

Potentially Harmful International Cooperation on Global Public Good Provision

By WOLFGANG BUCHHOLZ†, RICHARD CORNES‡ and DIRK RÜBBELKE††

†University of Regensburg and CESifo Munich ‡Australian National University

††Basque Centre for Climate Change and IKERBASQUE

Final version received 13 August 2013.

Experience from climate policy suggests that full cooperation among all countries is not a likely outcome. In this paper we therefore consider the case where only members of a subgroup of countries cooperate by reciprocally matching their public good contributions. In a two-stage game, matching rates are set at stage 1 then national contributions are chosen at stage 2. In the case of small coalitions, negative matching may result in the subgame-perfect equilibrium that decreases global public good provision and outsiders' welfare. Moreover, a growing number of countries may paradoxically entail a reduction of equilibrium public good supply.

INTRODUCTION

International cooperation is a prerequisite for the attainment of an efficient outcome in the presence of global public goods. However, international negotiations suffer from divergent national interests, different conceptions of fair burden or effort sharing, and a lack of trust between the negotiating countries. Such obstacles are regularly observed in negotiations in various contexts such as international trade agreements, regulation of financial markets, development aid and international environmental protection.

A prominent example for the limited success of international negotiations is climate change policy. Up to now, negotiators have been unable to achieve a global agreement following the Kyoto Protocol. A 'grand coalition' that credibly commits all polluting parties to domestic emission reductions is out of sight. As a precursor of a global agreement to be reached by 2020, only a rather informal 'road map' has been agreed on at the Conference of Parties in Durban 2011, which casts serious doubts on the prospects of a top-down approach in future global climate policy. In practice, a bottom-up approach that is based on international agreements among subgroups of like-minded countries—e.g. the EU member countries—seems likely to be the more realistic perspective.¹

In this paper we analyse the outcomes of international cooperation in the provision of a global public good like climate protection in which the set of countries participating in an international environmental agreement is of only limited extent. Buchholz *et al.* (1998) show that offsetting or 'crowding-out' behaviour of the outsiders will weaken incentives to cooperate and to form a coalition when the group of 'willing' countries is of limited size. In contrast, Vicary (2012) recently found conditions under which even a small group of agents can gain from unilaterally forming a coalition. Neither paper, however, refers to explicit equilibrium concepts that permit inferring the optimal design of collective action within some cooperating coalition. In this paper, a full treatment of equilibrium solutions is given by linking the partial cooperation issue with the matching approach that has gained considerable attention in public economic theory.

Matching mechanisms have been considered in different fields of economics for a long time. So in the theory of fiscal federalism, matching grants flowing from the central

government to lower-level jurisdictions serve as a standard instrument for internalizing interregional spillovers (see Oates (1972) as the classical reference).² In public good theory, matching mechanisms that aim to improve efficiency of public good supply were first suggested by Guttman (1978, 1987) and have since been refined in various ways by many authors.³ Barrett (1990), Falkinger *et al.* (1996), Rübbelke (2006), and Boadway *et al.* (2007, 2011) have applied matching particularly in the context of global environmental protection as an outstanding example of a global public good.

The basic idea underlying the matching approach is that the public good contributions of an agent, i.e. of a country in this paper, are subsidized by other agents. This reduces the cost to this agent of providing an additional unit of the public good and thus increases her incentives to contribute when matching rates are positive. But our model allows matching rates also to become negative, which means that the public good provision of a coalition is penalized within this group of countries, e.g. through a subsidy on fossil fuel use. Negative matching raises the individual marginal cost of public good provision, thus leading the coalition to reduce its collective contribution to the public good. In the strategic context considered here, this may benefit the coalition members by provoking a higher contribution from the non-members, on which the coalition then is able to take a free ride.

Our analysis of matching as the instrument of partial cooperation applies the Aggregative Game Approach developed by Cornes and Hartley (2003, 2007), which provides a very convenient tool in the search for equilibrium outcomes in the context of voluntary public good provision. In particular, we use this approach to complement the literature on coalition formation (e.g. Carraro and Siniscalco 1993; Barrett 1994; Hoel 1992; Hoel and Schneider 1997; Finus *et al.* 2005). However, we focus on the profitability of coalitions and not on their stability, which—as in the case without matching—has to be secured through separate mechanisms such as issue linkage (e.g. Barrett 1997) which lie outside of our model. But we briefly note that stability may be improved for not so small coalitions when the effect coined as ‘harmful cooperation’ in this paper is taken into account.⁴ We are also not concerned with the precise choice of domestic policy instruments that each country may use to achieve its mandated matching level, which may include domestic regulations, taxes and subsidies or carbon pricing.

We proceed as follows. Section I presents our model, in which there are two different groups of countries. One is a coalition of like-minded cooperating countries whose members are mutually matching their public good provision, and the other consists of outsiders that—without any matching—act non-cooperatively, playing Nash against each other and against the coalition. Section II applies the Aggregative Game Approach to determine the Nash equilibria at the second stage of the whole two-stage game when the matching rate within the cooperating coalition is exogenously given. Section III considers the strategically optimal choice of the matching rate through which the coalition maximizes utility of its members at the first stage of the game. We thus find the subgame-perfect equilibria of the two-stage game in which the matching rate is chosen at stage 1 and a Nash game in public good contributions is played at stage 2 in which the coalition and all outsiders are the players. In Section IV, as throughout the whole paper, we illustrate the major results of our analysis with a simple Cobb–Douglas example. As a specific result, we show that partial cooperation of a rather small number of countries may reduce public good supply as compared to the conventional Nash equilibrium without matching and in this sense it may be harmful for global environmental quality. The reason for this rather unexpected effect is—as will be shown by our analysis—that the matching mechanism works as a commitment device that helps the coalition to attain a leader position in

the game of public good provision. Moreover, multiple equilibria may arise in which public good supply and utility of the countries may differ extremely. Section V concludes.

I. THE MODEL

We consider the standard public good economy of Bergstrom *et al.* (1986) and Cornes and Sandler (1986), which we apply to the case of a global public good. There are n identical countries that all have the same utility function $u(x_i, G)$ and the same private good endowment ('income') w . By x_i we denote private consumption in country $i = 1, \dots, n$, and G is the quantity of a global public good (e.g. 'climate protection'). The utility function has the usual properties, i.e. it is strictly monotone increasing and twice continuously differentiable in both arguments, and both the private and the public good are assumed to be non-inferior. The technically given marginal rate of transformation between the public and the private good, and thus the marginal costs of greenhouse gas abatement, are assumed to be constant and can therefore be normalized to 1. Then the aggregate budget constraint comprising the incomes of all n countries reads

$$(1) \quad G = \sum_{i=1}^n (w - x_i) \quad \text{or} \quad G + \sum_{i=1}^n x_i = nw.$$

The essential assumption of our analysis now is that the whole world is exogenously divided into two groups. One subgroup, the coalition K , consists of $k \geq 2$ cooperating countries that jointly provide the public good. These countries may have a higher intrinsic ethical motivation to improve global environmental quality that is not reflected in cost–benefit calculations and is not part of our model. Alternatively, the coalition may have been more successful in establishing common institutions to enforce public good contributions within the coalition. In this case of enhanced political integration between the coalition members, the stability problems that preoccupy the literature on cooperative public good provision are absent from the outset since it is supposed that a central authority with some coercive power has been created. The process by which such a political federation is formed also lies outside our analysis.

As a specific feature of our model, partial cooperation among the participants of the coalition is portrayed through reciprocal matching of public good contributions within the coalition. The matching approach is quite common in the theory of global public goods. From a technical viewpoint, it helps to facilitate the analysis because it allows us to apply the Aggregative Game Approach and thus which is able to grasp, in a direct way, the change of the effective public good price that is implied by matching within the coalition. In an alternative formulation of the model, one could use the level of the contributions to the public good as the strategic variable on which the coalition decides. This would leave some of the results unchanged but would alter others. This point will be taken up later in some more detail.

The symmetric matching mechanism that underlies our analysis is characterized by the common (positive or negative) matching rate μ , and works as follows. If some country $i \in K$ has chosen g_i as its basic ('flat') contribution to the public good, then this expense is augmented by the other members of the coalition by μg_i if $\mu > 0$. If, however, there is negative matching with $\mu < 0$, then a coalition member's flat contribution is depleted by the other members of group K by μg_i . Irrespective of whether matching is positive or negative, the flat contribution g_i of country i thus induces

$(1 + \mu)g_i$ as public good contribution. The marginal rate of transformation between the private and the public good then becomes $1 + \mu$ for each member of the coalition, which equivalently means that each country $i \in K$ is confronted with the effective public good price ρ , where

$$(2) \quad \rho = \frac{1}{1 + \mu}.$$

Clearly, $\mu > 0$ (< 0) implies $\rho < 1$ (> 1).

To give an example for the functioning of such a matching mechanism, suppose that the matching rate is $\mu = 2$. If country i then gives up one unit of the private good to directly contribute to the public good, then this sacrifice in total will provide three extra units of the public good, two of which are indirectly contributed by the other countries in the coalition. The effective price of the public good that this country then has to pay is reduced to $\rho = 1/3$ through matching.

We could equally choose ρ or μ as the variable on which the outcomes produced by a matching scheme depend. To facilitate the exposition, throughout the paper we will use the effective public good price ρ that is induced by matching for each country in K , and will call it the ‘matching parameter’. This matching parameter coincides with the conventional Lindahl price $1/k$ for group K when the matching rate is $\mu = k - 1$ and there are no outsiders. This special case will be considered more closely in the next section.

Corresponding to the symmetric structure of the model, we moreover assume that the subsidy payment μg_i for country i is shared equally by the other coalition members except country i . Therefore if (g_1, \dots, g_n) is the vector of flat contributions, total contributions of country i amount to

$$z_i = g_i + \frac{\mu \sum_{j \in K, j \neq i} g_j}{k - 1},$$

where the second term gives the matching expenses that country i has to pay for the flat contributions of the rest of the coalition. In a symmetric solution where the aggregate and the flat contributions of all coalition members are identical and equal to z_K and g_K , respectively, we have $z_K = (1 + \mu)g_K$.

Note that for a negative matching rate $\mu < 0$, national public good contributions of coalition members are not subsidized but taxed or penalized. In the case of climate protection, such a penalty might consist in fossil fuel support policies, like, for example, a subsidization of fossil fuel inputs, for example through the central government of the federation, causing additional greenhouse gas emissions. In OECD countries, among the public fossil fuel support policies are ‘direct subsidies, intervention in markets in ways that affect costs or prices, assumption of a part of companies’ financial risks, tax reductions or exemptions, and under-charging for the use of government-supplied goods, services or assets’ (OECD 2011b, p. 3).⁵ From the perspective of the ‘Green Paradox’ (Sinn 2008), political measures that increase supply of fossil fuels in the coalition (e.g. through subsidization of new technologies for fossil fuel use like ‘fracking’) could also be included here.

The other subgroup M of non-cooperating outsiders consists of $m = n - k \geq 1$ countries that independently play Nash. Their effective public good price is the technically given marginal rate of transformation between the public and the private good that has been assumed to be 1.

II. MATCHING EQUILIBRIA

In this section we apply the *Aggregative Game Approach* (Cornes and Hartley 2003, 2007) to determine public good supply $\hat{G}(\rho)$ at a *Nash equilibrium under partial matching* (NEPM) when the effective public good price for the coalition members is changed to ρ through matching. This approach relies heavily on the use of income expansion paths. So, given preferences $u(x_i, G)$ and a public good price π , let $e(G, \pi)$ be the income expansion path along which the marginal rate of substitution between the public and the private good is equal to π . We assume that these expansion paths are defined on $[0, \infty)$ with $e(0, \pi) = 0$, that $\lim_{G \rightarrow \infty} e(G, \pi) = \infty$ for any $\pi > 0$, and that $\lim_{\pi \rightarrow \infty} e(G, \pi) = \infty$ for any $G > 0$. Our assumptions on utility functions imply that the income expansion paths are continuously differentiable and strictly monotone increasing.

To give an example, consider the case of Cobb–Douglas preferences, i.e. $u(x_i, G) = x_i^\alpha G$, where $\alpha > 0$ indicates the preference intensity for the private relative to that for the public good. For any public good price π , the income expansion path then becomes

$$(3) \quad e(G, \pi) = \pi\alpha G.$$

With the help of income expansion paths, we can now describe three types of *potential* NEPM. In these allocations, either both groups K and M (type a) or only one of these groups (type b or type c) actively contribute to the public good.

Type a: potential interior Nash equilibria

Candidates for an interior NEPM, in which all countries in K and M make strictly positive flat contributions to the public good, are characterized by a special version of the aggregate budget constraint (1), i.e.

$$(4) \quad G^a(\rho) = k(w - e(G^a(\rho))) + m(w - e(G^a(\rho), 1)),$$

where $G^a(\rho)$ denotes public good supply in a potential NEPM of type a (see, for example, Cornes and Hartley (2007) for a general elaboration of the Aggregative Game Approach, and Buchholz *et al.* (2011) for an application to matching). The level of the public good as defined by (4) differs from $\hat{G}(\rho)$ if the true NEPM is a corner solution with zero contributions from either insiders or outsiders.

Condition (4) provides the necessary condition for an interior NEPM since in this case all coalition members face the public good price $\pi = \rho$ as altered by matching, while the public good price for the outsiders is still $\pi = 1$. This implies that to have an interior NEPM, the position of each coalition member must lie on the income expansion path $e(G, \rho)$, while that of each outsider is on the income expansion path $e(G, 1)$. Otherwise, given the public good price with which a country inside or outside the coalition is confronted, it would have an incentive either to raise or to lower its flat contribution to the public good so that no Nash equilibrium could prevail. Then each country in K has private consumption $x_K^a(\rho) = e(G^a(\rho))$ and thus in total (through its own flat contribution and through matching the flat contributions of the other coalition members) spends $z_K^a(\rho) = w - x_K^a(\rho)$ for the public good. Each country in group M , however, has private consumption $x_M^a(\rho) = e(G^a(\rho), 1)$ and contributes $z_M^a(\rho) = w - x_M^a(\rho)$ to the public good. The aggregate contributions of both groups then make up for public good supply.

For all income levels $w > 0$, existence of a public good level $G^a(\rho)$ that satisfies (4) is implied by the Intermediate Value Theorem since $\lim_{G \rightarrow 0} e(G, \pi) = 0$ and $\lim_{G \rightarrow \infty} e(G, \pi) = \infty$ for any effective public good price π . Uniqueness is ensured by the strict monotonicity of the income expansion paths, which follows from strict normality of the underlying preferences.

For $\rho = 1$, equation (4) characterizes the standard Nash equilibrium with voluntary public good provision as a special case where utility $u^a(1) = u_K^a(1) = u_M^a(1)$ is the same for insiders and outsiders.

To infer how a change of the matching parameter ρ alters an allocation of type a , we use Figure 1, whose right (left) part describes the position of a coalition member (an outsider). In this diagram, a reduction of ρ (from ρ' to ρ'') implies that the income expansion path of a coalition member rotates upwards while the income expansion path of an outsider remains unchanged. Observing the aggregate budget constraint (4), it follows that the new position of an outsider after the fall of ρ must lie to the left of the old one, i.e. public good supply and private consumption of an outsider both increase. Otherwise, $G^a(\rho'') < G^a(\rho')$ and thus $x_M^a(\rho'') < x_M^a(\rho')$ and $x_K^a(\rho'') < x_K^a(\rho')$ (by the shift of the income expansion path) would imply $G^a(\rho') + kx_K^a(\rho'') + mx_M^a(\rho'') < (k + m)w$. It is then a further implication of the budget constraint that private consumption of a coalition member must decrease.

These comparative statics results are summarized in the following proposition.

Proposition 1 Public good supply $G^a(\rho)$ and private consumption of the outsiders $x_M^a(\rho)$ rise, and private consumption of the coalition members $x_K^a(\rho)$ falls, if the matching parameter ρ in the cooperating coalition is reduced.

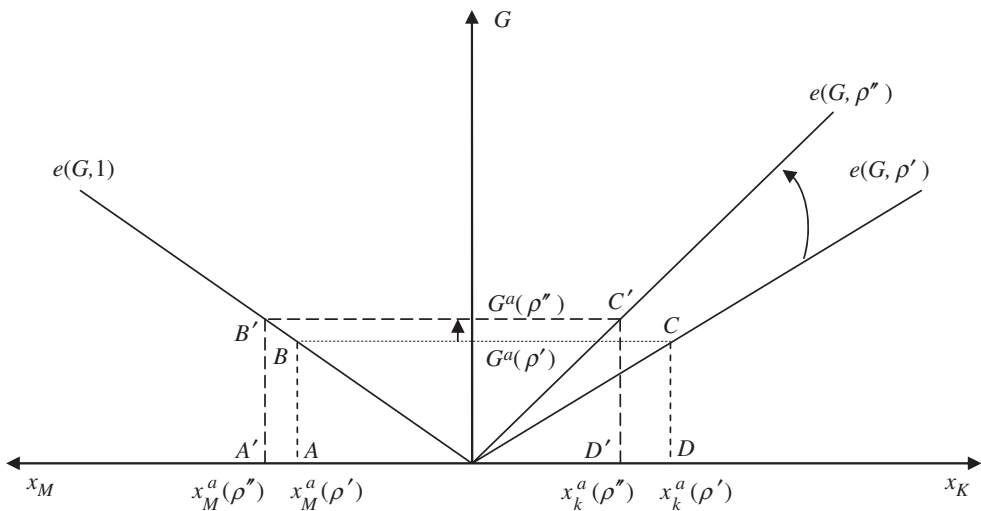


FIGURE 1. The effects of a changing matching parameter on a type a allocation.

As a next step we analyse how utility $u_K^a(\rho) := u(x_K^a(\rho), G^a(\rho))$ of a coalition member is changed by a variation of the matching parameter ρ . To facilitate notation, we adopt the following abbreviations for marginal changes of income expansion paths and utility levels as effected by a marginal change of the matching parameter:

$$\begin{aligned}
e_{K1}(\rho) &:= \frac{\partial e(G, \rho)}{\partial G} \Big|_{G=G^a(\rho)} > 0, & e_{K2}(\rho) &:= \frac{\partial e(G, \rho)}{\partial \rho} \Big|_{G=G^a(\rho)} > 0, \\
e_{M1}(\rho) &:= \frac{\partial e(G, 1)}{\partial G} \Big|_{G=G^a(\rho)} > 0, \\
u_{K1}^a(\rho) &= \frac{\partial u(x, G)}{\partial x} \Big|_{(x, G)=(x_K^a(\rho), G^a(\rho))} > 0, & u_{K2}^a(\mu) &= \frac{\partial u(x, G)}{\partial G} \Big|_{(x, G)=(x_K^a(\rho), G^a(\rho))} > 0
\end{aligned}$$

For the marginal effect of a variation of the matching parameter on the utility, we then have

$$\begin{aligned}
\frac{\partial u_K^a(\rho)}{\partial \rho} &= \frac{\partial u(e(G^a(\rho), \rho), G^a(\rho))}{\partial \rho} \\
&= u_{K1}^a(\rho) \left(e_{K1}(\rho) \frac{\partial G^a(\rho)}{\partial \rho} + e_{K2}(\rho) \right) + u_{K2}^a(\rho) \frac{\partial G^a(\rho)}{\partial \rho}.
\end{aligned}$$

Observing that $\rho u_{K1}^a = u_{K2}^a$ and applying (4), it follows that

$$(5) \quad \frac{\partial G^a(\rho)}{\partial \rho} = - \frac{ke_{K2}(\rho)}{1 + ke_{K1}(\rho) + me_{M1}(\rho)}.$$

Equation (5) directly implies the following result.

Proposition 2 If the matching parameter is marginally varied, then $\partial u_K^a(\rho)/\partial \rho < 0$ (> 0 , $= 0$) if and only if

$$(6) \quad \rho k > 1 + me_{M1}(\rho) \quad (< 0, = 0)$$

By $\tilde{\rho}^a = (1 + me_{M1}(\tilde{\rho}^a))/k$, we denote the levels of ρ for which equality in (6) holds. Clearly, maximization of utility $u_K^a(\rho)$ of coalition members requires $\rho = \tilde{\rho}^a$. In the Cobb–Douglas case—by applying (3)—condition (6) becomes

$$(7) \quad \rho k > 1 + m\alpha \quad (< 0, = 0)$$

Conditions (6) and (7) can be interpreted as follows. A unilateral increase of the public good contributions of coalition K , which follows from positive matching $\rho < 1$, induces a reduction of the contributions made by the outsider group M . This crowding-out effect is small if the group K is large while M is small, which makes partial matching profitable for the coalition K . If the income expansion path $e(G, 1)$ is rather steep at $G = G^a(\rho)$, i.e. if $e_{M1}(\rho)$ is rather small, then the outsiders' reactions to variations in public good contributions made by the coalition are rather weak. This implies that offsetting by the outsiders is not too strong, which favours positive matching in the coalition. Especially in the Cobb–Douglas case, it is a high preference for the public good, i.e. a low level of α , which gives a steep income expansion path $e(G, 1)$.

In the matching context, these results reflect those in Buchholz *et al.* (1998). Essentially, they would also hold if the coalition did not choose the level of the matching parameter but directly chose the level of its public good contribution.

Type b: potential Nash equilibria with coalition members as the sole contributors

By $G^b(\rho)$ we denote public good supply in the ‘standalone’ matching equilibrium of coalition K given the matching parameter ρ . Applying the Aggregative Game Approach, public good supply $G^b(\rho)$ is defined by $G^b(\rho) + k e(G^b(\rho), \rho) = kw$, which by letting $m = 0$ represents a special case of (4). For private consumption in such an allocation we have $x_K^b(\rho) = e(G^b(\rho), \rho)$. In an x_r - G -diagram in which the position of a country in K is depicted, the point $(x_K^b(\rho), G^b(\rho))$ lies on the budget line with slope $-1/\rho$ passing through $(w, 0)$. It is a straightforward implication of normality that public good supply $G^b(\rho)$ is a decreasing function of the matching parameter ρ . In allocations of type b , both public good supply and utility of the coalition members increase when the coalition size k increases.

Applying Proposition 2 to the case $m = 0$ directly yields $\tilde{\rho}^b k = 1$ and also shows that utility $u_K^b(\rho) = u(x_K^b(\rho), G^b(\rho))$ of a country in K is maximized if $\tilde{\rho}^b = 1/k$ since utility is increasing in ρ if $\rho < 1/k$ and decreasing if $\rho > 1/k$. This solution is the standalone Lindahl equilibrium for coalition K with $\tilde{\rho}^b = 1/k$ as the Lindahl price of each country.

Type c: potential Nash equilibria with outsiders as the sole contributors

By G^c we denote public good supply in the standard non-cooperative standalone Nash equilibrium that would be attained by the outsider group M if the cooperating coalition K were to contribute nothing to the public good. With help from the Aggregative Game Approach, G^c is defined by $G^c + m e(G^c, 1) = mw$, which is a special case of equation (4) letting $k = 0$. Each outsider then has private consumption

$$x_M^c = e(G^c, 1)$$

and attains utility $u_M^c = u(x_M^c, G^c)$. The members of coalition K then have private consumption $x_K^c = w$, which gives them utility $u_K^c = u(w, G^c)$. In allocations of type c , which—from their definition—are clearly independent of the level of the matching parameter ρ , public good supply and utility of insiders and outsiders is increased when the number m of outsiders grows.

Having described these three different types of potential NEPM, we now show how it depends on the level of the matching parameter ρ whether the NEPM is of type a , b , or c . To that purpose we define two threshold levels $\underline{\rho}$ and $\bar{\rho}$ for the matching parameter ρ by letting

$$(8a) \quad x_M^a(\underline{\rho}) = e(G^a(\underline{\rho}), 1) = w,$$

$$(8b) \quad x_K^a(\bar{\rho}) = e(G^a(\bar{\rho}), \bar{\rho}) = w.$$

Looking at Figure 1 these threshold levels can be explained as follows. If, starting from $\rho = 1$, the matching parameter ρ is reduced, then private consumption of the

outsiders in a type a allocation increases so that their public good contribution falls. Finally, if ρ has fallen to $\underline{\rho}$, then the outsiders stop making positive contributions to the public good. Conversely, if ρ becomes larger than $\rho = 1$, then private consumption of the coalition members increases, and consequently their public good contributions decrease. At $\rho = \bar{\rho}$ it is the coalition members who would no longer contribute to the public good. Since $x_M^a(\rho) < w$ may hold for all $\rho < 1$, it is possible that a strictly positive lower threshold $\underline{\rho}$ does not exist. For the sake of completeness we set $\underline{\rho} = 0$ in this case. (A formal proof for the existence and uniqueness of $\underline{\rho}$ and $\bar{\rho}$ is given in online Appendix A1.⁶)

Using these two threshold levels, the NEPM can now be characterized as follows. If ρ lies between $\underline{\rho}$ and $\bar{\rho}$, then we have $x_K^a(\rho) < w$ and $x_M^a(\rho) < w$, which leads to an NEPM of type a , while $x_M^a(\rho) > w$ holds for all $\rho < \underline{\rho}$ and $x_K^a(\rho) > w$ for all $\rho > \bar{\rho}$. At $\rho = \underline{\rho}$ all outsiders stop to make a positive contribution to the public good in an allocation of type a , and the NEPM then shifts from type a to type b . At $\rho = \bar{\rho}$, however, the contributions of the coalition members in the allocation of type a fall to zero, and the NEPM shifts from type a to type c . At $\rho = \underline{\rho}$, the type a allocation coincides with the type b allocation, and for $\rho \geq \bar{\rho}$ with the type c allocation, i.e. $\bar{G} := G^a(\underline{\rho}) = G^b(\underline{\rho})$ and $\underline{G} := G^a(\bar{\rho}) = G^c$.

Even though our previous considerations make these assertions quite intuitive, a precise proof is needed to show that the allocations of types a , b and c satisfy the conditions required for an NEPM. In the following, private consumption of a coalition member in the NEPM is denoted by $\hat{x}_K(\rho)$, and private consumption of an outsider is denoted by $\hat{x}_M(\rho)$ if the matching parameter is ρ .

Proposition 3

- (i) If $\rho \leq \underline{\rho}$, then the unique NEPM is an allocation of type b with $\hat{x}_K(\rho) = x_K^b(\rho)$, $\hat{x}_M(\rho) = w$ and $\hat{G}(\rho) = G^b(\rho)$.
- (ii) If $\rho \in (\underline{\rho}, \bar{\rho})$, then the unique NEPM is an allocation of type a with $\hat{x}_K(\rho) = x_K^a(\rho)$, $\hat{x}_M(\rho) = x_M^a(\rho)$ and $\hat{G}(\rho) = G^a(\rho)$.
- (iii) If $\rho \geq \bar{\rho}$, then the unique NEPM is of type c with $\hat{x}_K(\rho) = w$, $\hat{x}_M(\rho) = x^c$ and $\hat{G}(\rho) = G^c$.

Proof

- (i) Let again $z_K^b(\rho) = w - x_K^b(\rho)$ denote the aggregate public good contribution (i.e. the flat contribution plus the subsidy payments to the contributions of the other countries) of each country in K when the matching parameter is ρ . Then, being confronted with the symmetric matching mechanism, each member of the coalition K will choose the flat contribution $\hat{g}_K(\rho) = \rho z_K^b(\rho)$ as its best response to zero contributions by the outsiders to attain a position on the expansion path $e(G, \rho)$. Since $\rho \leq \underline{\rho}$ gives $G^b(\rho) \geq \bar{G}$ and thus $e(G^b(\rho), 1) \geq w$, the optimal reaction of each outsider is to contribute nothing to the public good if the coalition provides $G^b(\rho)$. This shows that in this case an NEPM is reached in an allocation of type b . This NEPM clearly is unique when public good contributions of the outsiders are zero. To complete the uniqueness proof we therefore assume that there is some other NEPM of type a or c where public good supply is $\hat{G}'(\rho)$ and the outsiders make a positive contribution to the public good. Then $\hat{G}'(\rho) < \bar{G}$ or, equivalently, $e(\hat{G}'(\rho), 1) < w$ is required to induce positive contributions by the outsiders. But as

- $e(\hat{G}^i(\rho), \rho) < e(\bar{G}, \rho)$, total public contributions in this hypothetical NEPM would even exceed \bar{G} , so the budget constraint (1) would be violated.
- (ii) If $\rho \in (\rho, \bar{\rho})$, then it follows from the monotonicity of the income expansion paths that $x_K^a(\rho) < w$ and $x_M^a(\rho) < w$. Then if each country in K chooses the public good contribution $\hat{g}_K(\rho) := \rho z_K^a > 0$ (with $z_K^a(\rho) = w - x_K^a(\rho)$) and each country in M chooses $\hat{g}_M(\rho) := z_M^a(\rho)$ (with $z_M^a(\rho) = w - x_M^a(\rho)$), the insiders attain a position on the income expansion path $e(G, \rho)$ while the outsiders' position is on $e(G, 1)$. Since the aggregate budget constraint is satisfied, $\hat{g}^K(\rho)$ and $\hat{g}^M(\rho)$ are mutually best responses so that an interior NEPM of type a with public good supply $G^a(\rho)$ and private consumption levels $x_K^b(\rho)$ and $x_M^a(\rho)$ for insiders and outsiders would arise. This is clearly the unique interior NEPM in this case. Now assume that there is an alternative NEPM of type b in which public good supply is $G^i(\rho)$ and only the coalition contributes to the public good. Then $\hat{G}^i(\rho) > G^a(\rho)$ must hold, as otherwise the outsiders would still have an incentive to contribute to the public good. As the coalition members in this NEPM were situated on $e(G, \rho)$ and $e(\hat{G}^i(\rho), \rho) > e(G^a(\rho), \rho) = x_K^a(\rho)$ holds, the contributions of coalition K would be smaller than $G^a(\rho)$, and the aggregate budget constraint (1) would be violated. The possibility that only the outsiders contribute to the public good in an NEPM of type c can be excluded in a similar way.
- (iii) The proof is analogous to that of part (i). \square

We now consider the two-stage game in which the members of the coalition cooperatively determine the matching rate at stage 1 and then play the non-cooperative Nash game at stage 2.

III. SUBGAME-PERFECT EQUILIBRIA

Based on Proposition 3, we will now consider the choice of the matching parameter ρ made by coalition K at stage 1 of a two-stage game. The coalition commits to this matching parameter, which then determines the NEPM obtained at stage 2. At stage 1 the coalition seeking to maximize utility of its members determines through its choice of ρ where on the combined reaction curve of the non-members it would like to locate. This strategically optimal matching rate characterizes the subgame-perfect equilibrium of the two-stage game.

In this context we impose an additional assumption **SP** by which it is postulated that the utility of the coalition members, as a function of the matching parameter ρ , has a single peak in an allocation of type a . This condition is fulfilled, for example, when the countries have Cobb–Douglas preferences, but also holds in the more general case where the slopes of the income expansion paths do not change too much when ρ varies (see online Appendix A2 for details).

Condition SP. There is a unique matching parameter for which the condition $\tilde{\rho}^a = (1 + me_{M1}(\tilde{\rho}^a))/k$ is satisfied. Utility $u_K^a(\rho)$ of the coalition members is strictly monotone increasing for $\rho < \tilde{\rho}^a$ and strictly monotone decreasing for $\rho > \tilde{\rho}^a$.

In order to determine subgame-perfect equilibria, we first compare utility of coalition members in allocations of type a and of type b . For this auxiliary technical result it is not important whether an allocation of type a or b will arise as the NEPM.

Proposition 4 Given **SP**, the matching parameter that maximizes $u_K^a(\rho)$ is greater than the matching parameter that maximizes $u_K^b(\rho)$, i.e. $\tilde{\rho}^a > \tilde{\rho}^b = 1/k$. Moreover, $u_K^b(\rho) > u_K^a(\rho)$ ($< 0, = 0$) if and only if $\rho < \underline{\rho}$ ($> \underline{\rho}, = \underline{\rho}$).

Proof The first part of the proposition follows because $e_{M1}(\rho) > 0$ implies

$$\tilde{\rho}^a = \frac{1 + me_{M1}(\tilde{\rho}^a)}{k} > \frac{1}{k} = \tilde{\rho}^b.$$

Concerning the second part, we write the aggregate budget constraint for an allocation of type a as

$$(10) \quad kx_K^a(\rho) + G^a(\rho) = kw + m(w - x_M^a(\rho)).$$

If $\rho < \underline{\rho}$ and thus $x_M^a(\rho) > w$, then we have $kx_K^a(\rho) + G^a(\rho) = ke(G^a(\rho), \rho) + G^a(\rho) < kw$. Since $ke(G, \rho) + G$ is increasing in G and $kx_K^b(\rho) + G^b(\rho) = ke(G^b(\rho), \rho) + G^b(\rho) = kw$, it follows that $G^b(\rho) > G^a(\rho)$. Consequently, $x_K^b(\rho) = e(G^b(\rho), \rho) > e(G^a(\rho), \rho) = x_K^a(\rho)$ and thus $u_K^b(\rho) = u(x_K^b(\rho), G^b(\rho)) > u(x_K^a(\rho), G^a(\rho)) = u_K^a(\rho)$. The other parts of the assertion are proven in a similar way. \square

The intuition behind the second part of this result is as follows. If $\rho < \underline{\rho}$, then—in an allocation of type a —the outsiders make a negative contribution to the public good at the cost of the coalition. In an allocation of type b where the coalition stands alone, this exploitation by the outsiders is avoided. Conversely, if $\rho > \underline{\rho}$, then the outsiders also contribute to the public good in an allocation of type a , which makes the coalition members better off than in a standalone solution of type b .

There are three candidates for the matching parameter ρ^* that maximizes utility of coalition K . Either $\rho^* = \tilde{\rho}^a$ when an interior solution is attained at stage 2, or $\rho^* = 1/k$ if only the coalition contributes to the public good, and $\rho^* = \bar{\rho}$ if only the outsiders contribute to the public good. Now precise conditions for the occurrence of the different types of Nash equilibria with partial matching will be provided. First we deal with the case in which the coalition benefits from some positive matching and thus a marginal increase of its public good contribution. Here, only two of the three possible solutions emerge as subgame-perfect equilibria.

Proposition 5 Let condition **SP** be fulfilled and assume $\tilde{\rho}^a \leq 1$. The strategically optimal matching parameter ρ^* for coalition K then is:

- (i) $\rho^* = 1/k$ if either $\tilde{\rho}^a \leq \underline{\rho}$, or $\tilde{\rho}^a > \underline{\rho} > 1/k$ and $u_K^b(1/k) > u_K^a(\tilde{\rho}^a)$;
- (ii) $\rho^* = \tilde{\rho}^a$ if either $\tilde{\rho}^a > \underline{\rho} > 1/k$ and $u_K^b(1/k) < u_K^a(\tilde{\rho}^a)$, or $\underline{\rho} \leq 1/k$.

Proof If $\tilde{\rho}^a < 1$, then **SP** implies that the coalition will not want to choose a matching parameter $\rho > 1$. Therefore we can restrict attention to matching parameters $\rho \leq 1$. The proof of the two parts of the proposition now is as follows.

- (i) If $\tilde{\rho}^a \leq \underline{\rho}$ and since **SP** holds, the coalition increases utility of its members in an allocation of type a by lowering ρ until $\underline{\rho}$ is reached where the NEPM switches from type a to type b . It follows from Proposition 4 that at $\underline{\rho}$, utility $u_K^b(\rho)$ must increase if ρ is

further reduced, which gives $\underline{\rho} \geq 1/k$. Therefore the NEPM for $\rho = 1/k$ is of type b , which then will be chosen at stage 1. If instead $\tilde{\rho}^a \geq \rho > 1/k$, the coalition can attain the best NEPM of type a and b , respectively. If then $u_K^b(1/k) > u_K^a(\tilde{\rho}^a)$, the coalition will also choose $\rho^* = 1/k$ at stage 1.

- (ii) If $\underline{\rho} > 1/k$, the proof is analogous to the second part of (i). If $\underline{\rho} \leq 1/k$, Proposition 4 combined with **SP** implies that $u_K^b(\rho) < u_K^b(\underline{\rho}) = u_K^a(\underline{\rho}) \leq u_K^a(\tilde{\rho}^a)$ holds for all $\rho \leq \underline{\rho}$ and thus for every attainable NEPM of type b , which proves the assertion. \square

In Proposition 5, essentially three cases are distinguished. If $\underline{\rho} \geq \tilde{\rho}^a$, the NEPM switches to an allocation of type b before the utility maximum in an allocation of type a is reached. Then it is immediately clear from Proposition 5 that the optimal choice for the coalition is $\rho^* = 1/k$. If, however, $\tilde{\rho}^a > \underline{\rho} > 1/k$, the NEPM that give the coalition members maximal utility among the allocations of types a and b are both attainable, and the coalition's optimal choice depends on the comparison between the utility levels $u_K^a(\tilde{\rho}^a)$ and $u_K^b(1/k)$. Finally, if $\underline{\rho} < 1/k$, the allocation that would maximize utility of the coalition members among all allocations of type b is not attainable. Moreover, it follows from Proposition 4 that in all attainable allocations of type b , utility is lower than $u_K^a(\tilde{\rho}^a)$, which implies $\rho^* = \tilde{\rho}^a$ in this case.

In the second case, $\tilde{\rho}^a > 1$, some negative matching entails a reduction of the coalition's public good contributions, which benefits the coalition members. Then all three candidates of type a , b or c allocations may constitute the subgame-perfect equilibrium, and harmful effects of partial cooperation may arise.

Proposition 6 Let condition **SP** be fulfilled, and assume $\tilde{\rho}^a > 1$. The strategically optimal matching parameter ρ^* for coalition K then is:

- (i) $\rho^* = \tilde{\rho}^a$ if $\tilde{\rho}^a < \bar{\rho}$ and either $u_K^a(\tilde{\rho}^a) > u_K^b(1/k)$ or $\underline{\rho} \leq 1/k$;
- (ii) $\rho^* = \bar{\rho}$ if $\tilde{\rho}^a \geq \bar{\rho}$ and either $u_K^c > u_K^b(1/k)$ or $\underline{\rho} \leq 1/k$;
- (iii) $\rho^* = 1/k$ otherwise.

Proof Condition **SP** implies that among all matching parameters $\rho < \underline{\rho}$ the coalition maximizes utility of its members in the second stage equilibrium by choosing $\rho = \tilde{\rho}^a$ when $\tilde{\rho}^a \leq \bar{\rho}$ or by choosing $\rho = \bar{\rho}$ when $\tilde{\rho}^a > \bar{\rho}$. The coalition will clearly make these choices at stage 1 when its best allocation of type b , which is reached by setting $\rho = 1/k$, can be attained as an NEPM but utility is smaller there. In the case $\underline{\rho} < 1/k$, however, it follows from **SP** and Proposition 4 that in all attainable type b allocations, utility is lower, than $u_K^a(\tilde{\rho}^a)$, and also smaller than u_K^c if $\bar{\rho} \leq \tilde{\rho}^a$. This gives (i) and (ii). If $1/k < \underline{\rho}$, the optimal allocation of type b can be reached, and it will be chosen when it gives the coalition members higher utility than $u_K^a(\tilde{\rho}^a)$ or u_K^c , respectively. \square

There is some asymmetry between the two cases described in Propositions 5 and 6: If $\tilde{\rho}^a < 1$, i.e. if the coalition would benefit from marginally positive matching, Proposition 5 says that also in the subgame-perfect equilibrium the coalition definitely will choose a positive matching rate and will make higher contributions to the public good than in the standard Nash equilibrium without matching. But, as seen from Proposition 6, positive matching in the subgame-perfect equilibrium, i.e. $\rho^* = 1/k$, also shows up as a possible outcome if $\tilde{\rho}^a > 1$, i.e. if marginally negative matching would be in the interest of the coalition.

In both cases treated in Propositions 5 and 6, multiple equilibria may arise if the relevant utility levels coincide, e.g. if $u_K^b(1/k) = u_K^a(\bar{\rho}^a)$ in the situation described in Proposition 5(i).

It is a direct consequence of Proposition 1 that public good supply in a subgame-perfect equilibrium, which results in cases (i) and (ii) of Proposition 6, is lower than in the conventional Nash equilibrium without matching, which aggravates the underprovision problem. Then partial cooperation within coalition K is harmful from both the perspective of global public good provision and the welfare position of non-cooperating countries, while utility of the coalition members clearly is higher than in the Nash equilibrium without matching. This result reminds us of a paradoxical effect described by Hoel (1991) in a different framework: unilateral action of a country or a group of countries from which at first sight more greenhouse gas abatement should be expected may in the end have the opposite effect, i.e. entail further environmental degradation.

If preferences and the income level are fixed, the size k of the coalition for which such undesirable effects may result in the subgame-perfect equilibrium is limited irrespective of the size m of the outsider group M . This is shown by the next proposition, which also does not depend on whether $\bar{\rho}^a \leq 1$ or $\bar{\rho}^a > 1$.

Proposition 7 Given a utility function $u(x, G)$ and an income level w , there exists some critical level \underline{k} for the coalition size so that $\rho^* = 1/k$ for all $k \geq \underline{k}$ and all $m \geq 1$.

Proof For all $m \geq 1$, public good supply in an NEPM of type a or type c is bounded above by the level \bar{G} for which $e(\bar{G}, 1) = w$ holds, since otherwise the outsiders would not contribute to the public good in a Nash equilibrium (see Andreoni (1988) for a similar argument on conventional Nash equilibria without matching). Therefore if $\rho > 1$, then in any NEPM of type a or type c , utility of a coalition member will be smaller than $u(w, \bar{G})$. Now choose some $\underline{k} \geq 2$ for which $u^b(1/\underline{k}) > u(w, \bar{G})$ holds. This condition is clearly fulfilled if the budget line through $(w, 0)$ with slope $-1/\underline{k}$ intersects the indifference curve through (w, \bar{G}) . For any $k \geq \underline{k}$ and all $m \geq 1$, then $G^b(1/k) > \bar{G} = G^b(\underline{\rho})$. Hence $1/k < \underline{\rho}$ and $\rho^* = 1/k$. \square

Note that harmful effects of partial cooperation do not occur if—in a one-stage game—the countries in coalition K played Nash against the outsiders by collectively choosing their public good contributions. For an explanation, assume that G_M are the aggregate public good contributions made by the outsiders. By $g^r(G_M)$ we denote the public good contribution of each country in K when all coalition members react in an isolated and identical manner to G_M , and by $\tilde{g}_K^r(G_M)$ the individual contribution of a country in the coalition if the coalition collectively determines its optimal response with identical contributions of all of its members. Clearly,

$$g^r(G_M) = \arg \max_{g \geq 0} u(w - g, G_M + g)$$

$$\tilde{g}_K^r(G_M) = \arg \max_{g \geq 0} u(w - g, G_M + kg).$$

By applying the Aggregative Game Approach again, it can be demonstrated that $\tilde{g}_K^r(G_M) \geq g^r(G_M)$ holds for all $G_M > 0$ and that public good supply in the Nash equilibrium with partial cooperation cannot be less than public good supply in the

standard non-cooperative Nash equilibrium. Thus harmful cooperation is avoided if the coalition members use public good contributions as strategic variables.

The reason for the difference between these two instruments of cooperation is that internal matching acts as a kind of commitment device for the coalition, which puts the coalition in the position of a Stackelberg leader.⁷ That commitment may generate a strategic advantage is a well-known possibility in the industrial organization literature, where, for example, it is described how firms can improve their position in oligopolistic competition by committing to certain technologies (see, for example, Dixit 1980), and it has also been dealt with in other fields of public good theory (see, for example, Beccherle and Tirole 2011). In our context this commitment effect also implies that the coalition would prefer a matching mechanism over a collective decision on public good contributions. Even though determination of the levels of public good contributions might seem to be a more direct and simpler form of partial cooperation, it entails a strategic disadvantage for the coalition.

Harmful cooperation also may have consequences for the stability of coalitions. Consider first the case of harmful cooperation treated in Proposition 6(ii), where $u_K^c = u(w, G^c) > u^a(1) > u(x_M^c, G^c) = u_M^c$ and $u^a(1) (= u_K^a(1) = u_M^a(1))$ again denotes utility of all countries in the Nash equilibrium without matching. After a country has left the original coalition, the NEPM continues to be of type c . On the one hand, G^c and thus utility of a coalition member in a type c allocation increase when the outsider group becomes larger. On the other hand, in an allocation of type b , which would be the alternative choice for the coalition, utility of a coalition member is reduced when the coalition becomes smaller. In the new NEPM of type c , the deviating country attains a utility level that lies below $u^a(1)$ and is thus lower than its utility as a member of the former coalition. Therefore any coalition that brings about harmful cooperation is internally stable—which is a rather unwelcome stability result.

But the threat of harmful cooperation also may, under certain circumstances, improve the stability of beneficial coalitions. The reason is that a coalition that has chosen positive matching in the initial NEPM may be induced to shift to the harmful strategy when one country leaves. Then the deviating country would attain a utility level as an outsider that is smaller than the level in the Nash equilibrium without matching and hence also smaller than its utility as a coalition member in the original NEPM. Anticipating this decline in utility will deter any member of the original coalition from leaving, which ensures internal stability.

This effect of harmful cooperation that—unlike the standard case—might also stabilize comparatively large coalitions deserves some further analysis; however, that lies outside the scope of this paper.⁸ An example in which the threat of harmful cooperation is able to stabilize a coalition of size $k = 11$ in the public good framework is provided in the online Appendix.

IV. THE COBB–DOUGLAS EXAMPLE

The general results obtained in the previous section will now be illustrated for the Cobb–Douglas case, where $u(x_i, G) = x_i^\alpha G$. In the allocations of types a , b and c , the coalition members then attain the utility levels

$$(9) \quad u_k^a(\rho) = (\rho\alpha)^\alpha \left(\frac{(k+m)w}{(k\rho+m)\alpha+1} \right)^{\alpha+1},$$

$$(10) \quad u_K^b(\rho) = (\rho\alpha)^{\alpha} \left(\frac{k\omega}{k\rho\alpha + 1} \right)^{\alpha+1},$$

$$(11) \quad u_K^c = \frac{m\omega^{\alpha+1}}{m\alpha + 1}.$$

These utility levels, which will be used throughout our treatment of the Cobb–Douglas example, are calculated in online Appendix A3. Applying conditions (8a) and (8b) to the Cobb–Douglas case gives the threshold levels $\underline{\rho} = 1 - 1/k\alpha$ and $\bar{\rho} = 1 + 1/m\alpha$.

To determine the subgame-perfect equilibria, we must—as in Propositions 5 and 6—consider the two cases $\tilde{\rho}^a \leq 1$ and $\tilde{\rho}^a > 1$ separately. Observing (7), these conditions boil down to $k \geq m\alpha + 1$ and $k < m\alpha + 1$ in the Cobb–Douglas case. (For details, see online Appendix A4.)

Case 1. If $k \geq m\alpha + 1$, it follows from Proposition 5 that the subgame-perfect equilibrium is given by either $\rho^* = \tilde{\rho}^a = (m\alpha + 1)/k$ or $\rho^* = 1/k$.

An intricate analysis in the online Appendix shows that in the case $\alpha \geq 1$, i.e. when private consumption weighs not less in utility than the public good, $\rho^* = 1/k$ always holds. If, however, $\alpha < 1$, it is possible that $\rho^* = \tilde{\rho}^a$, which results if $\underline{\rho} < \tilde{\rho}^a$ (or equivalently $k < m\alpha + 1 + 1/\alpha$) holds and $u_K^b(1/k) < u_K^a(\tilde{\rho}^a)$ (which follows from comparing the expressions in (9) and (10)).

Case 2. If $k < m\alpha + 1$, then it follows from Proposition 6 that $\rho^* = \tilde{\rho}^a = (m\alpha + 1)/k$, $\rho^* = 1 + 1/m\alpha$ or $\rho^* = 1/k$.

Among all $\rho > 1$, the coalition will prefer $\rho = \tilde{\rho}^a = (m\alpha + 1)/k$ if $\tilde{\rho}^a < \bar{\rho}$, or equivalently $k > m\alpha$, and $\rho = \bar{\rho} = 1 + 1/m\alpha$ if $k \leq m\alpha$. If then—as given by a comparison between (9), (10) and (11)—either $u_K^a(\tilde{\rho}^a)$ or u_K^c exceeds $u_K^b(1/k)$, harmful cooperation will occur in the subgame-perfect equilibrium.

Harmful cooperation with $\rho^* = \bar{\rho} = 1 + 1/m\alpha$ arises, for example, if $\alpha = 1$, $k = 2$ and $m \geq 3$. But—as shown in the online Appendix— $\rho^* = \bar{\rho}$ is possible only if the coalition has no more than three members while the number of outsiders may be very large and even go to infinity.

Unlike Case 1, the interior outcome $\rho^* = \tilde{\rho}^a = (m\alpha + 1)/k$ now can also result even if $\alpha > 1$. This is shown by the example $\alpha = 1.1$, $k = 2$ and $m = 1$, where $\rho^* = (1.1 + 1)/2 = 1.05$. The possibility of having $\rho^* = \tilde{\rho}^a$ is, however, quite restricted since it requires $k \leq 3$ and $m \leq 2$ (see again the online Appendix).

Further interesting observations are made by looking more closely at the specific example with $\alpha = 1$ and $k = 3$. Here, the matching parameter which characterizes the subgame-perfect equilibrium is made dependent on the number m of outsiders and is denoted by $\rho^*(m)$.

If $m = 1$ or $m = 2$, we are in the situation described by Case 1 and $\rho^*(1) = \rho^*(2) = \frac{1}{3}$. The NEPM is of type b with $G^b(\frac{1}{3}) = \frac{3}{2}w > w$, $x_K^b(\frac{1}{3}) = \frac{1}{2}w$, $x_M^b(\frac{1}{3}) = w$ and $u_K^b(\frac{1}{3}) = \frac{3}{4}w^2$.

For any $m \geq 3$, however, we are in the situation of Case 2. If, specifically, $m = 3$ by choosing $\tilde{\rho}^a = \bar{\rho} = \frac{4}{3}$, the coalition attains an allocation of type c with $G^c = \frac{3}{4}w$, $x_K^c(\frac{4}{3}) = w$, $x_M^c(\frac{4}{3}) = \frac{3}{4}w$ and $u_K^c(\frac{4}{3}) = \frac{4}{3}w^2$. Hence the coalition is indifferent between choosing $\rho = \frac{4}{3}$ leading to an allocation of type c , and $\rho = \frac{1}{3}$ leading to an allocation of type b . This gives rise to multiple subgame-perfect equilibria: On the one hand there is a ‘good’

equilibrium with $\rho^*(3) = \frac{1}{3}$ where public good supply and utility of outsiders are higher than in the conventional Nash equilibrium without matching, while on the other hand there is a ‘bad’ equilibrium with $\rho^*(3) = \frac{4}{3}$ where partial cooperation is harmful and public good supply and utility of the outsiders are lower than in the conventional Nash equilibrium.

If $m > 3$, we have $\rho^*(m) = \bar{\rho} = (m + 1)/m$. Public good supply in the subgame-perfect equilibrium then is $G^c = mw/(m + 1) < w$. Now assume that initially there are $m = 1$ or $m = 2$ outsiders. If the number of outsiders then jumps to any $m \geq 2$, the level of public good supply in the subgame-perfect equilibrium falls from $G^b(\frac{1}{3}) = \frac{3}{2}w$ to a level below w . This effect is against intuition, which suggests that increasing the size of a public good economy will automatically raise public good supply. It is the strategic effect coined as harmful partial cooperation that is responsible for this paradoxical effect.

V. CONCLUSIONS

In this paper we have in a two-stage public good game investigated the implications of partial cooperation for the provision of global public goods. In this context we have added three novel elements to the existing contributions. So we have done the following.

- We considered matching mechanisms as the instrument for cooperation within a coalition of like-minded countries. Matching is by now seen to be a potentially appealing way to improve the level of international environmental protection and generates a strategic advantage for the coalition as compared to the simpler strategy of fixing contribution levels.
- We applied the Aggregative Game Approach to characterize matching equilibria in flat contributions at the second stage of the game. Thereby we linked the theory of coalition formation with some recent developments in public good theory.
- We explored the coalition’s optimal strategy for setting the matching rate at a first stage, paying special attention to corner solutions at the second stage where either the outsiders or the coalition members do not contribute to the public good.

In particular, we have shown that the two types of corner solution are of a very different quality. If only the coalition contributes to the public good, then partial cooperation generates a Pareto improvement as compared to the standard voluntary provision equilibrium: public good supply and utility of all countries are increased. If, however, only the outsiders contribute in the subgame-perfect equilibrium, then partial cooperation has harmful effects: although the utility of coalition members is increased, public good supply and utility of the outsiders fall as compared to the standard Nash equilibrium without matching. In the context of global warming, the coalition lowers its public good contribution (by increasing greenhouse gas emission through subsidization of fossil fuels) in order to induce the outsiders to increase their contributions. This result provides some caveat against the commonly shared belief that any form of international cooperation will be beneficial for climate protection and global environmental quality.

Depending on the specific circumstances, it is possible that the cooperating coalition’s strategically optimal choice of the matching parameter at the first stage of the game steers the economy into either of these two corner solutions. Harmful partial cooperation thus may occur in the subgame-perfect equilibrium.

Further results of our analysis, which we have demonstrated through examples in the Cobb–Douglas case, are that multiple subgame-perfect equilibria may arise and

that—paradoxically—public good supply in the subgame-perfect equilibrium may decrease when the number of countries is increasing.

Moreover, it is precisely the presence of a small coalition that generates the unfavourable outcome. So in the Cobb–Douglas example when the private good is valued higher than the public good, the harmful effect is definitely excluded as soon as the coalition size exceeds 3. This might provide an additional reason why, even if—in the spirit of the bottom-up approach—pragmatic reasons motivate a particular interest in partial cooperation among a small number of countries, it is still important to keep in mind the possibilities afforded by larger coalitions. It is solely large coalitions that definitely avoid the harmful effects of cooperation and ensure a Pareto improvement over the standard voluntary provision equilibrium.

ACKNOWLEDGMENTS

This paper is part of the research project ‘RECAP 15—Re-thinking the Efficacy of International Climate Agreements after COP 15’ financed by the German Federal Ministry of Education and Research BMBF (grant number: 01LA1139A). Financial support from BMBF research grant 01LA11043 is also gratefully acknowledged. Richard Cornes holds the F. H. Gruen Chair of Economics at the Australian National University. The authors thank the participants of the conference on ‘New Directions in the Voluntary Provision of International Public Goods’ at the ZEW Mannheim, 17–18 April 2012, and the participants of the 19th EARE conference in Prague in June 2012 for their helpful comments.

NOTES

1. With respect to international environmental agreements, Finus and Tjøtta (2003) find that grand coalitions do not necessarily bring about the social optimum, and recommend analysing the formation of subcoalitions.
2. With some reference to public good theory and matching, see also the more recent review of the literature on fiscal federalism in Oates (2005). An application to international public goods is in Buchholz *et al.* (2013).
3. See Boadway *et al.* (1989), Danziger and Schnytzer (1991), Althammer and Buchholz (1993), Varian (1994), Andreoni and Bergstrom (1996), Falkinger (1996), Kirchsteiger and Puppe (1997), Falkinger and Brunner (1999), and Buchholz *et al.* (2011, 2012).
4. For a general treatment of this phenomenon without any reference to the standard public good model, see Beaudry *et al.* (2000).
5. In a recent study, the OECD has identified more than 250 individual mechanisms that effectively support fossil fuel production or consumption in the considered 24 OECD countries (OECD 2011a, p. 3).
6. Details of the online Appendix are available from the authors on request.
7. For some general treatment of the differences between the Nash and the Stackelberg scenarios, see Mallozzi and Tijs (2012).
8. See Eichner and Pethig (2013) for a treatment of this issue in a Stackelberg model that incorporates international trade.

REFERENCES

- ALTHAMMER, W. and BUCHHOLZ, W. (1993). Lindahl-equilibria as the outcome of a non-cooperative game. *European Journal of Political Economy*, **9**, 399–405.
- ANDREONI, J. (1988). Privately provided public goods in large economies: the limits of altruism. *Journal of Public Economics*, **35**, 37–73.
- and BERGSTROM, T. C. (1996). Do government subsidies increase the private supply of public goods? *Public Choice*, **88**, 295–308.
- BARRETT, S. (1990). The problem of global environmental protection. *Oxford Review of Economic Policy*, **6**, 68–79.
- (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, **46**, 878–94.
- (1997). The strategy of trade sanctions in international environmental agreements. *Resource and Energy Economics*, **19**, 345–61.

- BEAUDRY, P., CAHUC, P. and KEMPF, H. (2000). Is it harmful to allow partial cooperation? *Scandinavian Journal of Economics*, **102**, 1–21.
- BECCHERLE, J. and TIROLE, J. (2011). Regional initiatives and the cost of delaying binding climate change agreements. *Journal of Public Economics*, **95**, 1339–48.
- BERGSTROM, T. C., BLUME, L. and VARIAN, H. (1986). On the private provision of public goods. *Journal of Public Economics*, **29**, 25–49.
- BOADWAY, R., PESTIEAU, P. and WILDASIN, D. (1989). Tax-transfer policies and the voluntary provision of public goods. *Journal of Public Economics*, **39**, 157–76.
- , SONG, Z. and TREMBLAY, J.-F. (2007). Commitment and matching contributions to public goods. *Journal of Public Economics*, **91**, 1664–83.
- , ——— and ——— (2011). The efficiency of voluntary pollution abatement when countries can commit. *European Journal of Political Economy*, **27**, 352–68.
- BUCHHOLZ, W., CORNES, R. C. and RÜBBELKE, D. (2011). Interior matching equilibria in a public good economy: an aggregative game approach. *Journal of Public Economics*, **95**, 639–45.
- , ——— and ——— (2012). Matching as a cure for underprovision of voluntary public good supply. *Economics Letters*, **117**, 727–9.
- , HASLBECK, C. and SANDLER, T. (1998). When does partial co-operation pay? *FinanzArchiv/Public Finance Analysis*, **55**, 1–20.
- , HAUPT, A. and PETERS, W. (2013). International environmental agreements, fiscal federalism, and constitutional design. *Review of International Economics*, **21**, 705–18.
- CARRARO, C. and SINISCALCO, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, **52**, 309–28.
- CORNES, R. C. and SANDLER, T. (1986). *The Theory of Externalities, Public Goods and Club Goods*. Cambridge: Cambridge University Press.
- and HARTLEY, R. (2003). Risk aversion, heterogeneity and contests. *Public Choice*, **117**, 1–25.
- and ——— (2007). Aggregative public good games. *Journal of Public Economic Theory*, **9**, 201–19.
- DANZIGER, L. and SCHNYTZER, A. (1991). Implementing the Lindahl voluntary-exchange mechanism. *European Journal of Political Economy*, **7**, 55–64.
- DIXIT, A. (1980). The role of investment in entry deterrence. *Economic Journal*, **90**, 95–106.
- EICHNER, T. and PETHIG, R. (2013). Self-enforcing environmental agreements and international trade. *Journal of Public Economics*, **102**, 37–50.
- FALKINGER, J. (1996). Efficient private provision of public goods by rewarding deviations from average. *Journal of Public Economics*, **62**, 413–22.
- and BRUNNER, J. K. (1999). Taxation in an economy with private provision of public goods. *Review of Economic Design*, **4**, 357–79.
- , HACKL, F. and PRUCKNER, G. J. (1996). A fair mechanism for efficient reduction of global CO₂-emissions. *FinanzArchiv/Public Finance Analysis*, **53**, 308–31.
- FINUS, M., ALTAMARINO-CABRERA, J. C. and van IERLAND, E. C. (2005). The effect of membership rules and voting schemes on the success of international climate agreements. *Public Choice*, **125**, 95–127.
- FINUS, M. and TJØTTA, S. (2003). The Oslo Protocol on sulfur reduction: the great leap forward? *Journal of Public Economics*, **87**, 2031–48.
- GUTTMAN, J. M. (1978). Understanding collective action: matching behavior. *American Economic Review*, **68**, 251–5.
- (1987). A non-Cournot model of voluntary collective action. *Economica*, **54**, 1–19.
- HOEL, M. (1991). Global environmental problems: the effect of unilateral actions taken by one country. *Journal of Environmental Economics and Management*, **20**, 55–70.
- (1992). International environmental conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics*, **2**, 141–59.
- and SCHNEIDER, K. (1997). Incentives to participate in an international environmental agreement. *Environmental and Resource Economics*, **9**, 153–70.
- KIRCHSTEIGER, G. and PUPPE, C. (1997). On the possibility of efficient private provision of public goods through government subsidies. *Journal of Public Economics*, **66**, 489–504.
- OATES, W. E. (1972). *Fiscal Federalism*. New York: Harcourt Brace Jovanovich.
- OECD (2005). Toward a second-generation theory of fiscal federalism. *International Tax and Public Finance*, **12**, 349–73.
- (2011a). *Inventory of Estimated Budgetary Support and Tax Expenditures for Fossil Fuels*. Paris: OECD.
- (2011b). *Joint Report by IEA, OPEC, OECD and World Bank on Fossil-fuel and Other Energy Subsidies: An Update of the G20 Pittsburgh and Toronto Commitments*, Prepared for the G20 Meeting of Finance

- Ministers and Central Bank Governors (Paris, 14–15 October 2011) and the G20 Summit (Cannes, 3–4 November 2011); available online at www.oecd.org/env/49090716.pdf (accessed 17 December 2013).
- MALLOZZI, L. and TIJS, S. (2012). Stackelberg assumption vs. Nash assumption in partially cooperative games. *Czech Economic Review*, **6**, 5–13.
- RÜBBELKE, D. (2006). An analysis of an international environmental matching agreement. *Environmental Economics and Policy Studies*, **8**, 1–31.
- SINN, H.-W. (2008). Public policies against global warming: a supply side approach. *International Tax and Public Finance*, **15**, 360–94.
- VARIAN, H. (1994). Sequential contributions to public goods. *Journal of Public Economics*, **53**, 165–86.
- VICARY, S. (2012). Is three a crowd? Small groups and the provision of a public good. Mimeo.