \alpha_1 \tag{6}$$

Namely, once an agent is sufficiently poor to choose $x_i < G$ in the Kantian scheme we describe below, then the valuation of the public by such agents is very low.

The countries involve in production defined as $f_i(y_i) = y_i - b_i(y_i)^2 + m_i$ if $y_i \leq \overline{y}_i$ and $f_i(y_i) = \overline{y}_i$ $-b_i(\overline{y}_i)^2 + m_i$ if $y_i \leq \overline{y}_i$, with $\overline{y}_2 = \infty$. $m_i > 0$ represents rent (excess profit) from production. b_i's ($b_1 > b_2$) are assumed to be very small so it would not appear in (approximate) calculation except that, in any efficient level of interior productions we have:

$$b_1 y_1^* = b_2 y_2^* \tag{7}$$

The three proposed allocations are transformed to the case when citizens in each Southern countries have income $W - \Delta \equiv W_1 > 0$, and a citizen in Northern countries has income $W + \frac{n_1}{n_2}\Delta \equiv W_2$ for $\Delta > 0$. The resource constraint is:

$$n_1 x_1 + n_2 x_2 - Y + Z = n_1 m_1 + n_2 m_2 + Z \equiv nW. \tag{8}$$

where we set $W_i = \hat{y}_i + m_i + r_i \hat{G}$ (the disagreement equilibrium's production and the share r_i of the excess profit of the Foley firm) to describe the ratio allocation, $W_i = m_i + r_i \widetilde{G}$ for the TP-Lindahl allocation, and $W_i = m_i + y_i^* + r_i G^*$ for no-transfer-efficient allocation. $Z = \widetilde{G}$ would warrant the above resource constraint. Δ above is conditioned by the decrease of m_1 and the increase of m_2.

We also assume that $\overline{y}_1 + m_1$ is sufficiently small so that the pollution from the Southern country may reach its upper limit in the disagreement equilibrium. Applying $MRS_i(x, x) = 1$ at non-differential points, $\hat{x}_1 = \overline{y}_1 + m_1$ and:

$$MRS_1(\hat{x}_1, \widetilde{G} - \hat{Y}) \leq 1 = MRS_2(\hat{x}_2, \widetilde{G} - \hat{Y}) \text{ and } (8). \tag{9}$$

When the Southern citizens' pollution may reach its maximum, increasing W_2 and reducing W_1 increases the level of public good (in other words, it reduces $\hat{Y}$).

Lemma 1. The TP-Lindahl equilibrium by Uzawa (2003) and the ratio ewuilibrium by Chander and Tulkens (1997) and the no-transfer efficient allocation correspond to the respective Lindahl equilibrium for public-good problem: the representative citizens maximize the utility subject to $x_i + p_i G = W_i$ and (p_1, p_2) such that $p_i = r_i p$ generates the utility-maximizing level $G_1^L = G_2^L = \widetilde{G} - (n_2 y_2 - n_1 \min\{\frac{b_2}{b_1} y_2^*, \overline{y}_1\})$.

4.1 Inequality and the provision of public good

Suppose that Δ is sufficiently high and we suppose that the Lindahl equilibirum with superscript L satisfies $x_1^L < G^L < x_2^L$ (see the complete of description in the Appendix).

The assumption of $x_1^L < G^L < x_2^L$ generates $x_i^L = (1 - \alpha_i) W_i$ so (2) is rearranged to, for $n_1 \alpha_1 + n_2 \alpha_2 = n\overline{\alpha}$:

$$n_1 \alpha_1 (W - \Delta) + n_2 \alpha_2 (W + \frac{n_1}{n_2} \Delta) = n\overline{\alpha} W + n_1 (\alpha_2 - \alpha_1) \Delta = G^L, \tag{10}$$

which also satisfies the resource constraint (8) and we assume such G^L is on $[(1 - \alpha_1)(W - \Delta), (1 - \alpha_2)(W + \frac{n_1}{n_2} \Delta)]$. Notably, this assumption is satisfied when α_1 and α_2 are not different, n_1 / n_2 is large, and Δ is large.

Buchholz and Rübbelke (2023) showed that the provision of the public good declines as income becomes more equally distributed, but Baumgärtner et al. (2017) showed that social willingness to pay for environmental goods may decrease with income inequality. The purpose of the present analysis is that this conclusion similar to that of Buchholz and Rübbelke (2023). For the TP-Lindahl solution, the increase of Δ decreases r_2 increases r_1, so the difference of virtual income $m_i + r_i \widetilde{G}$ increases.

For the ratio solution, there is ambiguity. When the pollution by the southern countries reaches its maximum, as we argued above, $\hat{y}_1 = \bar{y}_1$ and $\hat{y}_2 = \frac{1-\alpha_2}{n_2}(\widetilde{G} - n_1 \bar{y}_1) - \alpha_2 m_2$ so $\Delta \hat{y}_2 + \Delta m_2 > 0$ and $\Delta W = 0$. With the same logic as above, income inequality raises the level of G in the ratio allocation. However, when the pollution level in the disagreement equilibrium is in its interior, the inequality-invariance by Warr (1983), Warr (1982) and Bergstrom et al. (1986) apply, so the disagreement equilibrium is invariant in its consumption level. In turn, the ratio allocation exhibits distributional neutrality. The no-transfer solution, in contrast, is characterized by $n_1 r_1 (x_1, \widetilde{G} - n_1 x_1 - n_2 x_2) + n_2 r_2 (x_2, \widetilde{G} - n_1 x_1 - n_2 x_2) = 1$ and $x_1 = y_1^* + m_1 \leq \overline{y}_1 + m_1$ and $x_2 \geq m_2$.[5] If the inequality is represented by the decrease of $\overline{y}_1 + m_1$ and the increase of m_2, then, as long as that inequality is made up with the increase of y_1^* and the decrease of y_2^*, then this solution is invariant with respect to

[5] Since the marginal pollution is constant for a wide range of allocations, there are multiple solutions without transfers. With strict concavity of the production, such an allocation is unique, and the distribution invariance still holds.

the inequality of incomes. This is Warr's (1982, 1983) type of neutrality with respect to endowment inequality.

Proposition 1. (i) For the TP-Lindahl allocation, the level of the public good is greater when the inequality of initial endowment represented by Δ is greater.

(ii) If the inequality of m_i is moderate, then if (y_1^*, y_2^*) generates a no-transfer efficient solution (x_1^*, x_2^*, G^*), reducing m_1 by Δ that makes $\overline{y}_1 + m_1 - \Delta \geq x_1^*$ in exchange of the increase of the increase of m_2 by $\frac{n_1}{n_2}\Delta$ such that $m_2 + \frac{n_1}{n_2}\Delta \leq x_2^*$, then (x_1^*, x_2^*, G^*) is a non-transfer efficient solution after the increased inequality.

(iii) For the ratio allocation, inequality invariance holds of the inequality of m_i is moderate. If southern countries' pollution reach their maximum in the disagreement equilibrium, then the increased inequality increases the public food provision in the ratio allocation.

The proof of this proposition, as well as those for the following propositions, is a logical conclusion of the above argument.

In our formulation of Kantian ethics, the emergence of inequality paradox in part (i) of the above proposition is explained as follows: if richer countries are more concerned about the global environment, then, though equality invites by contribution by poorer, less environmentally concerned countries, the outcome may ultimately harm the environment more. This is because redistribution in such a setting can reduce the aggregate level of public good provision due to the strategic substitutability of individual contributions. However, the conclusions are different in the the no-transfer allocation.

4.2 Kantian structure in North-South economy

In contrast, we now consider the Kantian equilibrium that applies to the individual budget constraint $x_i = W_i - g_i$. g_i represents the abatement from $W_i = \hat{y}_i + m_i + r_i\hat{G}$ in the ratio allocation, and $W_i = m_i + r_i\widetilde{G}$ for the TP-Lindahl allocation, and $W_i = m_i + y_i^* + r_iG^*$ for no-transfer-efficient allocation. As such, the three allocations are distinct in principle.

Definition 2. Given β_1 and β_2, consider the following utility $V_i(g_i) \equiv u_i(W_i - g_i, n_ig_i + n_j(\beta_ig_i))$ for $j \neq i$. A strategy (g_1^K, g_2^K) consists of Kantian equilibrium with (β_1, β_2) if g_i^K maximizes $V_i(g_i)$ and $\beta_ig_i = g_j$ for $i = 1,\ 2$ and $j \neq i$.

Here we formulate Kantian moral reasoning, i.e., "what if everyone acted as I do" as follows: (i) every Southern (Northern) citizen behaves symmetrically; (ii) every Southern citizen has "as if" behavior for Northern citizen based on different conditional preference for the public good and different income; every Northern citizen has similar "as if" behavior for Southern citizens; (iii) the Kantian type of cooperative behavior is formulated as the utility-maximizing contribution to the public good.

The first-order condition of optimality given β_i is as follows:

$$\frac{dV_i}{dg_i} \propto -G + n_i \frac{\alpha_i}{1-\alpha_i} x_i + n_j \frac{\alpha_i}{1-\alpha_i} x_i \beta_j = 0. \tag{11}$$

The second term in the middle corresponds to the gain through coordinated contribution of g_i among citizens with the same circumstance, and in the last part, Southern citizens think of the suitable proportional behavior to the wealthier and a more public-good preferring citizens ($\alpha_2 > \alpha_1$), conditional on their wealth. A symmetric explanation applies to β_2.

Proposition 2. In this economy, if we set the coefficient $\beta_i = \frac{\alpha_j}{\alpha_i}\frac{W_j}{W_i}$ and hence the agent i expect an agent with different circumstance to behave as: $g_j = \frac{\alpha_j}{\alpha_i}\frac{W_j}{W_i} g_i$ as Kantian imperative for all the three proposed allocations with the restrictive definition of W_i above. Namely, the assumption is that each agent contributes according to income and the preference, with the latter being income-dependent in the present scenario.

In a Cobb–Douglas economy used for illustration, we impose a Kantian structure of moral universalizability. Our purpose is to show the structure that the resulting allocation is Pareto efficient. The form written as $\beta_i = \frac{\alpha_j/(1-\alpha_j)}{\alpha_i/(1-\alpha_i)}\frac{x_j^*}{x_i^*}$ applies for the Kantian imperative. In the terminology by Edgeworth (1897), the Kantian proportionality is applied to the post-contribution income x_i. For the TP-Lindahl and the ratio allocations, contributions are proportional to income and a circumstance-contingent preference. For the TP-Lindahl allocation and the ratio allocation, the imputed value of the clean air ($p_j\tilde{G}$) is a part of virtual income W_j, and p_j is dependent on j's preference. The no-transfer allocation departs from Roemer and Silvestre's (2021) Lindahl-Foley equilibrium by the appearance of $-\eta_i^2 = r_i C(G)$ in Table 1. The Kantian coefficient has to be proportional to the endogenously determined consumption level.

5 Voluntary cooperation and Membership Kantian

In the absence of coercive supreme body, the move towards cooperation should include unanimous agreement — namely, Pareto domination — from the disagreement equilibrium. However, we can show the following:

Proposition 3. In the ratio equilibrium associated with (3), defined by Chander and Tulkens (1997):

(i) some countries may emit more pollution than under disagreement equilibrium.

(ii) if we consider the tradable-permit scheme, no natural restriction allows θ_i^{CT} to be in $0 < \theta_1^{CT} < 1$.

Under quasi-linear utility with respect to x_i, Chander and Tulkens (1997) showed that the ratio allocation belongs to "γ-core", allocations which are immune to coalitional deviation, with Pareto improvement from disagreement equilibrium as a necessary condition.

Proposition 4. There exists an economy such that no Pareto efficient allocation with tradable permits with $\theta_i \in [0,\ 1]$ Pareto dominates the disagreement equilibrium.

Yang (2008) made "reverse-engineering" argument further implies that the implementation of any Pareto-efficient allocation based on a specified Negishi weighting scheme requires explicit redistribution of resources across countries.[6]

Using the definition by Eichner and Pethig (2024), we can have another justification of Pareto domination of the disagreement equilibrium. Suppose that $C_S(G_S) = \min \sum_{i=1,2} n_i^S(f_i(\hat{y}_i) - f_i(y_i))$ $s.t.$ $\sum_{i=1,2} n_i y_i = \hat{Y}_S - G_S$, and the hat denotes the aggregate emission by coalition S.

Definition 3. (Membership Kantian equilibrium) Given β_1^S and β_2^S and a set S that consists of $n_i^S \leq n_i$ members from group i, consider the following utility $V_i^S(g_i, y_{-S}) \equiv u_i(W_i - g_i, \widetilde{G} - \hat{Y}_S + G_S^{-1}(\hat{Y}_S - n_i^S g_i - n_j^S(\beta_i g_i)) - \sum_{k \notin S} y_k)$ for $j \neq i$. A strategy $(g_1^{K,S}, g_2^{K,S})$ consists of a membership Kantian equilibrium with (β_1^S, β_2^S) if: (i) $g_i^{K,S}$ maximizes $V_i^S(g_i, y_{-S})$ and $\beta_i^S g_i = g_j$ for $i = 1,\ 2$ and $j \neq i$. (ii) Each agent $k \notin S$ maximizes $u_k(f_k(y_k), \widetilde{G} - Y_{-k} - y_k)$ given Y_{-k} and $Y = Y_{-k} + y_k$.

The Kant-Samuelson condition corresponding to (11) for S is applied to conclude that, if $\beta_i^S = \frac{\partial u_j/\partial G}{\partial u_i/\partial G}\frac{\partial u_i/\partial x_i}{\partial u_j/\partial x_j}$, then the membership Kantian equilibrium in Definition 3 is Chander-Tulkens' (1997) partial agreement equilibrium on which their γ-core concept, similar to Uzawa's (1999) core, is defined.

Corollary 1. Suppose that preferences are $u_i = x_i - v_i(Y)$ and Chander and Tulkens' (1997) Assumption 1", $\sum_{i \in S} MRS_i(Y^*) \geq MRS_j(\hat{Y})$ for all $j \in S$ defined at a Pareto efficient Y^* and the disagreement equilibrium $\hat{Y}$.

Then for any subset S that consists of $n_i^S \leq n_i$ members from group i, the aggregate payoff $\sum_{i \in S} u_i(x_i, G)$ that the ratio equilibrium yields to the members of S is larger than the payoff that S

[6]His usage of "Lindahl principle" is different from conventional Lindahl framework since the latter, as shown by Buchholz and Rübbelke (2023) and others, does not necessary Pareto dominate the disagreement equilibrium.

can achieve at any membership Kantian equilibrium with $\beta_i^S = \frac{\partial u_j/\partial G}{\partial u_i/\partial G}$. However, no corresponding property can be obtained at any Pareto efficient allocation with tradable permits with $\theta_i \in [0,\ 1]$.

Namely, the core problem is translated to the one whether the proposed allocation is better than the *actual* utility from Kantian decision-making for all S. The proof of Corollary 1 follows from Chander and Tulkens' (1997) Theorem 2 for the ratio allocation, and the example in the next section with $n_1 = 1 = n_2$ (where the membership Kantian equilibrium from a singleton coincides with the disagreement equilibrium).

5.1 Illustration

To prove Propositions 3 and 4, we use another numerical example that belongs to the global-public good model:

$$n_1 = 1 = n_2,\ u_i = x_i - \frac{0.9}{2}Y^2 \text{ when } x_i \geq 0.5 \text{ or } x_i \geq -\frac{3}{11}Y + 0.5,\ u_i = x_i - \frac{0.1}{2}Y^2 \text{ otherwise}$$
$$f_1(y_1) = 0.2(y_1)^{0.5},\ f_2(y_2) = (y_2)^{0.5} \tag{12}$$

This example amends that of Nishimura (2008) where agents' preference on clean environment is a luxury, and the preference is identical and dependent on income level.

In the disagreement equilibrium, agents equate their marginal disutility of pollution to the marginal productivity of pollution; the pollution-avert Northern citizen ends up with $\hat{y}_2 = (f_2')^{-1}(-\partial u/\partial Y) = 0.258$ whereas the consumption-oriented South has (circumstance-oriented) high willingness to pollute despite low fuel efficiency: $\hat{y}_1 = 0.836$. In the Pareto efficient allocation, marginal cost of pollution is equated; the high productive Northern technology is utilized to the greater extent than at the disagreement equilibrium so that $y_2^* = 0.614$ and $y_1^* = 0.025$ (where $f_2'(y_2^*) = f_1'(y_1^*)$ and the reduction of the pollution at Pareto efficient allocations is made by Southern country).[7] This is Proposition 3(i).

The conclusions corresponding to Buchholz and Rübbelke (2023) and others hold in that Southern citizens become worse-off at the Lindahl equilibrium than at the disagreement point ($u_1^L < \hat{u}_1$; the current numerical example is the independent proof needed in this particular circumstance). However, the conclusion related to tradable permits is sharper. Here, Southern citizens' utility through tradable permit trading becomes the highest when $\theta_1 = 1$ and $\theta_2 = 0$ in the tradable-permit allocation in Section 3. However, this maximal utility by Southern citizens is below the utility level of the disagreement equilibrium since the agent who minds pollution less

[7]Since the utility is quasi-linear with respect to x_i, the Pareto efficient level of pollution is unique.

damages the global climate in the disagreement equilibrium. This confirms Proposition 4. Since $u_1^{CT} > \hat{u}_1 > u_1^L$ by representing the transfer that Southern citizens receive at the ratio allocation as $p(\theta_1^{CT}Y^* - y_1^*)$, $\theta_1^{CT} > 1$ has to be the case.

6 Implementing Pareto inefficient allocations and altruism

Tadenuma (2005) showed that, when $\hat{\theta}_i = \hat{y}_i/\hat{Y}$ for all i (*grandfathering*) and $\pi_i(p) = \max_{y_i} f_i(y_i) - py_i$, by changing Y, the allocations with $x_i = \pi_i(p) + p\hat{\theta}_i Y$ contains those that Pareto dominate the disagreement equilibrium. The logic is straightforward since $x_i = f_i(\hat{y}_i)$ and $Y = \hat{Y}$ is feasible, so simply by setting the market-clearing price associated with $\hat{Y}$ and achieving production efficiency Pareto dominates the disagreement equilibrium. Proposition 4 is applicable so that Pareto efficient choice of Y with grandfathering would not Pareto dominate the disagreement equilibrium. The expression in Table 1 is similar to the TP-Lindahl equilibrium except that p_i is not a Lindahl price.[8]

Restricting to the grandfathering and setting the cost-minimizing production (4) such that $y_i^* = y_i(\widetilde{G} - Y)$ and p to be the unique market-clearing choice with Y, *the constrained efficient outcomes with grandfathering* is obtained by maximizing the social welfare function $W = u_1\gamma + u_2$ such that, in the modified Cobb-Douglas case in Section 4:

$$\frac{dW}{dY} \propto G\left(\frac{\partial x_1}{\partial Y} n_2 \gamma + \frac{x_1/(1-\alpha_1)}{x_2/(1-\alpha_2)}\right) - \frac{(1-\alpha_2)G/x_2}{1-\alpha_1} x_1 n_1 \frac{\partial x_1}{\partial Y} - n_2\left(\frac{\alpha_2}{1-\alpha_1} + \frac{\alpha_1}{1-\alpha_1}\gamma\right) x_1 = 0. \quad (13)$$

In the unconstrained Pareto concept, the Samuelson condition (2) is compatible with a Hicksian budget set such that $\frac{\partial x_i}{\partial Y} = MRS_i f_i'(y_i^*)$ and the Negishi weight $\gamma = \frac{\partial u_2/\partial x_2}{\partial u_1/\partial x_1}\frac{n_1}{n_2}$.

On the other hand, the constrained efficiency with $x_i - \pi_i(p) = p\hat{\theta}_i Y$ is a kind of proportionality but $\frac{\partial x_i}{\partial Y} = \frac{\partial p}{\partial Y}(-y_i^* + \hat{\theta}_i Y) + p\hat{\theta}_i$. But the Roemer and Silvestre (2023)-type of Kantian justification, (11), is hard to be obtained. The interpretation is more straightforward to assume that agent 2 (a Northern citizen) has an altruistic preference.

7 Conclusion

We suggest a reinterpretation of Lindahl's (1919) justice notion and Roemer's (2010) reciprocity principle in international public finance. However, by integrating ethical reasoning with income

[8]Denoting $u_i^{gf}(Y)$ be the outcome of the tradable permits with grandfathering $\hat{\theta}_i$ and pollution Y, in the example of Section 5.1 we have $u_1^{gf}(\hat{Y}) > \hat{u}_1 > u_1^{gf}(Y^*)$ and $u_2^{gf}(Y^*) > u_2^{gf}(\hat{Y})$. In general, the constrained efficiency frontier may or may not include $\hat{Y}$ but it includes $Y^* < \hat{Y}$ (but Y^* may not Pareto dominate the disagreement equilibrium).

and preference heterogeneity, decentralized ownership of firms or the disagreement level as an initial endowment rather weakens the reasons to be compatible with universalizable ethics. We also showed that equitable and cooperative outcomes cannot arise without explicit redistribution. This issue makes the Southern countries worse-off than the disagreement equilibrium in both Uzawa (2003) and Shiell (2003). It seems that this feature is not known in the literature, but this is in line with Buchholz and Rübbelke (2023) in the pure public-good model. In contrast, Buchholz and Rübbelke's (2023) another conclusion that the public good provision in the Lindahl equilibrium increases with inequality does not hold under the formulations of Chander and Tulkens (1997) and Shiell (2003).

Appendix

Along the utility possibility frontier, we consider the following way. Consider first $x_1 = x_2 = x^*$ and we consider $\max u_i(x^*, n(W - x^*))$. If $x^* > n(W - x^*)$, agent prefers $\frac{x^*/W}{1-\alpha_2} = \frac{1-\frac{x^*}{W}}{\alpha_2}$ so $x^*/W = 1 - \alpha_2$, which is violated since $\alpha_2 > 0.5$. If $x^* < n(W - x^*)$, agent prefers $\frac{x^*/W}{1-\alpha_1} = \frac{1-\frac{x^*}{W}}{\alpha_1}$ so $x^*/W = 1 - \alpha_1$, which holds if $n\alpha_1 > 1 - \alpha_1$ which we exclude by assumption.

Therefore, the only possibility is $x^* = n(W - x^*)$. If one changes x^* by dx, the utility changes by $(\frac{1-\alpha_i}{x^*} - \frac{\alpha_i}{x^*}n)dx$ with $i = 1$ if $dx < 0$ and $i = 2$ if $dx > 0$. The utility increases in neither direction. This allocation is the Lindahl allocation with $p_i = \frac{1}{n}$ and $W_i = W$, which is Pareto efficient.

Then consider a Pareto efficient allocation with $\widetilde{x}_1$ in the neighborhood of x^* but $u_1(\widetilde{x}_1, \widetilde{G}) < u_1(x^*, x^*)$. Clearly we have $\widetilde{x}_1 < \widetilde{G}$ so one can reduce $\widetilde{x}_1$ from x^* and there exists $(\widetilde{x}_2, \widetilde{G})$ that satisfies $u_2(\widetilde{x}_2, \widetilde{G}) > u_2(x^*, x^*)$. $\widetilde{x}_2 < \widetilde{G}$ would not fit the Samuelson condition so $\widetilde{x}_2 \geq \widetilde{G}$. If $\widetilde{x}_2 > \widetilde{G}$, then $\widetilde{x}_2$ and $\widetilde{G}$ corresponds to x_2^L and G^L in the text, and otherwise the allocation is determined by $\widetilde{x}_2 = \widetilde{G} > x^*$, $\frac{\alpha_1}{1-\alpha_1} = \frac{\widetilde{G}}{\widetilde{x}_1}$ and the resource constraint. In both cases, the level of the public good increases when $\widetilde{x}_1$ is reduced.

References

Baumgärtner, S., Drupp, M.A., Meya, J.N., Munz, J.M., Quaas, M.F., 2017. Income inequality and willingness to pay for environmental public goods. Journal of Environmental Economics and Management 85, 35–61.

Bergstrom, T., Blume, L., Varian, H., 1986. On the private provision ofpublic goods. Journal of Public Economics 29, 25–49.

Bordignon, M., 1990. Was kant right?: voluntary provision of public goods under the principle of unconditional commitment. Economic Notes , 342–372.

Bradford, D., 2007. Improving on Kyoto: Greenhouse Gas Control as the Purchase of a Global Public Good. The Design of Climate Policy, edited by Roger Guesnerie and Henry Tulkens, MIT Press.

Buchholz, W., Rübbelke, D., 2023. Improving public good supply and income equality: Facing a potential trade-off. FinanzArchiv 79, 146–163.

Chander, P., Tulkens, H., 1997. The core of an economy with multilateral environmental externalities. International Journal of Game Theory 26, 379–401.

Chander, P., Tulkens, H., 2011. The Kyoto protocol, the Copenhagen Accord, the Cancun Agrements, and beyond. LIDAM Discussion Papers CORE, 2011-51 .

Danziger, L., Schnytzer, A., 1991. Implementing the lindahl voluntary-exchange mechanism. European Journal of Political Economy 7, 55–64.

Edgeworth, F., 1897. The pure theory of taxation. reprinted in: R.A. Musarave and A.T. Peacock. eds., Classics in the theory of public finance (Macmillan, New York, 1958) .

Eichner, T., Pethig, R., 2024. International environmental agreements when countries behave morally. Journal of Environmental Economics and Management 125, 102955.

Eyckmans, J., 1997. Nash implementation of a proportional solution to international pollution control problems. Journal of Environmental Economics and Management 33, 314–330.

Foley, D., 1970. Lindahl's solution and the core of an economy with public goods. Econometrica 38, 66–72.

Guesnerie, R., 2007. The Design of Post-Kyoto Climate Schemes: Selected Questions in Analytical Perspective. The Design of Climate Policy, edited by Roger Guesnerie and Henry Tulkens, MIT Press.

Guttman, J., 1978. Understanding collective action: matching behavior. American Economic Review, Papers and Proceedings 68, 251–255.

Kaneko, M., 1977. The ratio equilibrium and a voting game in a public goods economy. Journal of Economic Theory 16, 123–136.

Murty, S., 2007. Design of Climate Change Policies: A Discussion of the GPGP Approach of Bradford and Guesnerie. The Design of Climate Policy, edited by Roger Guesnerie and Henry Tulkens, MIT Press.

Nishimura, Y., 2008. A Lindahl solution to international emissions trading. Queen's Economics Department Working Paper, No. 1177 .

Roemer, J.E., 2010. Kantian equilibrium. Scandinavian Journal of Economics 112, 1–24.

Roemer, J.E., Silvestre, J., 2021. Kant and Lindahl. Cowles Foundation Discussion Paper2278 .

Roemer, J.E., Silvestre, J., 2023. Kant and Lindahl. Scandinavian Journal of Economics 125, 517–548.

Shiell, L., 2003. Equity and efficiency in international markets for pollution permits. Journal of Environmental Economics and Management 46, 38–51.

Tadenuma, K., 2005. Possibility and optimality of agreements in international negotiations on climate change. No. 2004-13. Graduate School of Economics, Hitotsubashi University, 2005. .

Uzawa, H., 1999. Global warming as a cooperative game. Environmental Economics and Policy Studies 2, 1–37, reprinted in Uzawa (2003), Chapter 7.

Uzawa, H., 2003. Economic Theory and Global Warming. Cambridge University Press.

Walker, M., 1981. A simple incentive compatible scheme for attaining lindahl allocations. Econometrica 49 , 65–71.

Warr, P., 1982. Pareto optimal redistribution and private charity. Journal of Public Economics 19, 131–138.

Warr, P., 1983. The private provision of a public good is independent of the distribution of income. Economic Letters 13, 207–211.

Yang, Z., 2008. Strategic Bargaining and Cooperation in Greenhouse Gas Mitigations. MIT Press.