Environ Resource Econ (2017) 67:591–607
DOI 10.1007/s10640-016-9997-2

Successful Leadership in Global Public Good Provision: Incorporating Behavioural Approaches

Wolfgang Buchholz[1,2] · **Todd Sandler**[3]

Accepted: 31 December 2015 / Published online: 30 January 2016

Abstract In the standard model of voluntary public good provision and other game theoretic models, climate-friendly leadership of a country is not successful: A unilateral increase of this country's greenhouse gas abatement measures, i.e., contributions to the global public good of climate protection, will not lead to a positive reaction by the other countries but instead trigger a reduction of their abatement efforts and thus a crowding-out effect. In this paper it is shown how this undesired consequence need no longer occur when elements of behavioral economics are incorporated in the otherwise standard model of public good provision. In particular, strategic complementarities between the public good contribution of the leading country and those of the follower may result either if the follower has specific non-egoistic or other-regarding preferences or if the leader's contribution positively affects the follower's beliefs, i.e., his conjectural variations, about the leader's behaviour.

Keywords Climate protection · Voluntary public good provision · Other-regarding preferences · Conjectural variations

JEL Classification C72 · D03 · H41 · H87 · Q54

Wolfgang Buchholz gratefully acknowledges financial support from the German Federal Ministry of Education and Research BMBF through the research projects RECAP 15 (FKZ 01LA1139A) and ECCUITY (FKZ 01LA1104B). Todd Sandler received support from the Vibhooti Shukla endowment at the University of Texas at Dallas. Lisa Dippl, Michael Eichenseer and Dirk Rübbelke have given helpful comments on previous versions of this paper. We gratefully acknowledge the comments and suggestions of two referees that helped improve the paper.

✉ Wolfgang Buchholz
wolfgang.buchholz@ur.de

Todd Sandler
tsandler@utdallas.edu

1 University of Regensburg, Regensburg, Germany

2 CESifo Munich, Munich, Germany

3 University of Texas at Dallas, Richardson, TX, USA

1 Introduction

From its very beginning the demand for leadership has played an important role in global climate politics. Already in 1992, in relation to the normative principle of "common but differentiated responsibility", it was postulated in Article 3 of the United Nations Framework Convention on Climate Change (UNFCCC) that "…, the developed country parties should take the lead in combating climate change and the adverse effects thereof." Originally, attributing a leading role to industrialized countries was mainly motivated by distributional motives since these rich countries have a higher capacity to bear the burden of mitigation measures and, at the same time, have been responsible for the major share of greenhouse gas emissions in the past. But, in the spirit of the "bottom-up-approach", the postulate for leadership is also based on the expectation of a "motivational push" caused by the leaders: By adopting a pioneering role and by setting a good example, these countries should induce other countries to follow suit in the provision of the global public good "climate protection". In practice the leadership idea has become prominent in the EU, and especially in Germany, since it serves as a justification for far-reaching activities in climate and energy policy, e.g., for the 20:20:20 goal and for the subsidisation of renewable and also nuclear energy in European countries.

According to the theory of public goods that has become an important tool for the theoretical analysis of global climate policy (e.g., Sandler 1997; Barrett 2005, pp. 1462–1465; Cornes 2015), a positive effect from leadership is not the normal case. Rather, the standard model of voluntary public good provision predicts that increased public good contributions by the leaders have a negative effect on the public good contributions of the followers. Against the background of this pessimistic view, the main objective of this paper is to show how, by enriching the standard model of voluntary public good provision with elements of behavioural economics, positive effects on the followers' public good contributions may result so that leadership becomes successful, which in theoretical terms means that the follower's reaction path is upward sloping and the contributions of both countries are strategic complements.

The paper is organized as follows: In Sect. 2, we provide a review of the existing theoretical literature on the effects of leaders' behaviour on followers going beyond the public good model. In Sect. 3, we present and interpret various types of psychological preferences as different kinds of inequality aversion, a desire for reciprocity and impure altruism in the sense of warm-glow-of-giving. In Sect. 4, we then show how, in an otherwise standard static two-country model of voluntary public good provision, the inclusion of such psychological components into the follower's preferences may entail a positively sloped reaction function of the follower, thus making leadership successful. In Sect. 5, we consider a public goods model with conjectural variations to show, how increased public good contributions by the leader may influence the expectations of the follower so that he is motivated to increase his public good contributions. In Sect. 6, these results are illustrated by an example. Finally, in Sect. 7, we briefly discuss some implications for climate policy, which follow from our theoretical analysis.

2 Leadership in Public Good Provision from the Perspective of Economic Theory: A Short Review of the Literature

From the perspective of standard models which describe public good provision by independent agents, the prospects for successful leadership appear rather poor: In a simple one-shot game with a discrete strategy space, the cooperation problem is usually modelled as a pris-

oners' dilemma or a chicken game (e.g., Lipnowski and Maital 1983; Cornes and Sandler 1996, pp. 305–319; with application to international environmental problems Finus 2001, pp. 21–41; DeCanio and Fremstad 2013). In both of these games, a unilateral cooperative effort by a leader is met with non-cooperation by the follower. In the model of voluntary public good provision with a continuous strategy space (Bergstrom et al. 1986; Cornes and Sandler 1996), a self-interested Stackelberg leader will choose a smaller public good contribution than that chosen in a simultaneous Nash game, everything else being equal. The reason is that strategic behaviour enables the leader to free-ride even more on the contributions of the follower so that the underprovision problem is aggravated (Varian 1994). A negative outcome of the sequential game becomes even more likely if it is the agent with the higher preferences for the public good who moves first (Nosenzo and Sefton 2011). In this public-good scenario, the leader does not act as a pioneer for better climate protection, but uses his first-mover advantage in an environmentally detrimental way.

If, instead, an altruistically motivated country unilaterally increases its public good contributions, the standard model of private public good provision predicts that the other countries will react not by increasing, but by reducing their abatement efforts (e.g., Cornes and Sandler 1996; Buchholz et al. 1998)—which, as a special type of the carbon leakage phenomenon, has been vividly described by Sinn (2012, p. 143) as "grabbing from the collection box". In theoretical terms, such an offsetting behaviour of the outsiders means that reaction curves are negatively sloped and the agents' public good contributions are strategic substitutes. Thus, a crowding-out effect arises and part of the leader countries' increased mitigation efforts evaporates. The crowding-out effect may be reinforced when climate change damages are uncertain and the followers are sufficiently risk-averse (Auerswald et al. 2011). In a similar way, a crowding-out effect results in the non-cooperative scenario when a country takes unilateral measures to develop and improve "green" technologies as, e.g., through subsidization of renewable energy (Buchholz and Konrad 1994; Ihori 1996).

But also if the public good is provided cooperatively, total public good supply may fall after a country's unilateral increase of its public good contribution (or a unilateral R&D-measure), since a unilateral action changes the threat point of a Nash bargaining game. Unless the level of public good supply is the same for all Pareto optimal solutions, public good supply in the bargaining outcome may decrease through the shift of the threat point, which is the famous paradox of Hoel (1991)[1]. For unilateral R&D-activities a similar result was obtained by Buchholz and Konrad (1994). Intuitively speaking, a pioneering country may lose the power to push through its ambitious environmental goals when it acts unilaterally before negotiations start. In case of uncertain abatement costs, it also becomes less likely that an efficient bargaining outcome is attained if one country has unilaterally committed to higher mitigation efforts (Konrad and Thum 2014).

Furthermore, according to Sinn's "green paradox" (e.g., Sinn 2012),[2] producers of fossil fuels may counteract the leading countries' attempts to curb CO_2 emissions by shifting supply from the future to the present by lowering current fuel prices, which represents an intertemporal leakage effect. Interregional leakage effects, caused by a fall of the world

[1] Conditions on preferences, which ensure constant public good supply along the utility possibility frontier, are described by Bergstrom and Cornes (1981). See also Cornes and Sandler (1996, pp. 199–201). If these conditions are satisfied, the Hoel paradox is excluded.

[2] The occurrence of this effect hinges on several conditions: e.g., the exact design of climate policy or the ability of producers to increase the supply of fossil fuels. Moreover, a crucial, but contestable, assumption underlying the "green paradox" is that all resources in the ground will finally be dug out and exhausted. For an encompassing review of the "green paradox" and its various versions, see the collection of essays in Pittel et al. (2014).

market prices of fossil fuels or by a shift of carbon-intensive industries to countries with a less ambitious climate policy, can also neutralize abatement activities. The cost of leadership may be magnified through structural changes in the economy, de-industrialization, depreciation of capital and job losses, so that pioneering countries may further lose bargaining power.

The danger of ineffective or even detrimental pioneering activities is particularly great if the leading coalition is small since then the crowding-out effect caused by the outsiders is strong (Buchholz et al. 1998) or, in case of supply-side reactions according to the "green paradox", a larger group of outsiders can more easily absorb the additional supply of fossil fuels.

Theoretical economics, however, also provides some reasons for a more positive assessment of leadership in public good provision in general and climate protection in particular (see also, e.g., section 4.2 in Edenhofer et al. 2015). So, a one-shot prisoners' dilemma game may turn into a stag hunt or an assurance game (e.g., Fehr and Schmidt 2009) if certain fairness preferences prevail in the follower country (Pittel and Rübbelke 2012) that either reflect a positive valuation of reciprocal behaviour (Rabin 1993; Dufwenberg and Kirchsteiger 2004) or a preference for a more equal distribution of the game's payoffs (Fehr and Schmidt 1999; Bolton and Ockenfels 2000). In the context of climate change, an assurance game may result from specific abatement technologies, risk preferences or perceptions of the climate problem, e.g., if the follower country is sufficiently risk averse and believes that it can prevent a climate catastrophe by "topping" the efforts of the leader (e.g., Endres and Ohl 2001; Buchholz et al. 2014). In a repeated prisoners' dilemma game, a large initial public good contribution by some country may initiate continual cooperation, which is, e.g., implemented by a tit-for-tat strategy (Axelrod 1984; specifically in the context of climate change see Barrett 2005, pp. 1489–1493). In an evolutionary game setting, positive reactions by the follower can also be triggered through learning effects (e.g., De Oliveira et al. 2005).

In a world of uncertainty, pioneering efforts of leading agents may reveal the costs and benefits of a public good to other agents, which reduces their uncertainty about the effects of their own efforts and thus increases their willingness to contribute to the public good (Hermalin 1998; with application to climate change Brandt 2004). Irreversible public good contributions by the leader may also relieve the follower of the fear that he is cheated by the leader and thus helps to make subsequent communication easier and leadership more successful (Barbieri 2012).

3 Different Types of Psychological Preferences

While in the standard public good model (with a summation technology for public good provision and normality of the public good in the follower's preferences) crowding-out results, some changes of this model may entail strategic complementarity and thus lead to crowding-in. Then, to vary Sinn's metaphor, the follower's behaviour is characterized not by "grabbing from" but by "putting additional money into the collection box". It is well-known from the literature that a positive reaction by the follower can be generated if there is a non-summation technology for public good production (see Arce M 2001; Kempf and Rota Graziosi 2010), if contributions to the public good entail private "ancillary" benefits for the follower (e.g. Pittel and Rübbelke 2008), which is described by the "impure" public good model (Cornes and Sandler 1984a, 1994, 1996), or if the private good is inferior. We now extend the preference-related approach for explaining successful leadership by incorporating certain psychological motivations from behavioural economics into the follower's utility function.

Let country L be the leader (with income w_L and private consumption x_L) and country F be the follower (with income w_F and private consumption x_F) in a static game of voluntary public good provision. The public good contribution of country L is g_L and that of country F is g_F. As in the standard model, we suppose a summation technology for public good production, i.e., total public good supply is given by $G = g_L + g_F$, country L's private consumption is $x_L = w_L - g_L$ and that of country F is $x_F = w_F - g_F$. The preferences of country L are assumed to be only self-regarding by the utility function $u(x_L, G)$, which is twice differentiable with positive first derivatives, quasi-concave, and both the private and the public good are strictly normal.

What really matters for our analysis are the preferences $V(x_F, G, g_F, g_L)$ of the follower country F, in which the first two arguments x_F and G represent the standard utility components, while the additional behavioural or "psychological" preferences are captured by the third and fourth argument referring to country F's own public good contribution g_F and that of country L, i.e., g_L. This extended utility function V is assumed to be twice continuously differentiable in all four variables with V_1, V_2, V_3 and V_4 as its first-order partial derivatives and V_{ij} (with $i, j = 1, 2, 3, 4$) as its second-order partial derivatives for which our differentiability assumption gives symmetry, i.e., $V_{ij} = V_{ji}$. As in the standard model without psychological preferences, $V_1 > 0$ and $V_2 > 0$ clearly hold. For the second-order partial derivatives, we may assume that, as usual, $V_{11} < 0$ and $V_{22} < 0$. But also for the partial derivatives V_3 and V_4 of the psychological parts, a non-negative sign usually seems to be the most plausible assumption.

On the one hand, $V_3 \geq 0$ indicates that its own public good contribution provides additional psychological benefit for country F, which might be interpreted as some "warm-glow giving" in the sense of Andreoni (1990). It is plausible to assume decreasing marginal benefits also here, i.e., $V_{33} \leq 0$. On the other hand, $V_4 \geq 0$ can be interpreted as capturing that the people in country F experience some moral satisfaction if country L increases its public good contribution and, thus, is "doing its part" in financing the common project. A negative V_4, in contrast, then would mean that country F has envious or shameful feelings against larger public good contribution of country L, or would simply dislike "overachieving" by an overzealous leader. Such attitudes, however, are in contradiction with country F's material interests. The sign of the second partial derivative V_{44} is not straightforward to motivate, but it does not play any role for the results of this paper.

If $V_3 = 0$ (or $V_4 = 0$), we can abbreviate the utility function Vas $V(x_F, G, g_L)$ (or $V(x_F, G, g_F)$). If $V_3 = V_4 = 0$ we are back in the standard model without psychological utility components, while the case with $V_3 > 0$ and $V_4 = 0$ formally corresponds to Cornes and Sandler's (1994, 1996) impure public good model.

Various behavioural approaches suggest a more specific form for the utility function $V(x_F, G, g_F, g_L)$. So it is often supposed that the psychological utility component $\psi(g_F, g_L)$ enters the utility function V in an additively separable way, which gives

$$V(x_F, G, g_F, g_L) = v(x_F, G) + \psi(g_F, g_L). \tag{1}$$

If the extended preferences are of type (1), there is a clear distinction between material (or "egoistic" preferences) as described by $v(x_F, G)$ and the psychological preferences as described by $\psi(g_F, g_L)$, and $V_{13} = V_{14} = V_{23} = V_{24} = 0$ holds for the second order cross partial derivatives. Given (1) we have $V_1 = v_1$, $V_2 = v_2$, $V_3 = \psi_1$ and $V_4 = \psi_2$. Of special importance is the sub-case where

$$\psi(g_F, g_L) = \psi_m(g_F - g_L). \tag{2}$$

Then clearly $V_3 = -V_4 = \psi'_m$ holds. Two interpretations for (2) can be provided, both of which are related to specific fairness concepts: On the one hand, a utility function of type (1) may express inequality aversion of country F with respect to (w.r.t.) public good contributions. Then ψ'_m has no uniform sign: If $g_F > g_L$, an increase of g_F (or a decrease of g_L) increases inequality of public good contributions and thus reduces country F's utility so that $\psi'_m < 0$ holds. If, however, $g_F < g_L$, then $\psi'_m > 0$.

Similarly, but conceptually a little different, a utility function as described by (2) may indicate that country F has some intrinsic preference for reciprocity in the spirit of Rabin (1993).[3] With such a reciprocal attitude, country F has a bad conscience and thus experiences a utility loss if its own public good contribution falls below that of country L because it feels morally obliged to reciprocate country L's "kindness" as expressed through this country's public good contribution. If, however, its own public good contribution exceeds that of country L, then country F feels unfairly exploited by country L, which arouses a sentiment of anger in country F. With this interpretation, we again get $V_3 = \psi'_m < 0$ if $g_F > g_L$ and $V_3 = \psi'_m > 0$ if $g_F < g_L$. The feelings of bad conscience and anger, however, may have different welfare weights.

In a similar vein, but based on a precise interpretation of Rabin's (1993) approach, Nyborg (2014) has introduced a reciprocity component into the utility function, which essentially depends on benchmark values $\hat{g}_L$ and $\hat{g}_F$ for the public good contributions of country L and country F, respectively, and which leads to a psychological utility term of the form

$$\psi\left(g_F, g_L\right) = \psi_n\left(\left(g_F - \hat{g}_F\right) \cdot \left(g_L - \hat{g}_L\right)\right) \tag{3}$$

with $\psi'_n > 0$.[4] Given the reciprocity term $\psi_n((g_F - \hat{g}_F) \cdot (g_L - \hat{g}_L))$, a further increase of country F's public good contribution will increase country F's utility when both countries initially make contributions above their benchmarks, which can be interpreted as the fulfilment of some minimal obligations. But if country L is underachieving and violating its duties, then country F could reduce its anger about country L's unkind behaviour by reducing its own public good contribution. We later deal with Nyborg's utility function in more detail by discussing a specific example.

Country F's extended utility function V also reflects the case in which country L's utility directly affects utility of country F. Again assuming an additively separable form such 'other-regarding preferences' yield

$$V\left(x_F, G, g_F, g_L\right) = v\left(x_F, G\right) + \alpha u\left(w_L - g_L, G\right). \tag{4}$$

A positive value of the parameter α reflects altruistic preferences of country F, which means that country F shows a positive interest in the well-being of country L. Given a utility function of type (4), we have $V_1 = v_1$, $V_2 = v_2 + \alpha u_2$, $V_3 = 0$, and $V_4 = -\alpha u_1$.

There is also the possibility that (4) expresses inequality aversion of country F w.r.t. levels of material utility. Then α is positive or negative depending on the relation between the utility levels of leader L and follower F.[5] A positive (negative) α is obtained if in the initial allocation, country L's utility is smaller (larger) than that of country F. Then a

[3] For an encompassing discussion of the concept of preference-related "intrinsic" (in contrast to "instrumental") reciprocity, see Sobel (2005).

[4] Note that in the model of Nyborg (2014) the agents only face a dichotomous choice about contributing or not contributing to the public good.

[5] For special self-regarding utility functions, Kolstad (2012) explores the effects that such other-regarding preferences have on the Nash equilibrium of voluntary public good provision and on coalition formation in this context.

increase (decrease) of country L's self-regarding utility reduces inequality of these utility levels and leads to a rise of country F's extended utility (for different versions of psychological preferences with a concern for equality of outcomes see Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Charness and Rabin 2002; and the generalization by Benjamin 2015).[6]

4 Successful Leadership due to Follower's Psychological Preferences

Given a general extended utility function $V(x_F, G, g_F, g_L)$, we now want to determine in a unified way the conditions for successful leadership, i.e., for a positive slope of country F 's reaction function $g_F^r(g_L)$, defined by $g_F^r(g_L) = \arg\max_{g_F} V(w_F - g_F, g_F + g_L, g_F, g_L)$.[7] The first-order condition of this maximization problem is

$$\begin{aligned} &-V_1\left(w_F - g_F^r(g_L), g_F^r(g_L) + g_L, g_F^r(g_L), g_L\right) \\ &\quad + V_2\left(w_F - g_F^r(g_L), g_F^r(g_L) + g_L, g_F^r(g_L), g_L\right) \\ &\quad + V_3\left(w_F - g_F^r(g_L), g_F^r(g_L) + g_L, g_F^r(g_L), g_L\right) = 0. \end{aligned} \tag{5}$$

Omitting arguments and observing symmetry of second partial derivatives, we find that the second-order condition for maximization of V becomes $D := V_{11} + V_{22} + V_{33} - 2(V_{12} + V_{13} - V_{23}) < 0$, which is assumed to hold. Clearly, $V_{ii} < 0$ for $i = 1, 2, 3$ is favourable for the fulfilment of this condition.

Similar to Arce M and Sandler (2005, p. 17), implicit differentiation of (5) yields

$$\frac{\partial g_F^r}{\partial g_L} = \frac{V_{22} - V_{12} + V_{32} - V_{41} + V_{42} + V_{43}}{-D} \tag{6}$$

as the slope of country F's reaction path. From the second-order condition $D < 0$, it follows that the denominator of (6) is positive so that the sign of $\frac{\partial g_F^r}{\partial g_L}$ only depends on the sign of the numerator of (6). Thus, $\frac{\partial g_F^r}{\partial g_L} > 0$ if

$$V_{32} + V_{42} + V_{43} - V_{41} > V_{12} - V_{22}. \tag{7}$$

If condition (7) is satisfied in the initial allocation, starting from which country L unilaterally increases its public good contribution, then country F shows a positive reaction and leadership becomes successful. This outcome is more likely if the left-hand side of inequality (7) is large, i.e., if

- country F's psychologically motivated marginal valuations of its own and country L's contributions are much increased when public good supply increases (i.e., if $V_{32} + V_{42}$ is large).
- an increase of country L's public good positively affects the marginal psychological valuation of country F's own contribution, i.e., if $V_{34} = V_{43}$ is positive and large.
- an increase in country F's private consumption has a negative (or a positive but small) impact on country F's marginal valuation of country L's contribution, i.e., if V_{41} is small.

[6] To make this explicit, assume that the difference between the utility of the other country and its own utility directly affects extended utility of country F, i.e., $\tilde{V} = v + \beta(u - v)$. Inequality aversion of country F then is reflected by having $\beta < 0$ if $u > v$ and $\beta > 0$ if $u < v$ (and, e.g., $\beta = 0$ if $u = v$). If $\beta < 1$, the extended preferences of country F can, after a linear transformation, be represented by $V = v + \alpha u$ where $\alpha = \frac{\beta}{1-\beta}$.

[7] For an early discussion of successful leadership in the impure public good model, see Sandler (1996, esp. pp. 268–269).

In the case of additive separability according to (1), the interpretation of condition (6) becomes easier since $V_{32} = V_{42} = V_{41} = 0$ holds and (7) is reduced to

$$\psi_{12} = \psi_{21} = V_{43} > V_{12} - V_{22} = v_{12} - v_{22}. \tag{8}$$

When leadership starts from the standard Nash equilibrium of voluntary public good provision, we have $v_{12} - v_{22} > 0$.[8] Then condition (8) clearly shows the importance of strong reciprocal preferences for successful leadership: If $\psi_{12} = \psi_{21}$ is positive and large, country F is much more (less) willing to spend more on the public good when the other country spends more (less) on it.

Further specializing the utility function to type (2), condition (8) becomes

$$-\psi''_m > v_{12} - v_{22}, \tag{9}$$

which means that the negative marginal valuation of unequal public good contributions grows with the degree of inequality.

In the case of other-regarding preferences as described by (4), we get $V_{12} = v_{12}$, $V_{22} = v_{22} + \alpha u_{22}$, $V_{32} = V_{41} = V_{43} = 0$, and $V_{42} = -\alpha u_{12}$ so that the crowding-in condition (7) turns into $-\alpha u_{12} > v_{12} - (v_{22} + \alpha u_{22})$, which implies

$$\alpha\,(u_{22} - u_{12}) > v_{12} - v_{22}. \tag{10}$$

When, in addition, $u_{22} - u_{12} < 0$ holds (see also fn. 8 above), condition (10) is always satisfied if, in the initial allocation, the leader's material utility exceeds that of the follower and the follower's inequality aversion is sufficiently strong, i.e., α has a large negative value. If, however, α is positive, which occurs either when country F's material utility in the initial allocation exceeds that of country L or if country F is altruistic, condition (10) is not satisfied and crowding-out unavoidably results.

Finally, in the case of pure warm-glow-of-giving where clearly $V_{41} = V_{42} = V_{43} = 0$ holds, crowding-in results from condition (7) if

$$V_{32} > V_{12} - V_{22}. \tag{11}$$

This condition is satisfied if the warm-glow-of-giving effect increases sufficiently when public good supply increases. This effect can either be attributed to a pleasure from conforming with others' actions or by some reasoning such as: The greater is the aggregate contribution to the common good, the more I consider my own contribution as meaningful.

Apart from that, there is another possibility to explain the follower's positive reaction to an increased public good contribution of the leader in the case of warm-glow-of-giving. To this end we construct—for some given utility function V with $V_4 = 0$—a utility function for which the warm-glow-of-giving component no longer appears explicitly and which only depends on x_F and G. So, for all $x_F \in [0, w_F]$ and all $G \geq 0$, we define

$$\tilde{v}\,(x_F, G) = V\,(x_F, G, w_F - x_F)\,. \tag{12}$$

For this ancillary utility function, $\tilde{v}$ (which is of the same form as usually applied in public good theory), we suppose that it has the standard properties, i.e., that utility is increasing in both arguments, which specifically means that $\tilde{v}_1 = V_1 - V_3 > 0$, and that utility is quasi-concave. Furthermore, when the public good is non-inferior, the income expansion path, along which the marginal rate of substitution is equal to the marginal rate of transformation

[8] Given $\frac{v_1}{v_2} = 1$ it follows that $v_{12} - v_{22} = \frac{v_1}{v_2} v_{12} - v_{22} > 0$, where the second inequality follows from the normality assumption.

(= 1 as assumed), can be represented as a function of public good supply, i.e., as $x_F^e(G, 1)$. A short calculation shows that taking the derivative of $x_F^e(G, 1)$ w.r.t. G gives an expression, which has the same numerator as in (6) and whose denominator is positive. This implies that crowding-in (crowding-out) is entailed if $\frac{\partial x_F^e}{\partial G} < 0\,(>0)$ holds, i.e., if the private good is inferior (normal) for country F, given the ancillary utility function $\tilde{v}$.[9] This transformation of the impure public good scenario into a scenario with pure public goods gives a further explanation for the existence and uniqueness results obtained by Kotchen (2007) for impure public goods.

Note that in the warm-glow-of-giving scenario—unlike in the case of fairness preferences considered above—a positively sloped reaction function of country F can never occur when the psychological component additively enters the utility function V as in (1). Rather, it is only the cross derivative V_{32} that may cause crowding-in in this case and thus make leadership successful.

To illustrate some of the general results derived in this section, we consider several examples that build on country F's self-regarding utility function $v(x_F, G) = \ln x_F + \beta \ln G$, where $v_{12} = 0$ and $v_{22} = -\frac{\beta}{G^2}$. In the first two examples, the psychological component enters country F's extended preferences in an additive way as described by (1). Inequality aversion w.r.t. public good contributions according to (2) is reflected by

$$\psi_m\,(g_F - g_L) = -\frac{\alpha}{2}\,(g_F - g_L)^2\,. \tag{13}$$

Then, we have $V_{34} = -\psi_m'' = \alpha$ and condition (9) becomes

$$\alpha > \frac{\beta}{G^2}, \tag{14}$$

where G is public good supply in the original allocation for which the leader's unilateral increase of his public good contributions starts.

For Nyborg preferences, we specifically assume that

$$\psi_n\left(\left(g_F - \hat{g}_F\right) \cdot \left(g_L - \hat{g}_L\right)\right) = \alpha\left(g_F - \hat{g}_F\right)\left(g_L - \hat{g}_L\right). \tag{15}$$

Again, we have $V_{34} = \alpha$ so that it follows from (8) that country F's reaction path is positively sloped also if condition (14) is satisfied. According to this scenario, favourable conditions for successful leadership therefore are a large weight α of the psychological preference component, low preferences for the public good in the follower's material utility function and a large public good supply G in the initial allocation. If this allocation is the non-cooperative Nash equilibrium, condition (15) in particular implies that the chances for successful leadership are better when countries are rich since, everything else being constant, public good supply in the Nash equilibrium is increasing with the countries' income.

Concerning the case of other-regarding preferences with inequality aversion w.r.t. self-regarding utility levels, we assume that, analogous to country F, country L's utility function is $u(x_L, G) = \ln x_L + \gamma \ln G$, so that $u_{12} = 0$ and $u_{22} = -\frac{\gamma}{G^2}$. Condition (10) then becomes

$$-\alpha\gamma > \beta. \tag{16}$$

Therefore, in case of inequality aversion w.r.t. material utility levels and $\alpha < 0$, the prospects for successful leadership again are better when the follower has strong psychological preferences, but weak preferences for the public good. Furthermore, (16) shows that leadership is more likely to be successful when the leader has strong preferences for the public good.

[9] The relationship between impure public goods and inferiority of the private good has briefly been addressed by Cornes and Sandler (1985, pp. 128–129).

Finally, in the case of warm-glow-of-giving, we assume that country F's extended utility function is

$$V(x_F, G, g_F) = v(x_F, G) + \alpha g_F G, \tag{17}$$

which—by applying (11)—leads to the same condition for successful leadership as given by (14).

Activation of the follower's psychological preferences through a unilateral increase of the leader's public good contribution may also initiate political negotiations between the countries. An analysis of the dynamics of this negotiation process, however, lies outside the scope of the paper.

5 Successful Leadership due to an Induced Change of the Follower's Expectations

Crowding-in, and thus successful leadership, may also arise when larger public good contributions by the leader induce positive conjectures on the part of the follower about the leader's further reaction to marginal variations of the follower's public good contributions. By adapting an approach known from oligopoly theory, we can capture these beliefs for a static public good model with "conjectural variations", which indicate the marginal change b of the other agent's public good contribution that an agent expects if he marginally increases or decreases his own contribution.

Such belief parameters b may be positive or negative, depending on whether a parallel or a contrary reaction is expected. If $b = 0$, the follower does not suppose any influence of his own contribution on that of the leader, so that we are back to Nash behaviour. A special case for positive conjectural variations, which clearly reflects the idea of reciprocal behaviour, corresponds to "Kantian" behaviour, for which an agent chooses his optimal actions under the expectation that the other agents behave in a like manner, i.e., increase their public good contributions by the same relative or absolute amount (Roemer 2010, 2015).[10] Especially in relation to a leader-follower scenario, it has been shown in the conjectural variation scenario that public good supply at the Stackelberg equilibrium is even larger than in the standard Nash equilibrium if the follower has positive conjectural variations (Cornes and Sandler 1984b, 1996).

In the otherwise standard public good model with conjectural variations, crowding-out by the follower is never avoided when the follower's belief parameters b are exogenously given and fixed (which, as a by-product of our analysis, will be shown later on). This not so promising picture, however, changes when we allow for a specific endogenous determination of conjectural variations so that the leader is able to generate positive beliefs through larger public good contributions.

By modifying Cornes and Sandler's (1984b) approach, we make endogenous belief formation precise by assuming that the follower country F has a conjectural variation function $b(g_L, g_F)$, which is assumed to be partially differentiable in both variables with partial derivatives $b_1(g_L, g_F)$ and $b_2(g_L, g_F)$. Given the public good contributions g_L and g_F, country F thus expects that country L will change its public good contribution by $b(g_L, g_F)$ marginal units when country F increases its own public good contribution by one unit.

Concerning the sign of b_1 and b_2, we clearly may assume that $b_1 > 0$; i.e., a leader's public good contribution is perceived as an expression of goodwill (or "kindness" in the

[10] Some insightful hints at the relationship between Kantian behaviour and positive conjectural variations have already been given by Cornes and Sandler (1984b, p. 377) and Itaya and Okamura (2003, p. 56).

sense of Rabin 1993) and makes the follower believe that the leader is more willing to respond positively to an increase of the follower's contributions. In a certain sense, larger public good contributions by country L have a trust-building effect for country F.

The assumption $b_1 > 0$ captures recent insights in behavioural economics about the important role of beliefs and the formation of other agents' beliefs through own public good contributions (e.g., Croson 2007; Fischbacher and Gächter 2010). As put by Gächter and Renner (2014, p. 3): "Existing experimental evidence … suggests that people contribute on average more the more they believe others contribute. This observation implies that any factor that shifts beliefs will shift behaviour." Remarkably, long before the start of behavioural economics, Buchanan (1967), in an early contribution on voluntary public good provision in an economy with two agents called Tizio and Caio, already envisaged the possibility that "(b)y increasing … his own contribution, Tizio may hope that Caio will, …, follow suit and cooperate in response" (p. 113), thus deviating from the standard rationality assumption.[11]

The sign of b_2, however, is not determined a priori, but crucially depends on the reasoning upon which country F forms its beliefs: On the one hand, country F may expect that any increase of its own public good contribution will stimulate a greater positive reaction of country L. If country F believes in such a positive motivational effect, then we have $b_2 > 0$. On the other hand, country F may perceive the goodwill signal set by some g_L in relation to its own public good contribution g_F. Then, if g_F becomes larger, the positive signal flowing from a certain g_L becomes weaker, so that it reduces the follower's expectations and $b_2 < 0$ holds. Finally, if country F makes his conjectures about future reactions of country L only dependent upon country L's previous actions, $b_2 = 0$ is implied.

For a theoretical analysis of the effects that such endogenous belief have on the follower's behaviour, let—for any $x_F > 0$ and $G > 0$—$m(x_F, G) = \frac{v_1(x_F,G)}{v_2(x_F,G)}$ denote country F's marginal rate of substitution between the private and the public good at (x_F, G). If we assume normality of both the private and the public good, we clearly have $m_1 < 0$ and $m_2 > 0$ for the partial derivatives of m. An argument in the spirit of the Aggregative Game Approach (Cornes and Hartley 2007; Acemoglu and Jensen 2013; in relation to conjectural variations Possajennikov 2015) then implies that, given conjectural variations $b(g_L, g_F)$, country F can only be in an equilibrium position[12] when country L makes the public good contribution g_L, provided its own public good contribution g_F satisfies

$$1 + b\,(g_L, g_F) = m\,(w_F - g_F, g_F + g_L)\,. \tag{18}$$

If, however, the marginal rates of substitution and the perceived marginal rates of transformation were different, then country F—by anticipating a reaction of county L according to the conjectural variations given by $b(g_L, g_F)$—would expect to attain a greater utility level by some change of its public good contribution. With $b = 0$, the standard model of voluntary public good model would be obtained from (18) as a special case.[13]

[11] Another brief hint at the potential relevance of beliefs for reciprocal ("matching") behaviour already appears in Guttman (1978, pp. 254–255), where even some reference to early public good experiments is made.

[12] As always, such a static equilibrium may be considered as a reduced (and more easily tractable) form of a complex dynamic game (Cabral 1995; Itaya and Okamura 2003, pp. 52–53). Scafuri (1992) and Itaya and Okamura (2003) show how, either through learning or through the application of punishment strategies, conjectural variations may emerge in a repeated game.

[13] For a detailed discussion of the relationship between the conjectural variations model and the standard public good model, see Dasgupta and Itaya (1992). For a special focus on Warr neutrality, i.e., on the invariance of the equilibrium w.r.t. an income redistribution, see Sandler and Posnett (1991) and Buchholz (1993).

Taking the total derivative of (18) and again omitting all arguments yields the slope of the follower's reaction function as

$$\frac{\partial g_F^r}{\partial g_L} = \frac{b_1 - m_2}{m_2 - m_1 - b_2}. \tag{19}$$

If country F's conjectural variation is constant so that $b_1 = b_2 = 0$, (19) gives $\frac{\partial g_F^r}{\partial g_L} < 0$, which confirms the crowding-out effect, known from the standard public good model. But taking the behavioural effects into account, we see that (19) implies that country F may respond positively to an increased public good contribution of country L. This outcome, in which leadership of country L is successful, occurs if and only if the numerator and the denominator of (19) have the same sign. Concerning the conditions for $\frac{\partial g_F^r}{\partial g_L} > 0$, two cases are distinguished:

(i) $b_1 > m_2$ and $b_2 < m_2 - m_1$. In this case, the trust-building signalling effect provided by country L's public good contribution is strong, while country F's belief in the motivational effect triggered by its own public good contribution is relatively weak when it is non-negative. The denominator of (19) is also positive if country F evaluates country L's kindness in relative terms so that $b_2 < 0 < m_2 - m_1$.

(ii) $b_1 < m_2$ and $b_2 > m_2 - m_1$. In this case, the signalling effect caused by the leader is weak, while country F's belief in the motivational effect caused by its increased public good contribution is strong.

A negative reaction by country F obviously results if both behavioural effects, as described by b_1 and b_2, are weak. But, far more surprisingly, crowding-out is also observed if both behavioural effects, the signalling effect and the belief in the motivational effect are strong.

Note that the approach with endogenous belief formation outlined here is not subject to the usual objection against the use of the conjectural variation concept, which is based on the assumption of strict rational behaviour (Sugden 1985; Finus 2001, pp. 151–152; Figuières et al. 2004). Rather, in our model, we deliberately take determinants of behaviour into account that go beyond the standard rationality assumption.[14]

6 An Example

Inspired by the representation of conjectural variations in Cornes and Sandler (1984b, 1996), we provide a specific example where the perceived kindness of country L's public good contribution is also affected by country F's contribution, i.e. $b_2 < 0$ holds. In particular, the expectations of country F positively depend on country L's share of total contributions. Country F's belief function is

$$b\left(\frac{g_L}{G}\right) = b\left(\frac{g_L}{g_L + g_F}\right) \tag{20}$$

with $b' > 0$. From (19), the slope of country F's reaction function is easily calculated as

$$\frac{\partial g_F^r}{\partial g_L} = \frac{b'\frac{g_F}{G^2} - m_2}{b'\frac{g_L}{G^2} + m_2 - m_1}, \tag{21}$$

[14] Another objection (e.g., Sugden 1985) against the conjectural variations approach also does not hold when the conjectures are endogenized in the way as suggested in this paper, since it is possible that equilibria with positive conjectures and even Pareto optimal solutions satisfy the consistency requirement (Buchholz 2015). An early contribution on necessary and sufficient conditions for consistent conjectural equilibria in a public goods economy is Scafuri (1988).

which is positive if and only if $b'\frac{g_F}{G^2} > m_2$. If, given a positive reaction by the follower, $g_F \leq g_L$ holds in the initial allocation, the denominator of (21) exceeds the numerator and $\frac{\partial g_F^r}{\partial g_L} < 1$ follows. Then the induced increase of the follower's public good contribution is smaller than the leader's additional contribution.

For a further interpretation of (22), we abbreviate notation by setting $\sigma = \frac{g_L}{g_F}$ and $\tilde{b}(\sigma) := b(\frac{\sigma}{\sigma+1}) = b(\frac{g_L}{g_L+g_F})$. Furthermore, let $\varepsilon_{1+\tilde{b},\sigma} = \frac{\tilde{b}'\sigma}{1+\tilde{b}}$ and $\varepsilon_{m,g_L} = \frac{m_2 g_L}{m}$ denote the elasticities of country F's beliefs and its marginal rate of substitution, respectively, w.r.t. an increase in country L's public good contribution. Then, observing $m = 1 + b = 1 + \tilde{b}$, it follows from (21) that

$$\frac{\partial g_F^r}{\partial g_L} > 0 \Leftrightarrow \varepsilon_{1+\tilde{b},\sigma} > \varepsilon_{m,g_L}. \tag{22}$$

Condition (22) says that crowding-in results if an increase of g_L improves the follower's beliefs about the leader's behaviour more than it reduces its willingness to pay for the public good.

For a special belief function as given by (20), we now consider the effect of a unilateral increase of country L's public good contribution starting from an interior Nash equilibrium N, where the public good contributions of both countries are g_L^N and g_F^N, respectively. Furthermore, we assume $b\left(\frac{g_L^N}{g_L^N+g_L^F}\right) = 0$ for the conjectural variation function. Since in the Nash equilibrium zero conjectures prevail, the Nash equilibrium is also the equilibrium outcome, given these specific conjectural variations. We now assume that condition (22) is satisfied at N so that a marginal increase of country L's public good contribution triggers a positive reaction by the follower. Then a marginal increase of country L's public good contribution clearly leads to an increase in aggregate public good supply, which is larger than country L's additional contribution so that country L moves into a position above its indifference curve passing through (x_L^N, G^N). The same holds true for country F, which establishes that a Pareto improvement is attained through the unilateral action of country L.

We illustrate this result by assuming that both countries have the same utility function, $v(x_i, G) = x_i G$, and the same income level, $w_i = 1$ $(i = L, F)$. Then $m(x_i, G) = \frac{G}{x_i}$, $m_1(x_i, G) = -\frac{G}{x_i^2}$ and $m_2(x_i, G) = \frac{1}{x_i}$. Further, similar to Cornes and Sandler (1984b, 1996), the conjectural variation function can be given by

$$b\left(\frac{g_L}{G}\right) = \left(\frac{2g_L}{G}\right)^{\theta} - 1. \tag{23}$$

If $\theta > 0$, the follower has positive conjectures, and if $\theta < 0$, he has negative conjectures about the leader's behaviour, independent of the size of the leader's public good contribution. Given (23), condition (18) becomes

$$\frac{g_L + g_F}{1 - g_F} = \left(\frac{2g_L}{g_L + g_F}\right)^{\theta}. \tag{24}$$

According to (21), the slope of the reaction function, evaluated at the standard Nash equilibrium with $g_L^N = g_F^N = \frac{1}{3}$ and $x_L^N = x_F^N = G^N = \frac{2}{3}$, is

$$\frac{\partial g_F^r}{\partial g_L} = \frac{\theta - 2}{\theta + 4}, \tag{25}$$

since $b'(1) = \theta$, $m_1 = -\frac{3}{2}$ and $m_2 = \frac{3}{2}$. This implies strategic complements and thus crowding-in results if $\theta > 2$. Moreover, the larger θ, the stronger are the follower's positive reactions to the leader's increased public good contributions.

7 Implications for Climate Policy

This paper has shown that, by enriching the standard public good model with elements of behavioural economics, the otherwise occurring crowding-out effect is avoided so that unilateral increases of abatement efforts by some leading countries no longer are ineffective and meaningless. In this context, our analysis has emphasized the crucial role of other-regarding preferences, the formation of reciprocal beliefs and the building of trust, which are major features in behavioural economics. Both extensions of the public good model seem to be of great relevance for climate policy.[15]

On the one hand, ethical arguments lie at the heart of the whole debate on climate change and provide the motivation for more ambitious abatement measures. In particular, individuals, firms and countries want to have some fair burden sharing w.r.t. greenhouse gas abatement efforts, i.e., w.r.t. contributions to the global public good "climate protection". Moreover, these agents benefit from having a good conscience by acting responsibly towards the natural environment and future generations. These non-standard preference components, which are reflected in the type of utility function used in Sects. 3 and 4, may provide the basis for successful leadership and an improved global climate policy if those psychological preferences are sufficiently strong. But, interestingly, this positive outcome will not come about for all types of other-regarding preferences: If the follower is altruistic taking an interest in the leader's material well-being, leadership can never be successful, which might contradict intuition. At a general level, our considerations on the effects of psychological preferences also emphasize the eminent role "moral suasion" and thus the deliberate attempt at changing preferences could play for effective climate protection.

On the other hand, leading activities in climate policy shown in the past by the EU and especially by Germany, are accompanied by attempts to prove oneself as a reliable and trustworthy partner, being prepared to continue one's efforts for climate protection also in the future. This endeavour can explain the strong attempts of the German government to fulfil its obligations (which imply a reduction of greenhouse gas emissions by 40 % from 1990 till 2020)—even though the effect on global temperature is negligible. From this perspective, the European Emission Trading System (EU ETS) can also be considered as a commitment device to signal credibility of the EU's firm involvement in climate change issues (Buchholz and Peters 2005). It is the intention of this facet of leadership behaviour to foster positive expectations of potential followers and to create a general atmosphere of trust in the global community, which in general seems to be an essential pre-condition for effective cooperation among independent agents. How such belief formation may make leadership successful in a public good model has been demonstrated in Sects. 5 and 6.

Beyond these insights that are closely related to behavioural economics, there exist a lot of other factors by which leadership in global climate policy will be made more successful—and which are also well-known, not only from economic theory, but also from diplomacy in general and experiences in climate diplomacy in particular (e.g., Benedick 1999). So,

[15] The analysis in this paper has been purposefully restricted to the determination of conditions under which the follower will respond positively to increased public good contributions by the leader. To examine the equilibria of games of voluntary public good provision when agents have different types of psychological preferences would be an important topic in the same context which, however, is left to future research.

clearly, the danger of crowding-out is reduced when the leadership is provided by a large and powerful coalition, which could use issue-linkage (e.g., in the context of trade policy) to induce greenhouse gas mitigation by outsiders or even make credible threats to exclude them from international cooperation (e.g., Nordhaus 2015). The leading country can also use other (and more clever) instruments for its pioneering activities rather than raising direct public good contributions: It can apply a matching mechanism (e.g., Guttman 1978; Falkinger et al. 1996) to subsidize the public good contributions of other countries. It thus may avoid crowding-out behaviour by the followers and even attain Pareto optimal public good supply through unilateral action (Buchholz et al. 2015a). Moreover, pioneering activities may be conducted through e.g. "green" technological innovations and their transfer to other countries (e.g., Golombek and Hoel 2005; El-Sayed and Rubio 2014; Buchholz et al. 2015b), which may even have more environmental impact at lower cost than unilateral increases of abatement efforts. Such alternative approaches to successful leadership in global climate policy, however, lie outside the theoretical framework used in this paper.

References

Acemoglu D, Jensen MK (2013) Aggregative comparative statics. Games Econ Behav 81:27–49
Andreoni J (1990) Impure altruism and donations to public goods: a theory of warm-glow giving. Econ J 100:464–477
Arce M DG (2001) Leadership and the aggregation of international collective action. Oxf Econ Pap 16:17–25
Arce M DG, Sandler T (2005) The dilemma of the prisoners' dilemmas. Kyklos 58:3–24
Auerswald H, Konrad KA, Thum, M (2011) Adaption, mitigation and risk-taking in climate policy. CESifo working paper series 3320, CESifo Group Munich
Axelrod R (1984) The evolution of cooperation. Basic Books, New York
Barbieri S (2012) Communication and early contribution. J Public Econ Theory 14:391–421
Barrett S (2005) The theory of international environmental agreements. In: Mäler K-G, Vincent JR (eds) Handbook of environmental economics, vol 3. North-Holland, Amsterdam, pp 1457–1516
Benjamin DJ (2015) Distributional preferences, reciprocity-like behavior, and efficiency in bilateral exchange. Am Econ J Microecon 7:70–98
Benedick RE (1999) Contrasting approaches: the ozone layer, climate change, and resolving the Kyoto dilemma. WZB Discussion Paper, FS2 99–440
Bergstrom T, Cornes RC (1981) Independence of allocative efficiency from distribution in the theory of public goods. Econometrica 51:1753–1765
Bergstrom T, Blume L, Varian H (1986) On the private provision of public goods. J Pub Econ 29:25–49
Bolton GE, Ockenfels A (2000) ERC: a theory of equity, reciprocity, and competition. Am Econ Rev 90:166–193
Brandt US (2004) Unilateral actions: the case of international environmental problems. Resour Energy Econ 26:373–391
Buchanan JM (1967) Cooperation and conflict in public-goods interaction. West Econ J 5:109–121
Buchholz W (1993) A further perspective on neutrality in a public goods economy with conjectural variations. Public Finance Q 21:115–118
Buchholz W (2015) Implementing Pareto optimal solutions as consistent conjectural equilibria in a static game of voluntary public goods provision. Mimeo, University of Regensburg
Buchholz W, Cornes RC, Peters W, Rübbelke DTG (2015a) Pareto improvement through unilateral matching of public good contributions: the role of commitment. Econ Lett 132:9–11
Buchholz, W, Dippl L, Eichenseer M (2015b) Technological transfers in global climate policy. CESifo working paper series 5548, CESifo Group Munich
Buchholz W, Haslbeck C, Sandler T (1998) When does partial cooperation pay? FinanzArchiv 55:1–20

Buchholz W, Konrad KA (1994) Global environmental problems and the strategic choice of technology. J Econ 60:299–321
Buchholz W, Peters W (2005) A Rawlsian approach to international cooperation. Kyklos 58:25–44
Buchholz W, Peters W, Ufert A (2014) Binary climate games with reciprocal preferences. Mimeo, European University Viadrina, Frankfurt (Oder)
Cabral LM (1995) Conjectural variations as a reduced form. Econ Lett 13:397–402
Charness G, Rabin M (2002) Understanding social preferences with simple tests. Q J Econ 117:817–869
Cornes RC (2015) Aggregative environmental games. Environ Resour Econ. doi:10.1007/s10640-015-9900-6
Cornes RC, Hartley R (2007) Aggregative public good games. J Public Econ Theory 9:201–219
Cornes RC, Sandler T (1984a) Easy riders, joint production, and public goods. Econ J 94:580–598
Cornes RC, Sandler T (1984b) The theory of public goods: non-Nash behavior. J Public Econ 23:367–379
Cornes RC, Sandler T (1985) On the consistency of conjectures with public goods. J Public Econ 27:125–129
Cornes RC, Sandler T (1994) The comparative static property of the impure public good model. J Public Econ 54:403–421
Cornes RC, Sandler T (1996) The theory of externalities, public goods and club goods, 2nd edn. Cambridge University Press, Cambridge
Croson R (2007) Theories of commitment, altruism and reciprocity: evidence from linear public good games. Econ Inq 45:199–216
Dasgupta S, Itaya J-I (1992) Comparative statics for the private provision of public goods in a conjectural variations model with heterogeneous agents. Public Finance 47:17–31
DeCanio SJ, Fremstad A (2013) Game theory and climate diplomacy. Ecol Econ 85:177–187
De Oliveira AR, Faria RJ, Arce M DG (2005) Leading by example and international collective action. J Public Econ Theory 7:51–63
Dufwenberg M, Kirchsteiger G (2004) A theory of sequential reciprocity. Games Econ Behav 47:268–298
Edenhofer O, Flachsland C, Jakob M, Lessmann K (2015) The atmosphere as a global commons—challenges for international cooperation and governance. In: Semmler W, Bernard L (eds) The Oxford handbook on the macroeconomics of global warming. Oxford University Press, Oxford
El-Sayed A, Rubio SJ (2014) Sharing R&D investments in cleaner technologies to to mitigate climate change. Resour Energy Econ 38:168–180
Endres A, Ohl C (2001) International environmental cooperation in the one shot prisoners' dilemma. Schmollers Jahrbuch 121:1–26
Falkinger J, Hackl F, Pruckner G (1996) A fair mechanism for the efficient reduction of global CO_2-emissions. FinanzArchiv 53:308–331
Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. Q J Econ 114:817–868
Fehr E, Schmidt KM (2009) The economics of fairness, reciprocity and altruism—experimental evidence and new theories. In: Kolm S-C, Ythier JM (eds) Handbook of giving, altruism and reciprocity, vol 1. North-Holland, Amsterdam, pp 615–691
Figuières C, Jean-Marie A, Quérou N, Tidball M (2004) Theory of conjectural variations. World Scientific, Singapore
Finus M (2001) Game theory and international environmental cooperation. Edward Elgar, Cheltenham
Fischbacher U, Gächter S (2010) Social preferences, beliefs, and dynamics of free riding in public goods experiments. Am Econ Rev 100:541–556
Gächter S, Renner E (2014) Leaders as role models for the voluntary provision of public goods. CESifo working paper series 5049, CESifo Group Munich
Golombek R, Hoel M (2005) Climate policy under technological spillovers. Environ Resour Econ 31:201–227
Guttman J (1978) Understanding collective action: matching behavior. Am Econ Rev Pap Proc 68:251–255
Hermalin BE (1998) Toward an economic theory of leadership: leading by example. Am Econ Rev 88:1188–1206
Hoel M (1991) Global environmental problems: the effects of unilateral actions taken by one country. J Environ Econ Manag 20:55–70
Ihori T (1996) International public goods and contribution productivity differentials. J Public Econ 61:139–154
Itaya J-I, Okamura M (2003) Conjectural variations and voluntary public good provision in a repeated game setting. J Public Econ Theory 5:51–66
Kempf H, Rota Graziosi G (2010) Leadership in public provision: a timing game perspective. J Public Econ Theory 12:763–787
Kolstad CD (2012) Public good agreements with other-regarding preferences. NBER working paper 17017, Boston, MA
Konrad KA, Thum M (2014) Climate policy negotiations with incomplete information. Economica 81:244–256
Kotchen M (2007) Equilibrium existence and uniqueness in impure good models. Econ Lett 97:91–96

Lipnowski I, Maital S (1983) Voluntary provision of a pure public good as a game of chicken. J Public Econ 20:381–386
Nordhaus W (2015) Climate clubs: overcoming free-riding in international climate policy. Am Econ Rev 105:1339–1370
Nosenzo D, Sefton M (2011) Endogenous move structure and voluntary provision of public goods: theory and experiment. J Public Econ Theory 13:721–754
Nyborg K (2014) Reciprocal climate negotiators: balancing anger against even more anger. Mimeo, University of Oslo
Pittel K, Rübbelke DTG (2008) Climate policy and ancillary benefits—a survey and integration into the modeling of international negotiations on climate change. Ecol Econ 68:210–220
Pittel K, Rübbelke DTG (2012) Transitions in the negotiations on climate change: from prisoner's dilemma to chicken and beyond. Int Environ Agreem 12:23–39
Pittel K, van den Ploeg R, Withagen C (2014) Climate policy and nonrenewable resources—the green paradox and beyond. The MIT Press, Cambridge
Possajennikov A (2015) Conjectural variations in aggregative games: an evolutionary perspective. Math Soc Sci 77:55–61
Rabin M (1993) Incorporating fairness into game theory and economics. Am Econ Rev 83:1281–1302
Roemer JE (2010) Kantian equilibrium. Scand J Econ 112:1–24
Roemer JE (2015) Kantian optimization: a micro foundation for cooperation. J Public Econ 127:45–57
Sandler T (1996) A game-theoretic analysis of carbon emissions. In: Congleton RD (ed) The political economy of environmental protection. University of Michigan Press, Ann Arbor, pp 251–272
Sandler T (1997) Global challenges: an approach to environmental, political, and economic problems. Cambridge University Press, Cambridge
Sandler T, Posnett J (1991) The private provision of public goods: a perspective on neutrality. Public Finance Q 19:22–42
Scafuri AJ (1988) On consistency of conjectures in private provision of public goods. J Public Econ 37:395–398
Scafuri AJ (1992) Rational conjectures equilibria in the private provision of public goods. Public Finance Q 20:139–151
Sinn H-W (2012) The green paradox—a supply-side approach to global warming. MIT Press, Cambridge
Sobel J (2005) Interdependent preferences and reciprocity. J Econ Lit 43:392–436
Sugden R (1985) Consistent conjectures and voluntary contributions to public goods: why the conventional theory doesn't work. J Public Econ 27:117–124
Varian H (1994) Sequential contributions to public goods. J Public Econ 53:165–186