# Individual and Governmental Responsibility for Sustainability

## An Economic Analysis

Von der Fakultät Wirtschaftswissenschaften

der Leuphana Universität Lüneburg

zur Erlangung des Grades

Doktor der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)

genehmigte

Dissertation

von

Joachim Fünfgelt

aus

Heidelberg

Eingereicht am: 23. April 2012

Mündliche Prüfung am: 08. Juni 2012

Erstgutachter: Prof. Dr. Stefan Baumgärtner

Zweitgutachter: Prof. Dr. Thomas Wein

Prüfungsausschuss: Prof. Dr. Stefan Baumgärtner (Vorsitzender)

Prof. Dr. Thomas Wein

Prof. Dr. Markus Beckmann

Die einzelnen Beiträge des kumulativem Dissertationsvorhabens sind oder werden wie folgt in Zeitschriften veröffentlicht:

Fünfgelt, J. and S. Baumgärtner (2012). A utilitarian notion of responsibility for sustainability. University of Lüneburg Working Paper Series in Economics, Working Paper No. 234, March 2012.

Fünfgelt, J. (2012). Verantwortung für Nachhaltigkeit von Konsumenten. Submission planned for summer 2012.

Fünfgelt, J. and S. Baumgärtner (2012). Regulation of morally responsible agents with motivation crowding. University of Lüneburg Working Paper Series in Economics, Working Paper No. 241, May 2012.

Fünfgelt, J. and G.S. Schulze (2011). Endogenous environmental policy when pollution is transboundary. University of Lüneburg Working Paper Series in Economics, Working Paper No. 196, February 2011.

Elektronische Veröffentlichung des gesamten kumulativen Dissertationsvorhabens inklusive einer Zusammenfassung unter dem Titel:

Individual and Governmental Responsibility for Sustainability – An Economic Analysis

Veröffentlichungsjahr: 2012

Veröffentlicht 2012 im Onlineangebot der Universitätsbibliothek unter der URL: http://www.leuphana.de/bibliothek.html

# Contents

# Acknowledgements

I would like thank my first advisor Stefan Baumgärtner for his intellectual support, for his patience, and for his trust in this research project. I also thank my second advisor Thomas Wein for helpful comments, and for reviewing this thesis. Many thanks go to the co-author of my first paper Günther Schulze for teaching me to enjoy economic thinking.

Within the Sustainability Economics Group in Lüneburg, I would like to thank my colleagues Sebastian Strunz, Roland Olbrich, Nikolai Hoberg, Markus Groth, Klara Stumpf, Stefanie Glotzbach, John-Oliver Engler, Lutz Göhring, and Anna Magdalena Biehler. I enjoyed our scientific discussions, mutual encouragements, coffee breaks, and games of Siedler very much. Thanks for the good times and for becoming friends.

I am most grateful my parents Inge and Werner Fünfgelt who supported and believed in me throughout the years. An emotional home was also provided to me by many friends who encouraged me along the way. Thank you very much.

I would like to thank Katrin Riedel for her love, support and understanding which gave me the strength to go through with this thesis.

# Introduction

## 1. Motivation

In light of widespread sustainability problems such as climate change, biodiversity loss and deforestation it seems necessary that individuals and governments assume responsibility to "pass on a world of undiminished life opportunities to members of future generations" (Howarth 2007: 656). This necessity for responsibility for sustainability is well recognized and extensively debated. For example, the United Nations Environment Programme (UNEP) ranked for 2012 the issues of "Aligning Governance to the Challenges of Global Sustainability" and "Transforming Human Capabilities for the 21st Century" as the two most important emerging environmental issues (UNEP 2012). UNEP argues that the transition towards sustainability requires "major, new efforts by governments" as well as teaching individuals "to respond to global environmental change".

Responsibility is an action guiding concept which relates the abstract norm of sustainability with concrete action contexts. It thereby specifies what individuals and governments ought to do. However, individual and governmental means, for example fiscal resources or knowledge of the system are scarce. They may not be able to do everything that they ought to, and their responsibility is thus limited (Baumgärtner et al. 2006). Limitation of responsibility as a consequence of scarcities allows to consider responsibility for sustainability as an economic issue in the sense that economics "studies human behaviour as a relationship between ends and scarce means which have alternative uses" (Robbins 1932: 15). Considering responsibility as an economic issue holds big promises, as economic concepts and methods may shed light on the assignment and limitation of responsibility among political, economic and citizen actors (Baumgärtner and Quaas 2008: 8). Nonetheless, economic research on responsibility is still in its infancy. Fundamental questions on how the concept of responsibility and economics relate to and may mutually enrich each other are still unanswered.

In this thesis, I introduce the concept of responsibility to economic theory, focussing specifically on individual and governmental *responsibility for sustainability*. Traditionally, it is individual human beings who are considered to be the appropriate agents to bear responsibility (Gosselin 2006: 38). But there are reasons for locating responsibility for sustainability in institutions such as governments. One reason is that assigning responsibility to governments may be more effective since they have more resources and different means. Another reason is that assigning responsibility to governments may have a sheltering psychological effect on individuals (Shue 1988: 697). Since neither individuals nor governments will be able or willing to perfectly fulfill their responsibility, complete reliance on individual or governments responsibility alone is not feasible (Shue 1996: 61). From there, important economic questions arise: how should responsibility be distributed among agents?

How can agents, who are responsible for several normative aims, solve trade-offs? Do governmental policies affect individuals ability to assume responsibility? How can individuals efficiently induce governments to act responsibly?

I address these questions by analyzing normative and positive aspects of individual and governmental responsibility for sustainability in four research papers. My analysis comprises two levels of abstraction referring to the comprehensive multi-level approach by Baumgärtner et al. (2008), which distinguishes the abstracion levels of concepts, models and case studies. On the level of concepts, I develop a concept of responsibility for sustainability which, on the level of models, I use to derive general insights from stylized models. My analysis does not comprise the level of case-studies, but it provides concrete empirically testable results and orientation for further empirical research.

This introduction is organized as follows. Section 2 describes the aims and methodological approaches of the research papers. Section 3 discusses their results, shortcomings, and general contributions.

# 2. Research Papers

This thesis includes four research papers (Figure 1). I begin with normative analyses where I formalize individual responsibility in the first paper (**Chapter 1**) and investigate the individual and political responsibility of consumers in the second paper (**Chapter 2**). I then proceed with positive analyses, focusing on behavioral aspects of responsibility. In the third paper (**Chapter 3**), I analyze how governmental policies affect individuals' motivation to act responsibly. In the fourth paper (**Chapter 4**), I analyze how individuals' behavior impacts on governments' responsibility.

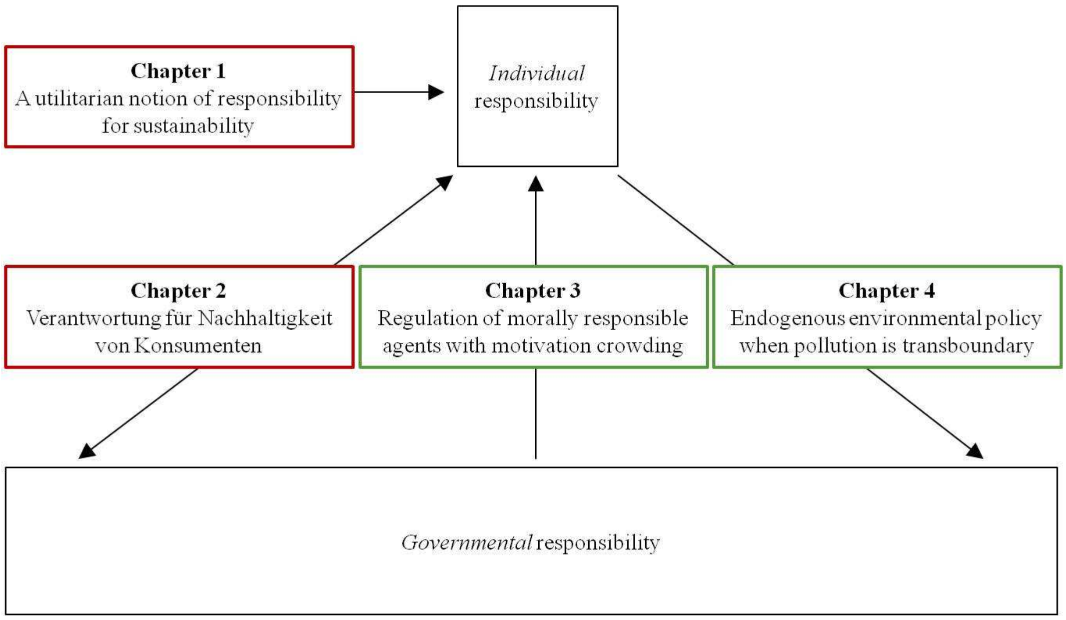



Figure 1: Papers of the thesis. Red boxes indicate normative analyses and green boxes indicate positive analyses.

In Chapter 1, **A utilitarian notion of responsibility for sustainability**, I conceptualize and formalize a utilitarian notion of responsibility for sustainability which I then relate to established normative criteria for assessing intertemporal societal choice.

The economic research on the normative foundation for responsibility focuses predominantly on backward-looking responsibility, that is “the idea that individuals are or should be held responsible, to some degree, for their achievements” (Fleurbaey 2008: 1). In contrast, the concept of forward-looking responsibility in the sense of an obligation (as in Baumgärtner et al. 2006) is only scarcely explored. Yet, this concept may fill the gap between the abstract norm of sustainability and the specific action contexts. Given that these contexts are typically characterized by limited feasability sets and knowledge, I explore in this paper whether responsibility considered in an economic framework can provide guidance for actions targeting sustainbility.

To this end, I develop a conceptualization of responsibility for sustainability on the basis of Singer’s (1972) principle and the Brundtland Commission’s notion of sustainability (WCED 1987). Subsequently, to illustrate the developed conceptualization, I formulate a simple model which comprises two non-overlapping generations sharing a natural resource which is used for the production of a good that allows for the satisfaction of basic needs and wants. I relate responsibility for sustainability to three common normative criteria for assessing intertemporal societal choice: Pareto-efficiency, (discounted) utilitarian welfare maximization, Brundtland-sustainability.

In Chapter 2, **Verantwortung von Konsumenten für Nachhaltigkeit**[1], I study consumers’ responsibility for sustainability. Particularly, I specify crucial components of this responsibility in order to analyze the relation of consumers’ private and political responsibility.

Consumption is a vital part of every day life. Every human being is a consumer, even if it is solely for satisfaction of basic needs such as nutrition, clothing or accommodation (König 2008). Not surprisingly, the question about the responsibility of consumers is being discussed both in the media (e.g. von Petersdorff 2011 or Langrock-Kögel 2012), and in the scientific discourse. In 2010, Grunwald started a discussion on the issue in the transdisciplinary journal GAIA, and argued that consumers ought to become active in the political sphere rather than changing their consumption patterns. He holds that sustainable consumption was neither effective nor normatively mandatory to achieve the aim of sustainability (Grunwald 2010) and political action is thus necessary. Petersen and Schiller (2011) share his view, arguing that consumption is and ought to remain private while sustainability is a political aim. Consumer responsibility is thus bound to influencing political decisions. Billharz et al. (2011) and Siebenhüner (2011) oppose this view arguing that consumers ought to change their consumption patterns as such a change would reduce the sustainability problem as well as create pressure in the political sphere. Altogether, all authors agree that consumers have a political responsibility for sustainability but disagree on the question of responsible

[1] English Title: Consumers’ responsibility for sustainability.

consumption and whether consumption can be politically relevant. The reason for the disagreement is that the concept of sustainability remains unclear in the discussion, so that the correlative responsibility remains unclear as well.

I therefore embark to clarify the above disagreement by specifying the concept of sustainability, specifically as a notion of intra- and intergenerational justice. As such, sustainability creates legitimate claims of human beings living today and in the future (Ott and Döring 2008). These sustainability claims entail a correlative responsibility to ensure that they are fulfilled. Ensuring the fulfillment of sustainability claims requires that responsibility for sustainability comprises at least three obligations (Shue 1996): First, the obligation to avoid depriving others from their claims. Second, the obligation to protect others from such deprivation. Third, the obligation to aid the deprived. Having specified the concept in such a way, I reconsider the aforementioned discussion and analyze which of these obligations can be ascribed to the responsibility of consumers and their consumption choices.

In Chaper 3, **Regulation of morally responsible agents with motivation crowding**, I focus on the impact of governmental policies on the motivation of an individual to assume moral responsibility. In particular, I study the regulation of a morally responsible individual with motivation crowding in the context of a negative externality.

Whenever negative externalities cause market failure, conventional economic wisdom (e.g. Pigou 1932, Baumol 1972) suggests the introduction of governmental policies (especially taxes or subsidies) to obtain an efficient outcome. Two critical assumptions are that preferences are purely self-centered and independent from governmental policies. However, both assumptions are not generally justified. First, human beings often assume moral responsibility, i.e. they respond to moral obligations (e.g. Sen 1977, Brekke et al. 2003). Second, evidence on Motivation Crowding Theory (Frey and Jegen 2001) has shown that governmental policies do affect preferences as they alter individuals' motivation to assume moral responsibility. It is thus questionable whether standard economic instruments lead to efficient outcomes. While there is a large body of empirical evidence on moral responsibility and motivation crowding (e.g. Gneezy et al. 2011), only few theoretical treaties on the issue exist. None of these theoretical studies has simultaneously considered negative externalities, moral responsibility and motivation crowding. Given these phenomena, the focus of this paper is to analyze whether a Laissez-faire, a tax policy, provision of information, and a complementary policy may lead to efficiency.

For this purpose, I develop an economic model based on the dual preferences model of Brekke et al. (2003) and the motivation crowding model of Frey and Oberholzer-Gee (1997). In my model, the consumption decision of individual *A* creates a negative externality which directly harms individual *B*. Individual *A* is motivated to assume moral responsibility as she weighs her self-directed utility from consumption with the knowingly inflicted externality on individual *B*. The model can accommodate motivation crowding since policy measures affect the relative weight individual *A* assigns the externality: a tax decreases the weight of the externality, reflecting a decrease in the motivation to assume responsibility (crowding-out); in contrast, provision of pefect information about the externality increases the weight, reflecting an increase in the motivation to assume responsibility increases (crowding-in). I use this model to evaluate four scenarios: a laissez-faire scenario, a tax-only scenario, an information-only scenario, and a complementary scenario where both tax is introduced and information is provided.

In Chapter 4, **Endogenous Environmental Policy when Pollution is Transboundary,** I examine how individuals which form lobby groups affect the determination of environmental policy when governments seek not only to maximize welfare – as is their responsibility to their voters – but simultaneous maximize support by lobby groups. More specifically, I consider the case in which two countries are linked through transboundary pollution.

There is ample evidence that environmental policy formation is indeed influenced by lobby groups since governments seek their support (Fredriksson et al. 2005). This leads to a distortion of the governments' responsibility to maximize welfare with the resulting equilibrium policies differing considerably from the Pigouvian rule (Aidt 1998). For national pollution, the distortion from lobbying has been broadly discussed in the literature (e.g. Fredriksson 1997, Schleich 1999, or Aidt 1998). For transnational pollution, an additional distortion may arise: When national environmental policies remain non-cooperative, even responsible governments internalize externalities only to the extent that they affect their own country (Markussen 1975). Is is largely unclear how the distortion arising from lobbying and the distortion arising from transnational pollution interact. This paper aims to fill this gap by focusing only on the interaction of a political distortion and transnational pollution.

To this end, I employ the common agency model developed by Bernheim and Whinston (1986) and introduced by Grossman and Helpman (1994) to the economic literature. I assume that there are two small open economies producing a pollution intensive good with perfectly transboundary pollution. The assumption of small open economies is crucial to avoid economic leakage effects (which have been discussed in Conconi 2003) and to focus on strategic interaction due to transboundary pollution. Furthermore, I assume that both national governments maximize a political support function which is composed of national welfare and contributions by lobby groups. An environmental lobby group advocating stricter environmental policies is formed by individuals who are afflicted with pollution. In contrast, an industrial lobby group advocating laxer environmental policies is formed by individuals who derive income from the production of the polluting good. The relative strength of the lobby groups, the preferences for environmental quality and the importance of national welfare for the governments' support may differ between the countries.

# 3. Discussion and conclusion

In Chapter 1, **A utilitarian notion of responsibility for sustainability**, I verify that responsibility for sustainability can be clearly and unambiguously conceptualized in economic models. Regarding the relation of responsibility for sustainability to established normative criteria, I find that sustainability and responsibility for sustainability are equivalent if and only if sustainability is feasible. If it is not, there still exists a responsible allocation which is

Pareto-efficient. This finding affirms that responsibility may provide action guidance even if the aim of sustainability is not feasible. Furthermore, utilitarian welfare maximization without discounting always fulfills the criterion of responsibility. Moderate discounting may be responsible if it leads to a sustainable allocation. If, however, sustainability is not feasible, discounting is not responsible.

One limitation of the analysis is that I focus on the satisfaction of basic needs and thus on a specific concept of sustainability and on a specific ethics. The satisfaction of basic needs, however, can only be regarded as a minimum requirement for sustainability. For other aspects of sustainability, the approach is less well suited, and hence other conceptualizations of responsibility are needed. A second limitation concerns the simplicity of the model. Introducing a second good into the model might allow the analysis to cover further issues, to gain further generality and to converge to standard economic models. For the purpose of the paper to provide general insights on responsibility, however, using a simple model is an adequate strategy (Baumgärtner et al. 2008: 9).

Generally, this paper demonstrates that the conception of responsibility adds specificity to the economic discussion about sustainability in two respects: (1) it clearly specifies how to act if sustainability is not feasible; (2) it specifies how to balance the trade-off between legitimate claims of present and future generations. In addition, the approach of this paper is useful for further economic research, especially regarding responsibility for sustainability under uncertainty which arises, for example, from technological progress.

In Chapter 2, **Verantwortung von Konsumenten für Nachhaltigkeit**, I argue that the responsibility for sustainability of consumers comprises three obligations. The first obligation concerns consumers' consumption choices where consumers ought not to deprive others from their legitimate sustainability claims. In contrast to the argument of Grundwald (2010), this obligation is relevant. For example, Meyer et al. (2010) estimate that moderate consumption changes in the fields of alimentation, construction and habitation could save up to 8.7% $CO_2$ emissions in Germany by 2020. The second obligation concerns the protection of others from the deprivation of their sustainability claims, particularly by the design of the economic system in such a way that strong incentives to deprive others are avoided. This obligation is especially relevant if consumers have trouble fulfilling the first obligation (avoid depriving others from their sustainability claims). This second obligation can be fulfilled by consumers directly by getting involved in the political process or indirectly by their consumption choices which influence political decision-makers. Yet, consumers are not obliged to fulfill this obligation through their consumption choices. The third obligation to aid the deprived is borne by consumers if failures *within* the economic system caused the deprivation. If deprivations were cause by external influences, for example by natural disasters, consumers do not bear the obligation to aid. Again, performing this obligation through consumption choices is not mandatory and only desirable if it is more efficient than other means such as giving directly to non-governmental organizations.

Due to its normative nature, my analysis cannot answer the question whether society should direct more attention to responsible consumption choices or to political efforts. I show, however, that both measures are relevant. Which measures currently deserves more attention, remains a matter of implementation. My conclusion is that both, individual consumption and political involvement, can and ought to be better utilituzed for sustainability and therefore no single measure should be trivialized.

In Chapter 3, **Regulation of morally responsible agents with motivation crowding**, I first show that moral responsibility is unlikely to lead to Pareto efficiency but may increase or decrease the market failure. Hence, the necessity for governmental policies remains although individuals assume moral responsibility. Second, I find in a tax-only scenario that there exists at least one efficient tax rate even if there is strong motivation crowding. The intuition is that once all moral motivation is crowded-out, a tax applies just as in the standard economic model. However, setting a tax at an ineffiently low level may exacerbate the market failure due to motivation crowding. Third, I show in an information-only scenario that provision of perfect information may not lead to Pareto-efficiency but may still diminish the extent of the market failure. Fourth, a complementary policy (joint use of a tax and provision of information) may require lower taxes, may reduce market failure due to inefficiently low taxes, and may lead to efficiency without fully crowding-out moral motivation. This result is highly contingent on parameter values so that a complementary policy is recommendable for some but not for all cases.

One limitation emerges from a lack of empirical evidence on crowding effects. While many studies demonstrate the existence of motivation crowding, the precise response of motivation crowding to policy measures is mostly unknown: Do crowding effects increase or decrease with a tax rate or the size of an information campaign? Does an information campaign affect the crowding effects of a tax? Since such questions are still open, the model had to remain very general and some outcomes, like a strong reversal of market failure, had to be ruled out by assumption.

A second limitation concerns the weaknesses of a tax-only policy in the paper. These weaknesses are not derived by the model, because this would require an unreasonably complex general equilibrium model. Instead, they are the result of general considerations. This reduces the appeal of the result and prevents a more detailed analysis of, for example, motivational spill-overs. Despite this limitation, I believe that the weaknesses of a tax-only policy stand on firm considerations and should thus be taken seriously.

The third limitation is that the paper does not provide new policy instruments to the problem of motivation crowding. It merely discusses the effects of well known instruments and concludes that they should be used with caution since motivation crowding impedes their proper application. Notwithstanding, the paper makes an important contribution by providing a general framework for further theoretical research as well as guidance for further empirical research.

Altogether, motivation crowding challenges the efficient application of tax policies, but not the efficiency of tax policies itself. To fully understand the regulation of morally responsible individuals, it will be necessary to shed further light on in how far higher taxes or levels of information cause stronger crowding than lower taxes or levels of information, and in how far complementary instruments affect the crowding effects of taxes.

In Chapter 4, **Endogenous Environmental Policy when Pollution is Transboundary**, I present three major findings. First, environmental policies adopted by self-interested (political support seeking) governments may be more stringent than by responsible (social welfare maximizing) governments. More precisely, the distortion arising from transboundary pollution may be alleviated by the distortion arising from lobbying if the influence of the environmental lobby group is higher than that of the industrial lobby group. Second, due to the interaction of distortions the space of optimal policies increases: politically optimal tax rates may be too high to optimally internalize the environmental externality (for strong environmental lobby groups) but they may also be too low (for strong industrial lobby groups). They may even be negative such that governments subsidize the production of the polluting good. Third, the interaction of the distortions may create instability of the equilibrium if the relative influence of either lobby group is too large. In the resulting corner solutions, one country cedes production of the polluting good to the other.

There are three limitations concerning this paper. First, the paper is not directly related to any specific specific real-world situation that it seeks to illuminate. One might argue that this limits its interest. I disagree with this view. There are, in fact, real-world situations adequately related to the assumptions of the paper, such as Scandinavian $SO_x$ depositions or the phenomenon of Asian Brown Clouds. In addition, this paper serves general interest as it has the purposes of theory-development and of understanding the functioning of the two discussed distortions.

Second, some assumptions of the model are very restrictive such as non-depletable externalities, only two countires with only one firm in each, exogenous lobby group membership and no economic leakages. This is due to the common agency model that I employed. While being a powerful instrument to anaylze endogenous policy making, it allows only for a narrow scope of analysis. Naturally, relaxing these assumptions could substantially enrich the scope of the analysis. Still, the analysis provides valuable insights regarding the relation of the two interacting distortions. These insights are applicable to a wide range of phenomena, including pollution stretching over more than two countries or depletable externalities.

Third, the result that an increase in the share of the industrial lobby group may lead to an increase in national social welfare is driven by the revenue effects operating through exogenous lump sum reimbursement. Ethier (2007) has correctly pointed out that such revenue effects do, in reality, not play a big role for shaping the policy preferences of lobby groups. The paper therefore overrates revenue effects due to the model characteristics.

Notwithstanding, if environmental concerns are systematically underrated (for example due to the occurence of transnational pollution and non-cooperating governments), a moderate and asymmetric relocation of political influence is welfare improving and reduces overall pollution. Restrictions on lobbying by polluting industries and support for lobbying by environmental groups could be brought about by a number of concrete measures. For example, donations to parties could be restricted as this would affect industry associations and corporations scope for lobbying disproportionately. On an individual level, the results of the paper provide a strong argument to join or support environmental lobby groups.

## References


Aidt, T. (1998). Political internalization of economic externalities and environmental policy. *Journal of Public Economics* 69: 1–16.

Baumgärtner, S., Faber, M. and Schiller, J. (2006). *Joint Production and Responsibility in Ecological Economics. On the Foundations of Environmental Policy*. Edward Elgar, Cheltenham.

Baumgärtner, S., Becker, C., Frank, K., Müller, B., and Quaas, M. (2008). Relating the philosophy and practice of ecological economics. The role of concepts, models and case studies in inter- and transdisciplinary sustainability research. *Ecological Economics* 67(3): 384–393.

Baumgärtner, S., and Quaas, M. (2010). What is sustainability economics? Ecological Economics 69(3): 445–450.

Baumol, W. (1972). On taxation and the control of externalities. *American Economic Review* 62(3): 307–322.

Bernheim, B. and Whinston, M. (1986). Menu auctions, resource allocation, and economic influence. *Quarterly Journal of Economics* 101(1): 1–31.

Bilharz, M., Fricke, V., and Schrader, U. (2011). Wider die Bagatellisierung der Konsumentenverantwortung. *GAIA* 20(1): 9–13.

Brekke, K., Kverndokk, S. and Nyborg, K. (2003). An economic model of moral motivation. *Journal of Public Economics* 87: 1967–1983.

Conconi, P. (2003). Green lobbies and transboundary pollution in large open economies. *Journal of International Economics* 59: 399–422.

Ethier, W. (2007). The theory of trade policy and trade agreements: a critique. *European Journal of Political Economy* 23: 605–623.

Fleurbaey, M. (2008). *Fairness, Responsibility, and Welfare*. Oxford University Press, Oxford.

Fredriksson, P. (1997). The political economy of pollution taxes in a small open economy. *Journal of Environmental Economics and Management* 33: 44–58.

Fredriksson, P., Neumayer, E., Damania, R., and Gates, S. (2005). Environmentalism, democracy, and pollution control. *Journal of Environmental Economics and Management* 49(2): 343–365.

Frey, B. and Oberholzer-Gee, F. (1997). The cost of price incentives: an empirical analysis of motivation crowding-out. *American Economic Review* 87(4): 746–755.

Frey, B. and Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys* 15(5): 589–611.

Gosselin, A. (2006). Global poverty and responsibility: identifying the duty-bearers of human rights. *Human Rights Review* 8(1): 35–52.

Grossman, G. and Helpman, E. (1994). Protection for sale. *American Economic Review* 84(4): 833–850.

Grunwald, A. (2010). Wider die Privatisierung der Nachhaltigkeit. *GAIA* 19(3): 178–182.

Howarth, R. (2007). Towards an operational sustainability criterion. *Ecological Economics* 63: 656–663.

König, W. (2008). *Kleine Geschichte der Konsumgesellschaft. Konsum als Lebensform der Moderne*. Stuttgart, Franz Steiner.

Langrock-Kögel, C. (2012). Wir Teilzeit-Helden. *Enorm Magazin* 1(2012): 16–26.

Markusen, J. (1975). Cooperative control of international pollution and common property resources. *Quarterly Journal of Economics* 89(4): 618–32.

Meyer, B., Meyer, M., and Meyer zu Holte, I. (2010). *Die ökologischen und ökonomischen Wirkungen eines nachhaltigeren Konsums in Deutschland*. Studienreihe Mut zur Nachhaltigkeit. Nr. 1. Retrieved from: http://www.mut-zur-nachhaltigkeit.de/images/dokumente/Infopool/studienreihe%20mzn%201_nachhaltiger%20konsumieren.pdf (April 3, 2012)

Nyborg, K. and Rege, M. (2003). Does public policy crowd out private contributions to public goods? *Public Choice* 115: 397–418.

Ott, K., and Döring, R. (2008). *Theorie und Praxis starker Nachhaltigkeit*. 2. Edition. Marburg, Metropolis Verlag.

Petersen, T. and Schiller, J. (2011). Politische Verantwortung für Nachhaltigkeit und Konsumentensouveränität. *GAIA* 20(3): 157–161.

Pigou, A. (1932)[1920]. *The Economics of Welfare*. 4th edition. London.

Robbins, L. (1932). *An Essay on the Nature and Significance of Economic Science*. London, Macmillan.

Schleich, J. (1999). Environmental quality with endogenous domestic and trade policies. *European Journal of Political Economy* 15(1) 53–71.

Shue, H. (1988). Mediating Duties. *Ethics* 98: 687–704.

Shue, H. (1996). *Basic rights: Subsistence, affluence and U.S. foreign policy*. 2. Edition. Princeton, NJ, Princeton University Press.

Siebenhüner, B. (2011). Kann die Politik es richten? Konsument(inn)en als politische Akteure. *GAIA* 20(1): 14–16.

Singer, P. (1972). Famine, affluence, and morality. *Philosophy and Public Affairs* 1(3): 229–243.

[UNEP] United Nations Environment Programme. (2012). *21 Issues for the 21$^{st}$ Century*: *Result of the UNEP Foresight Process on Emerging Environmental Issues*. Nairobi, Kenya.

von Petersdorff, W. (2011). Grüne Revolution – Die herzliche Ökodiktatur. *Frankfurter Allgemeine Zeitung*. 16.05.2011.

[WCED] World Commission on Environment and Development (1987). *Our Common Future*. Oxford University Press, Oxford.

# Paper 1: A utilitarian notion of responsibility for sustainability

JOACHIM FÜNFGELT[*] AND STEFAN BAUMGÄRTNER

Department of Sustainability Sciences and Department of Economics,
Leuphana University of Lüneburg, Germany

**Abstract:** We develop and formalize a utilitarian notion of responsibility for sustainability inspired by Singer's (1972) principle and the Brundtland Commission's notion of sustainability (WCED 1987). We relate this notion of responsibility to established criteria for the assessment of intertemporal choice: Pareto-efficiency, (discounted) utilitarian welfare maximization, and Brundtland-sustainability. Using a two-generations-resource-model, we find that sustainability and responsibility are equivalent if and only if sustainability is feasible. If it is not, there still exists a responsible allocation which is Pareto-efficient. The undiscounted utilitarian welfare maximum is responsible while discounting is responsible to a certain extent if and only if sustainability is feasible.

**JEL Classification:** D63, D90, Q01, G56

**Keywords:** basic needs, Brundtland, discounting, ethics, natural resources, Pareto efficiency, responsibility, Singer, sustainability, utilitarianism

[*] Corresponding author: Sustainability Economics Group, Leuphana University of Lüneburg, P.O. Box 2440, D-21314 Lüneburg, Germany, phone: +49.4131.677-2668, fax: +49.4131.677-1381, email: fuenfgelt@uni.leuphana.de

# 1 Introduction

Sustainability is a very broad conception of justice. As such it poses an imperative on presently living persons. This imperative of sustainability implies intra- as well as intergenerational justice. More specifically, as defined by the Brundlandt Commission (WCED 1987), sustainability refers to the satisfaction of basic needs of present and future generations. To realize sustainability, presently living persons ought to act sustainably which implies at least two obligations: one directed towards the present generation and the other towards future generations.

Acting sustainably means to take specific actions in accordance with the norm of sustainability in a concrete action context (Baumgärtner et al. 2010). An action context is characterized by a feasible set of actions, given system structure and dynamics, and knowledge of the system. This may create a gap between the general and abstract imperative to act sustainably, and the specific action context since the set of feasible actions and the knowledge of the system may be limited. This paper aims to fill this gap by conceptualizing a person's *responsibility* for sustainability.

The concept of responsibility – as it has emerged from modern practical philosophy, political science, and law – links abstract norms with specific action contexts (Baumgärtner et al. 2010). It is gaining importance in the normative assessment of public policy-making as well as of private decision-making, since the impacts of human actions have increased dramatically in modern times (Jonas 1979). Some impacts are irreversible and occur at remote places or far in the future, such as e.g. anthropogenic climate change or biodiversity loss. Furthermore, action contexts are often characterized by uncertainty and unidirectional power structures.

One crucial feature of responsibility is that it is limited – namely by the acting person's possibility of compliance as well as by the need to balance a plurality of normative obligations. Therefore, the imperative of sustainability cannot imply an absolute obligation to attain a particular (sustainable) state or development of the world. It does imply, though, a relative obligation to do one's best to live up to the challenge of sustainability. Responsibility thus provides a criterion for decision-making even if the aim of action (e.g. sustainability) is out of reach. Thereby, it adds to economic theory which traditionally assumes that the aim of action (e.g. utilitarian welfare maximization under some constraints) can always be reached. The crucial question of responsibility, then, is: what exactly does "one's best" mean?

In this paper, we develop and formalize a utilitarian notion of responsibility for sustainability which is inspired by Singer's (1972) principle and the Brundtland Commission's notion of sustainability (WCED 1987). To illustrate the meaning of the utilitarian notion of responsibility thus developed, we apply it in a simple model and relate it to established criteria for the normative assessment of intertemporal societal choice, namely Pareto-efficiency, (discounted) utilitarian welfare maximization, and Brundtland-sustainability. The model comprises two non-overlapping generations. They share a natural resource from which they produce a consumption good that allows them to satisfy their basic needs and wants. We thus model a simple resource allocation problem, yet with a unidirectional power structure: the first generation can decide which share of the resource to use for itself and which share to hand over to the second generation. This simple setup allows us to analyze and compare which allocations satisfy different normative criteria.

This study adds to the economic literature about responsibility. The normative strand of the literature focuses predominantly on retrospective responsibility, that is "the idea that individuals are or should be held responsible, to some degree, for their achievements" (Fleurbaey 2008: 1). We follow the idea of forward-looking responsibility in the sense of an obligation (as in Baumgärtner et al. 2006) and sharpen this idea as we formally implement it in economic modeling. Besides normative implications of responsibility, there further exists a descriptive strand in the literature, which analyizes the implications of individuals wanting to assume responsibility for the public good (e.g. Frey 1997, Brekke et al. 2003, Nyborg and Rege 2003, Heyes and Kapur 2011).

Our results show that sustainability and responsibility for sustainability are equivalent if and only if sustainability is feasible. If it is not, there still exists a responsible allocation which is also Pareto-efficient. Further, the utilitarian welfare maximum without discounting always fulfills the criterion of responsibility. Discounting may be rsponsible to a certain extent if sustainability is feasible. If sustainability is not feasible, discounting is not responsible. At a more general level, we demonstrate that responsibility can be formalized in economic models which adds specificity to the discussion about normative conceptions such as sustainability.

The paper is organized as follows. Section 2 defines and discusses the concepts of sustainability and responsibility, thus preparing the conceptual, normative basis for the analysis. Section 3 introduces the model. Section 4 gives formal definitions and characterizations, through necessary and sufficient conditions, of the normative criteria. Section 5 presents our results. Section 6 concludes.

# 2 Conceptual foundations

**Sustainability**

Sustainability, as we understand it, is a very broad conception of justice. It combines the ideas of global intragenerational justice and of intergenerational justice, and often also includes justice towards nature. We apply a specific anthropocentric notion of sustainability, namely the Brundtland Commission's definition: "Sustainable development is a development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (WCED 1987: 43). This definition includes elements of intra- and intergenerational justice but not towards nature. It is anthroprocentric and implies that only human beings deserve moral attention. By "generation" it thus refers to all human beings living at the same time period.[1] Furthermore, Brundtland-sustainability is in part result-

[1] From now on, we use the term "generation" in that sense.

oriented as it aims at the satisfaction of basic needs of the present generation, and in part prerequisite-oriented as it aims at maintaining the future ability to satisfy basic needs. In this paper, we focus on the aspect of intergenerational justice. However, our analysis can just as well be applied to intragenerational justice.

The term “basic needs” requires further specification.[2] In the Brundtland-definition it is further specified as “[…] the essential needs of the world's poor, to which overriding priority should be given” (WCED 1987: 43). Such basic needs include two elements. First, they include certain minimum requirements of a family for private consumption: adequate food, shelter and clothing are obviously included, as would be certain household equipment and furniture. Second, they include essential services provided by and for the community at large, such as safe drinking water, sanitation, public transport, and health and educational facilities (ILO 1976: 32). For our analysis, a crucial assumption is that at least some basic needs “can be set on the basis of scientific findings” (ILO 1976: 33) and that essential, “[…] fundamental human needs are finite, few and classifiable; and […] fundamental human needs […] are the same in all cultures and in all historical periods” (Max-Neef 1991: 18). These assumptions ensure the applicability of Singer’s principle in the context of intergenerational justice.

Being a conception of justice, Brundtland-sustainability defines legitimate claims of present and future generations with respect to the satisfaction of their basic needs. Thereby, it poses an imperative on presently living persons. Such persons ought to act in accordance with the norm of sustainability, that is, they ought to act such as to fulfill all these legitimate claims. Taking a specific action always occurs in a concrete action context in which there exists a set of feasible actions and in which knowledge about given system structures and dynamics are crucial to choose actions that deliver desired outcomes. There thus exists a gap between the abstract norm of sustainability and the specific action context, which needs to be closed in economic thinking and modelling.

**Responsibility**

This gap can be closed with the concept of responsibility. Responsibility is a multifarious notion. In the philosophical discussion of responsibility, at least three different aspects of the notion have been distinguished. (1) The primary meaning of responsibility is being the perpetrator of one’s own actions, that is, “[…] one ascribes an action to oneself and allows for it to be thus ascribed” (Baumgärtner et al. 2006: 227). The primary meaning is purely descriptive and has no moral relevance by itself. It simply states that $A$ is responsible for $X$ if and only if $A$ is the perpetrator of $X$.[3] This is a precondition of morality, as one can only be morally praised or blamed for an action that can be ascribed to oneself.

[2] We use the terms “basic” needs, “essential” needs and “fundamental” needs synonymously.

[3] The related and no less relevant question of which (future) consequences of one’s action can be ascribed to oneself poses a number of intricacies in a world where several actors interact and there are stochastic influences on system dynamics (Vallentyne 2008).

(2) Often, we use ‘responsibility’ as a synonym for obligation (Williams 2008: 458). This is what Baumgärtner et al. (2006) call the secondary meaning of responsibility. In this meaning, responsibility attains a moral significance when obligations exist which a person morally has to accept, that is, *A* ought to do *X* (positive responsibility) or ought not to do *X* (negative responsibility) for moral reasons. Such obligations arise for different reasons, one of which are the legitimate claims that some claim holders have due to principles of justice. In view of Brundtland-sustainability, there exists a positive responsibility (in the sense of: obligation to fulfill a legitimate claim), namely to satisfy the basic needs of the present generation, and a negative responsibility, namely to not compromise the ability of future generations to satisfy their basic needs.

(3) Williams (2008) defines a third meaning of responsibility: “Responsibility represents the readiness to respond to a plurality of normative demands” (Williams 2008: 459). In other words, responsibility is important whenever a person is facing a plurality of normative obligations. This becomes relevant for sustainability as the Brundtland-notion of sustainability contains two obligations: satisfying the basic needs of the present generation and not compromising the ability of future generations to satisfy their needs.

Our notion of responsibility for sustainability encompasses these three meanings. That is, our notion of responsibility is not purely descriptive (primary meaning) but is essentially normative, as it refers to an oblication that arises from some principles of justice (secondary and third meaning).

To further sharpen this notion of responsibility, we need to specify who bears responsibility for sustainability. In general, this could be every member of the present generation, e.g. an individual, a group of individuals, a corporation, a nation state and so on. The minimum requirement for being responsible is to be a person-like entity. Locke (1959: 264) defines a person as[4] “[a] thinking intelligent being that has reason and reflection and can consider itself as itself, the same thinking thing, in different times and places.” A person or a person-like entity[5] is thus defined by intelligence, capacity for reason, self-awareness and consciousness of time and space. With our focus on sustainability as intergenerational justice, we consider only presently living persons to bear responsibility for Brundtland-sustainability, while the whole present generation[6] and all future generations have legitimate claims due to Brundtland-sustainability.

[4] Locke’s idea of what constitutes a person is not undisputed (see Gertler 2010 for a discussion). Yet, it fits well with Singer’s understanding of a person.

[5] From here on, we use the term “person” in a broad sense including all “person-like entities” which satisfy Locke’s definition.

[6] Note that the two groups of (1) the presently living generation, who holds legitimate claims to the satisfaction of their basic needs according to Brundtland-sustainability, and (2) the presently living persons or person like entities, who bear responsibility for sustainability, do not need to be indentical. There may be members of the present generation who are not persons, and there may be persons who are not members of the present generation. As an example of the former, a presently living human infant has according to Brundtland-sustainability a legitimate claim that her basic

As we have now defined the subject (presently living persons), object (basic needs of present and future generation) and justification (sustainability as justice) of our notion of responsibility, we proceed with discussing the extent of responsibility. What are the limits of a person's responsibility for sustainability? There are two fundamental limits.

The first limit is the widely endorsed 'ought-implies-can' criterion according to which one can be only obliged to do what one actually can do.[7] Its rationale is that responsibility presupposes the possibility of compliance: "[a]ction-guiding principles must fit human capacities, or they become strange in a damaging way: pointless" (Griffin 1992: 123). The possibility of compliance implicates that a person has the power and the knowledge to comply. The power to comply refers to physical and mental abilities of the person as well as to the availability and effectiveness of instruments or resources. For example, imagine the situation of a drowning child. We do not hold a person responsible to save the child who is unable to swim, who is mentally paralyzed or who has no means to call somebody else to save the child. The knowledge to comply refers to situations in which a person cannot know the legitimate claims of others or the implications of her actions. In the example, a person cannot be held responsible to save the drowning child if she cannot know that the child is actually drowning. Hence, power and knowledge limit one's responsibility as they delineate the possibility of compliance. In this paper, we focus on the power to comply, as defined by a limited set of feasible actions, and leave questions related to knowledge for future research.[8]

The second limit concerns the legitimate claims of the person who bears responsibility. Conceptions of justice define legitimate claims of some individuals or collectives. In the case of Brundtland-sustainability, each member of the present and of future generations has a legitimate claim with respect to the satisfaction of his or her basic needs. To satisfy this claim is the responsibility of persons of the present generation. However, these persons have the same legitimate claim with respect to their own basic needs. It follows that there may arise a conflict between the obligation for the satisfaction of basic needs of others and the obligation for the satisfaction of one's own basic needs. But how exactly do the legitimate claims of a

---

needs are satisfied, because she is part of the present generation. However, we do not consider an infant as a responsible person, because she has not yet developed all characteristics of a person, such as reason and reflection. As an example of the latter, a business corporation that can be considered as a person-like entity because it has all the chartacteristics of a person (hence the name: "corporation") and, therefore, bears responsibility for sustainability, does not have any legitimate claim to the satisfaction of its "basic needs" because these are only defined for individual human beings.

[7] The ought-implies-can criterion goes back at least to Kant who maintained that responsibility as a duty or obligation presupposes the possibility of compliance: "it would not be a duty to strive after a certain effect of our will if the effect were impossible in experience" (Kant 1991: 62). Contemporary philosophers, such as Singer (1993) or Griffin (1992), argue in a similar way that it would be "absurd to say that we ought to do what we cannot do" (Singer 1993: 242).

[8] Krysiak (2009) discusses responsibility – yet, only in the primary meaning (i.e. ascription of consequences to actors) – for the case in which a present actor acts under uncertainty.

responsible person limit the responsibility for fulfilling the legitimate claims of others? An answer to this question is given by Singer’s principle.

**Singer’s principle**

The utilitarian ethicist Singer starts with the normative assumption that suffering – e.g. from lack of shelter or food or, more generally, from unsatisfied basic needs – is something bad. Singer’s principle then states that “if it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it“ (Singer 1972: 231). With “we”, Singer refers to persons in the sense of Locke (1959) as defined above. Hence, all persons are responsible to prevent suffering of others, e.g. from unsatisfied basic needs.

A crucial idea of Singer’s principle is that the claims of a responsible person are legitimate in limiting this responsibility only to the extent that they are “of comparable moral importance”. For instance, claims to consume cars, clothes, shoes, or concerts are, according to Singer, *not* of comparable moral importance compared to the basic needs of suffering persons. It follows that the obligation to prevent or remedy the suffering of others holds insofar as the responsible person is not also suffering from unsatisfied basic needs. More specifically, Singer defines that a responsible person ought to give[9] “to the point of marginal utility, at which by giving more one would cause oneself and one’s dependents as much suffering as one would prevent […]” (Singer 1972: 234). The point of marginal utility hence provides an explicit definition of the relation of legitimate claims of the responsible person and of those of other suffering persons. Responsibility for the latter extends up to the point where positive and negative marginal effects of giving more are equal.

Singer’s principle is a modified version of the utilitarian principle. It differs from standard utilitarianism as it states that minimizing suffering is morally more important than maximizing the satisfaction of wants, thus introducing a lexicographic ordering. In this sense, it is very well suited to specify the limits of responsibility for Brundtland-sustainability, as the latter only defines that basic needs should be satisfied and not what ought to be done beyond that point.

To apply the principle in the context of Brundtland-sustainability, we make the normative assumption that unsatisfied basic needs of present and future generations are something bad. All members of present and future generations suffer when their basic needs are not satisfied. With this assumption, we apply Singer’s principle and limit the responsibility of present persons to act responsibility by the point of marginal utility, at which by saving more resources for future generations present persons would cause themselves as much suffering as they would prevent in future generations.

[9] Singer discusses the context in which a person can remedy suffering of others by “giving” a donation. Hence his wording.

**Utilitarian notion of responsibility for sustainability**

To sum up, the imperative of sustainability cannot imply an absolute obligation to attain a particular (sustainable) state or development of the world. It does imply, though, the responsibility to use the best available knowledge and power to, according to Brundtland (WCED 1987), meet the needs of the present generation without compromising the ability of future generations to meet their needs. Our utilitarian notion of responsibility for sustainability can be summarized as follows:

*Presently living persons* are responsible for *meeting the basic needs* of the present generation and *not compromising the ability* of future generations to meet their basic needs to the extent of *presently living persons' possibility of compliance* and to the point of *marginal utility*.

# 3 Model

There are two non-overlapping generations $t = 1,2$.[10] Both have preferences over consumption $C_t$, represented by a utility function $U_t(C_t)$ which is characterized by positive and decreasing marginal utility.

In line with both Singer (1972) and Brundtland (WCED 1987), we assume that in the utility function, there is a distinction between consumption below and above a level $C^{\mathrm{BN}}$ at which basic needs are satisfied. $C^{\mathrm{BN}}$ is identical for both generations, normalized to 1, and yields a utility level $U_t\left(C^{\mathrm{BN}}\right) = U_t^{\mathrm{BN}}$. To the extent that their basic needs are not yet satisfied, that is for $C_t \leq C^{\mathrm{BN}}$, both generations have identical preferences. In terms of Singer's ethics, unsatisfied basic needs means that persons are suffering. The assumption thus states that any further unit of food, shelter or medicine has the same marginal effect on every suffering person. In other words, we assume persons to be equal in their suffering.

Beyond the threshold where basic needs are met, that is for $C_t > C^{\mathrm{BN}}$, their preferences may or may not be identical. The assumption of identical preferences below the basic needs level and diverging preferences beyond that point is not only central for Singer's ethics and the Brundtland notion of sustainability but is also in line with e.g. the arguments of Partridge (2003) who states that "[…] it is much easier to identify and address the causes of misery, than to promote the wellsprings of happiness. This is especially so with regard to the future. Their pains and ours can be traced to our common somatic needs and the status of the planetary ecosystem which sustains us both. Their pleasures and satisfactions will come from developments in culture, taste and technology that we cannot even imagine."

The utility functions are given by:

$$U_t(C_t) = \begin{cases} C_t^{\,\alpha} & \text{for } C_t \leq C^{\mathrm{BN}} \quad \text{for } t = 1,2 \\ C_t^{\,\alpha_t} & \text{for } C_t > C^{\mathrm{BN}} \quad \text{for } t = 1,2\ , \end{cases} \tag{1}$$

[10] For simplification, we assume that each generation consists of one representative person.

with $0 < \alpha_t < \alpha < 1$. Marginal utility from consumption is thus strictly larger if the basic needs are not met than if they are met. The utility functions are depicted in Figure 1:

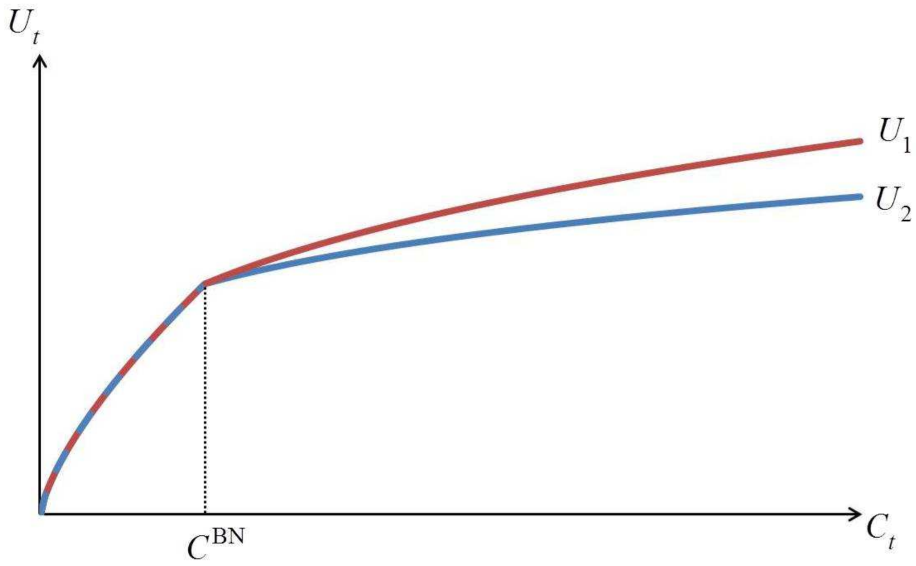



Figure 1: Utility functions

Consumption is being generated from the (consumptive) use of a resource stock $\overline{R} > 0$. This stock can be allocated between both generations such that each generation $t$ ($t$ = 1,2) has an endowment $R_t$:

$$R_1 + R_2 \leq \overline{R} \,. \tag{2}$$

Each generation has a simple linear production technology represented by the function:

$$C_1(R_1) = R_1 \,, \tag{3}$$

$$C_2(R_2) = \gamma R_2 \,. \tag{4}$$

$\gamma > 0$ is an exogenous factor which can be broadly interpreted: either as productivity change or as natural renewability/growth of the resource. There is no waste in production such that every unit produced will be consumed.

With these assumptions, there exists a minimal resource endowment $R^{\min}$ which exactly allows both generations to satisfy their basic needs:

$$R^{\min} = 1 + \frac{1}{\gamma} \,. \tag{5}$$

# 4 Definitions

Within this model, we now define resource allocations $(R_1, R_2)$ to be sustainable, responsible, Pareto-efficient, and Discounted-utilitarian-welfare-maximal. Further, we characterize each of these resource allocations with necessary and sufficient conditions. An allocation is *feasible* if the sum of the resource endowments is not larger than the total resource stock $\overline{\mathrm{R}}$ (Eq. (2)).

In line with the Brundtland-conception (WCED 1987), sustainable allocations are defined as meeting the basic needs of both generations.

**Definition 1 (Sustainable allocations)**
A feasible allocation $(R_1, R_2)$ is called *sustainable* if and only if it yields for all $t$ =1,2

$$C_t(R_t) \geq C^{\mathrm{BN}} = 1\ . \tag{6}$$

With this definition, sustainable allocations are characterized as follows.

**Lemma 1**
A feasible allocation $(R_1, R_2)$ is *sustainable* if and only if it meets the following conditions:

$$R_1 \geq 1 \text{ and } R_2 \geq 1/\gamma \tag{7}$$

The conditions for sustainable allocations are intuitive: both generations need a minimal resource endowment, as defined by Eq. (7), to be able to satisfy their basic needs. The minimal endowment of the second generation is contingent on $\gamma$. If $\gamma$ is large (e.g. due to high technological progress or natural resource growth), the second generation needs a small share of the resource. A small $\gamma$ (e.g. due to ecological degradation) requires a large resource share for the second generation. Further, Eq. (7) shows that existence of sustainable allocations requires that $\overline{R} \geq R^{\mathrm{min}}$.

Applying our notion of responsibility developed in Section 2 to this notion of sustainability, we continue with the formal definition and necessary and sufficient conditions of responsible allocations.

**Definition 2 (Responsible allocations)**
A feasible allocation $(R_1,R_2)$ is called *responsible* if and only if it yields for all $t$ =1,2

$$C_t(R_t) \geq C^{\mathrm{BN}} = 1 \qquad \text{for } \overline{R} \geq R^{\mathrm{min}}, \tag{8}$$

$$\frac{\mathrm{d}U_1\big(C_1(R_1)\big)}{\mathrm{d}C_1}\frac{\mathrm{d}C_1(R_1)}{\mathrm{d}R_1} = \frac{\mathrm{d}U_2\big(C_2(R_2)\big)}{\mathrm{d}C_2}\frac{\mathrm{d}C_2(R_2)}{\mathrm{d}R_2} \text{ and } \overline{R} = R_1 + R_2 \qquad \text{for } \overline{R} < R^{\mathrm{min}}. \tag{9}$$

Our definition of responsible allocations distinguishes situations in which it is feasible to satisfy the basic needs of both generations (Eq. (8)), and situations in which this is not feasible (Eq. (9)). If it is feasible, obviously all allocations in which basic needs of both generations are satisfied, are responsible. However, if the resource stock is too small, there still exists a responsible allocation: the whole resource stock must be allocated such that there are equal marginal utilities from consumption. This ensures that suffering in the sense of Singer is minimized.

With this definition, responsible allocations are characterized as follows.

**Lemma 2**

A feasible allocation $(R_1,R_2)$ is *responsible* if and only if it meets the following conditions:

$$R_1 \geq 1 \text{ and } R_2 \geq 1/\gamma \quad \text{for } \overline{R} \geq R^{\min} , \tag{10}$$

$$R_1 = \gamma R_2 \text{ and } \overline{R} = R_1 + R_2 \quad \text{for } \overline{R} < R^{\min} . \tag{11}$$

Lemma 2 shows that in the characterization of responsible allocations one needs to distinguish two cases: one (Eq. (10)) in which attaining the underlying normative aim (here: sustainability) is feasible, and one (Eq. (11)) in which it is not.

Now we define Pareto-efficient allocations.

**Definition 3 (Pareto-efficient allocations)**

A feasible allocation $(R_1,R_2)$ is called *Pareto-efficient* if and only if there does not exist another feasible allocation $(R_1',R_2')$ such that $U_t\big(C_t(R_t')\big) \geq U_t\big(C_t(R_t)\big)$ for all $t$ =1,2 and $U_t\big(C_t(R_t')\big) > U_t\big(C_t(R_t)\big)$ for at least one $t$.

With this definition, Pareto-efficient allocations are characterized as follows.

**Lemma 3**

A feasible allocation $(R_1,R_2)$ is *Pareto-efficient* if and only if it meets the following condition:

$$\overline{R} = R_1 + R_2 . \tag{12}$$

Since our model consists of one resource which can only be transformed into one good, and there are no externalities, all allocations which use the entire resource stock $\overline{R}$ must be Pareto-efficient.

Next we define allocations which are a discounted-utilitarian-welfare maximum.

**Definition 4 (Discounted-utilitarian-welfare maximum)**

A feasible allocation $(R_1, R_2)$ is called a *discounted-utilitarian-welfare maximum* if and only if it solves:

$$\max_{R_1,R_2} W = U_1\big(C_1(R_1)\big) + \delta U_2\big(C_2(R_2)\big) \text{ s.t. } \overline{R} = R_1 + R_2 \tag{13}$$

In this definition, $\delta \geq 0$ is a discount factor which is the weight of the utility of the second generation in the overall welfare function. The special case of $\delta = 1$ means that no discounting takes place.

With this definition, discounted-utilitarian-welfare maxima are characterized as follows.

**Lemma 4**

A feasible allocation $(R_1, R_2)$ is a *discounted-utilitarian-welfare maximum* if and only if it meets the following condition:

$$\alpha_1(R_1)^{\alpha_1} = \delta\alpha_2(\gamma R_2)^{\alpha_2} \text{ and } \overline{R} = R_1 + R_2 \quad \text{for } \overline{R} \geq R^{min}, \tag{14}$$

$$(R_1)^{\alpha} = \delta(\gamma R_2)^{\alpha} \text{ and } \overline{R} = R_1 + R_2 \quad \text{for } \overline{R} < R^{min}. \tag{15}$$

Discounted-utilitarian-welfare maxima are characterized by equal discounted marginal utility of both generations. Marginal utility of the second generation is weighed differently by the discount factor than marginal utility of the first generation.

Our analysis adds to the discussion about the ethical legitimacy of discounting. In general, there are three reasons for discounting (Gollier 2010). First, there is individual or societal impatience or pure time preference. Yet, ever since Pigou (1920) it is clear that while this argument may describe actual human behavior, it cannot be used normatively to justify discounting. Second, there is the assumption of decreasing marginal utility and future economic growth (Ramsey 1928). If there is higher consumption available in the future due to economic growth, and if marginal utility is decreasing with the level of consumption, intergenerational equity allows for discounting. Third, uncertainty about future outcomes allows for discounting as it makes future well-being uncertain. All taken together, there seems to be some ethical legitimacy in discounting, also in normative criteria of societal choice, at least to a certain extent.

# 5 Results

In this section, we present our results. First, we discuss the properties of responsible allocations. Further, we relate the necessary and sufficient conditions for responsible allocations with the conditions for sustainable, Pareto-efficient, and discounted-utilitarian-welfare maximum allocations.

**Proposition 1 (Reponsibility)**
If $\overline{R} \geq R^{min}$, there exist infinitely many responsible allocations, characterized by Condition (10). If $\overline{R} < R^{min}$, there exists a single responsible allocation, characterized by Condition (11).

*Proof*: Eq. (10) shows that there are infinite responsible allocations iff $\overline{R} \geq R^{min}$. Eq. (11) shows that there exists one responsible allocation iff $\overline{R} < R^{min}$.

This means, that in any case there exists a responsible allocation. If sustainability is feasible, that is if $\overline{R} \geq R^{min}$, there exist infinitely many responsible allocations. This is due to the Brundtland notion of sustainability which is blind for distributional aspects once all basic needs are satisfied. Our notion of responsibility adds to this as it defines one responsible allocation for $\overline{R} < R^{min}$ when sustainability is not feasible. At this allocation, $\overline{R}$ must be used completely ($\overline{R} = R_1 + R_2$) and marginal utilities from consumption must be equal as required by Singer's principle (which is the case for $R_1 = \gamma R_2$).

**Proposition 2 (Sustainability)**

If $\overline{R} \geq R^{min}$, each responsible allocation is also sustainable, and vice versa. In contrast, if $\overline{R} < R^{min}$, the responsible allocation is not sustainable. Responsibility for sustainability is, hence, equivalent to sustainability if and only if sustainability is feasible.

*Proof*: Eq. (10) shows that there are infinitely many responsible allocations for $\overline{R} \geq R^{min}$. Comparison of Eq. (7) with (10) shows that all allocations satisfying Eq. (10) must also satisfy Eq. (7). Comparison of Eq. (7) with (11) shows that an allocation satisfying Eq. (11) cannot satisfy Eq. (7). □

Our model illustrates the common and diverging properties of the criteria of sustainability and of responsibility for sustainability. They are equivalent whenever sustainability is feasible. If it is not, they differ since then a responsible allocation exists while a sustainable allocation does not exist. The criterion of responsibility thus provides action guidance even if it is not feasible to attain the underlying normative objective (here: sustainability).

**Proposition 3 (Pareto-efficiency)**

If $\overline{R} \geq R^{min}$, there exist some responsible allocations which are also Pareto-efficient. These are characterized by

$$\overline{R} = R_1 + R_2 \text{ and } R_1 \geq 1 \text{ and } R_2 \geq 1/\gamma \,. \tag{16}$$

Neither are all responsible allocations Pareto-efficient nor are all Pareto-efficient allocations responsible. If $\overline{R} < R^{min}$, the responsible allocation, which is characterized by Condition (11), is Pareto-efficient.

*Proof:* Comparison of Eq. (12) with (10) shows that some but not all allocations satisfying Eq. (10) also satisfy Eq. (12), e.g. $R_1 = 1$ and $R_2 = 1/\gamma$ for $\overline{R} > R^{min}$ satisfies Eq. (10) but not Eq. (12) while all $R_1 = 1 + \varepsilon$ and $R_2 = 1/\gamma$ for $\overline{R} = R^{min} + \varepsilon$ with $\varepsilon \geq 0$ satisfy Eq. (10) and Eq. (12). Comparison of Eq. (12) with Eq. (11) shows that an allocation satisfying Eq. (11) also satisfies Eq. (12), as Eq. (12) is part of Eq. (11). But not all allocations satisfying Eq. (12) satisfy Eq. (11), e.g. $R_1 = 1 - \varepsilon$ and $R_2 = 1/\gamma$ for $\overline{R} = R^{min} - \varepsilon$ with $\varepsilon \geq 0$. Eq. (16) follows straightforwardly from Eq. (12) and Eq. (7). □

Since the Brundtland notion of sustainability does not require Pareto-efficiency, the criterion of responsibility for sustainability does not require Pareto-efficiency if and only if sustainable allocations are feasible. The Brundtland notion merely defines a minimum standard and allows for wasteful allocations once the standard is achieved.

If sustainability is not feasible, the criterion of responsibility requires Pareto-efficiency in order to minimize suffering in the sense of Singer.

**Proposition 4 (Discounted-utilitarian-welfare maximum)**

There uniquely exists a discounted-utilitarian-welfare maximum, characterized by Condition (14) or (15). If no discounting takes place, $\delta = 1$, the discounted-utilitarian-welfare-maximum allocation is responsible. If, in contrast, discounting takes place, $\delta \neq 1$, the following holds:

For $\overline{\mathrm{R}} \geq \mathrm{R}^{\min}$, the discounted-utilitarian-welfare maximum is a responsible allocation iff

$$\delta^{\min} \leq \delta \leq \delta^{\max}. \tag{17}$$

with $\delta^{\min} = \frac{\alpha_1}{\alpha_2}\left(\overline{\mathrm{R}} - \frac{1}{\gamma}\right)^{\alpha_1 - 1}$ and $\delta^{\max} = \frac{\alpha_1}{\alpha_2}\left(\gamma(\overline{\mathrm{R}} - 1)\right)^{1-\alpha_2}$.

For $\overline{\mathrm{R}} < \mathrm{R}^{\min}$, the discounted-utilitarian-welfare maximum is not a responsible allocation.

*Proof*: For $\delta = 1$, comparison of Eq. (15) with Eq. (10) shows that all allocations satisfying Eq. (15) must also satisfy Eq. (10). The same holds for Eq. (14) and Eq. (10) since $\alpha_t < \alpha$ for all *t* =1,2. For $\delta \neq 1$, using $\mathrm{R}_2 \geq 1/\gamma$ from Eq. (10) and $\mathrm{R}_1 \geq 1$ in Eq. (14) yields Eq. (17). Comparison of Eq. (15) with Eq. (11) shows that an allocation satisfying Eq. (15) cannot satisfy Eq. (11). □

Let us first discuss the case without discounting, that is $\delta = 1$. If the resource stock is large enough ($\overline{\mathrm{R}} \geq \mathrm{R}^{\min}$) so that sustainable allocations exist, the discounted-utilitarian-welfare maximum must be sustainable and responsible since marginal utility of both generations is strictly larger when the basic needs are satisfied (see Eq. (1)). Any non-sustainable allocation, therefore, cannot be a discounted-utilitarian-welfare maximum. As there exist infinitely many sustainable and responsible allocations in this case, the discounted-utilitarian-welfare maximum is merely one out of many responsible allocations. If no sustainable allocations exist (i.e. $\overline{\mathrm{R}} < \mathrm{R}^{\min}$), Singer's principle requires that responsible allocations minimize suffering which is simply a negative formulation of maximizing happiness and thus of the principle of Utilitarianism. It follows that the responsible allocation in this case must be a discounted-utilitarian-welfare maximum.

Now, let us discuss discounting, that is $\delta \neq 1$. Discounting yields a sustainable and responsible allocation if and only if there exist sustainable allocations and the discount rate is within the range specified by Condition (17). The intuition is as follows. The Brundtland notion of sustainability merely defines a minimum standard of sustainability as satisfied basic needs. If this standard is feasible, discount rates that do not favor any generation too strongly yield sustainable allocations. Discount rates not satisfying Condition (17) however, yield allocations in which the basic needs of one generation cannot be satisfied and which are thus neither sustainable nor responsible.

The range specified by Condition (17) has the following intuitive properties. Intuitively, large technological progress ($\gamma$) allows for larger discounting of future utility to be responsible. A large resource stock ($\overline{\mathrm{R}}$) allows for a large discounting in general. Further, a large (small) ratio of $\alpha_1/\alpha_2$ allows for larger (smaller) discounting of future utility to be responsible, as it implies that marginal utility of the first generation is higher than of the second generation.

If the resource stock is so small ($\bar{R} < R^{min}$) that no sustainable allocation exists, discounting is not responsible. Any unequal valuation of utility between generations will not minimize suffering and, therefore, cannot be responsible. This result is interesting in light of the two ethically acceptable arguments for discounting: consumption growth with decreasing marginal utility, and uncertainty.

The argument of growth with decreasing marginal utility cannot be upheld in favor of discounting if sustainability is not feasible, because it is already included in the criterion of equal marginal utility. If there is growth in terms of a large $\gamma$, Eq. (11) shows that this yields a larger resource share for the first generation in the responsible allocation. Any further discounting can thus not be justified with this argument.[11] The case of uncertainty is different. In our model, we assume that there is no uncertainty. Uncertainty may thus not be an argument for discounting in our model. However, incorporating uncertainty, about e.g. $\gamma$, in the model, might very well justify discounting when sustainability is not feasible.

Figure 2 summarizes our main results for $\bar{R} \geq R^{min}$:

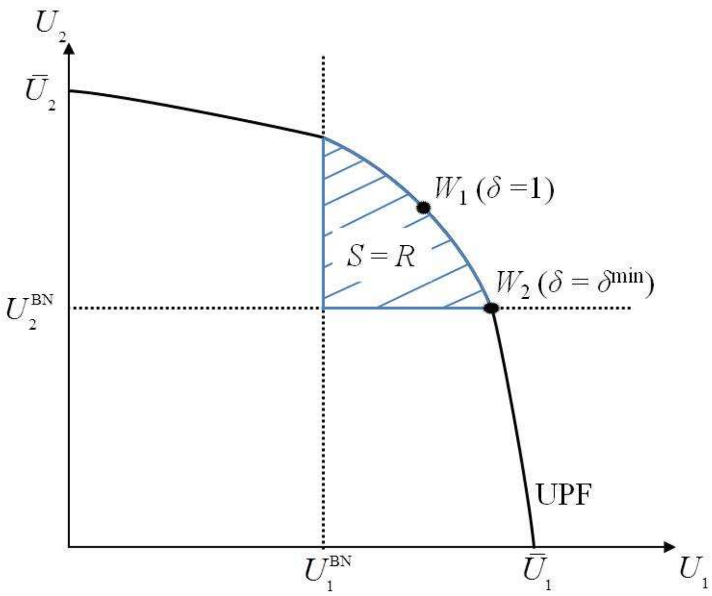



Figure 2: Illustration for $\bar{R} \geq R^{min}$ of responsible ($R$, dashed area), sustainable ($S$, dashed area), Pareto-efficient (UPF) and discounted-utilitarian-welfare maximum (for example $W_1$ and $W_2$) allocations.

The utility possibility frontier (UPF) – the curve connecting $(0, \bar{U}_2)$ and $(\bar{U}_1, 0)$ – contains all Pareto-efficient allocations. On the UPF, we find the discounted-utilitarian-welfare maxima

[11] See also Roemer (2011: 374): "And if [future] societies are indeed 'richer,' because of the technological progress that takes place […] and because we have saved the global commons for them, and it turns out that the optimal policy has their consuming more than we do, their average unit of consumption will not receive as much weight in the social-welfare function as our average unit […], which implements diminishing marginal utility. Why further discount their utility with positive discount rates?"

(for example $W_1$ and $W_2$). $W_1$ represents the special case of no discounting, that is $\delta = 1$. $W_2$ represents the special case of discounting utility of the future generation ($\delta = \delta^{\min} < 1$) such that $U_2 = U_2^{BN}$. In general, all discounted-utilitarian-welfare maxima lie on the UPF, with their exact position determined by the discount rate. Discounting of future utility ($\delta < 1$) leads to a discounted-utilitarian-welfare maximum somewhere on the UPF between $W_1$ and $(\bar{U}_1, 0)$. If the discount rate decreases below $\delta^{\min}$, the discounted-utilitarian-welfare maximum yields an allocation which is neither sustainable nor responsible.[12] Allocations that are sustainable (*S*) and responsible (*R*) are depicted by the dashed area which consists of the triangle delimited by $U_1 = U_1^{BN}$, $U_2 = U_2^{BN}$ and the UPF.

The picture changes fundamentally for $\bar{R} < R^{\min}$ as shown in Figure 3:

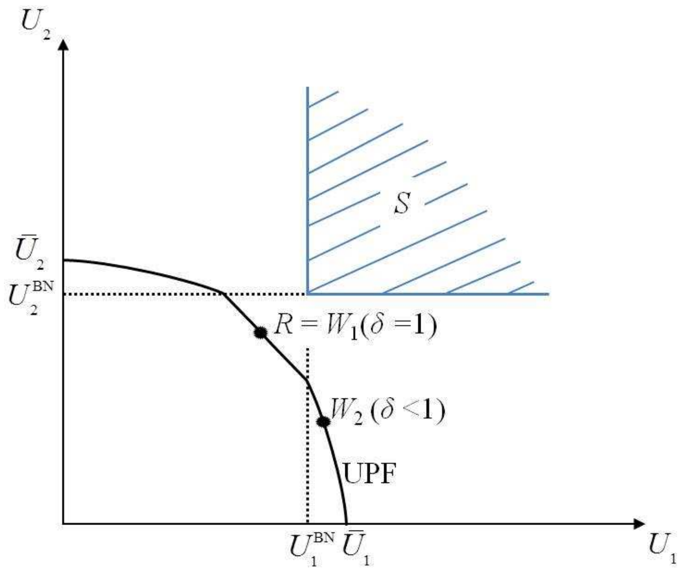



Figure 3: Illustration for $\bar{R} < R^{\min}$ of responsible (*R*), sustainable (*S*, dashed area), Pareto-efficient (*UPF*), and discounted-utilitarian-welfare maximum (for example $W_1$ and $W_2$) allocations.

Again, the UPF is connecting $(0, \bar{U}_2)$ and $(\bar{U}_1, 0)$ and contains all Pareto-efficient allocations. However, there is only one responsible allocation (*R*) which equals the discounted-utilitarian welfare maximum $W_1$ for $\delta = 1$. *R* and $W_1$ lie on the UPF but below satisfied basic needs levels. Since sustainability is not feasible, there are no sustainable allocations on or below the UPF, but all "sustainable" allocations (*S*, dahed area) would lie outside the UPF. We further see that for discounting the utilty of the future generation, $\delta < 1$, $W_2$ lies on the UPF somewhere between *R* (=$W_1$) and $(\bar{U}_1, 0)$, with the exact position again depending on the discount rate $\delta$. Analogously, discounting of utility of the present generation, $\delta > 1$, leads to a discounted-utilitarian welfare maximum on the UPF somewhere between *R* and $(0, \bar{U}_2)$. As any $\delta \neq 1$ yields a discounted-utilitarian-welfare maximum below or above *R*, the discounted

[12] The same reasoning applies for $\delta > 1$, i.e. discounting of utility of the present generation.

utilitarian welfare maximum cannot be responsible. But as shown in Figure 3, it may lead to an allocation in which one generation has its basic needs satisfied.

## 6 Conclusion

We have developed and formalized a utilitarian notion of responsibility which is inspired by Singer's (1972) principle and the Brundtland Commission's notion of sustainability (WCED 1987). Our results show that sustainabilty and responsibility for sustainability are equivalent if and only if sustainability is feasible. If it is not, there still exists a responsible allocation which is also Pareto-efficient. Further, the utilitarian welfare maximum without discounting always fulfills the criterion of responsibility. Discounting may be responsible to a certain extent if sustainability is feasible. If sustainability is not feasible, discounting is not responsible.

Our analysis demonstrates that reponsibility can be clearly and unambiguously conceptualized in economic models. Such a conception of responsibility is, albeit simple, neither trivial nor redundant, but adds specificity to the discussion about sustainability in two respects: (1) it clearly specifies how to act if sustainability is not feasible; (2) in any case, it specifies the balance between legitimate claims of present and future generations.

With these achievements, also the limits of our analysis are clear: we have built on a specific idea of sustainability and on a specific ethics, both of which focus on the satisfaction of basic needs (and, thus, go together very well). For other aspects of sustainability they are less well suited, and other notions of responsibility will be needed. More specifically, our results are essentially driven by the assumptions of the basic needs concept in Singer's ethics and in the Brundtland notion of sustainability: there exists a basic needs threshhold which is identical for all human beings and below which preferences are identical.

The conceptualization of responsibility with our approach lays out a broad basis for future research. In particular, the aspects of the possibility of compliance, namely the power and knowledge to comply, should be analyzed more deeply. With respect to the power to comply, there is the question of how the present generation can ensure that future generations are able to satisfy their basic needs given that the presently living persons have several options. With respect to knowledge, there immediately arises the problem of uncertainty, e.g. about technological progress, which affects the responsibility of the present generation. Uncertainty further raises the question of how much the present generation ought to invest in the reduction of uncertainty. We think that our approach can be helpful in adressing these issues.

## Acknowledgements

We are grateful to Nikolai Hoberg, Martin F. Quaas, Klara Helene Stumpf and Thomas Wein for critical discussion.

## References


Baumgärtner, S., Faber, M. and Schiller, J. (2006). *Joint Production and Responsibility in Ecological Economics. On the Foundations of Environmental Policy*. Edward Elgar, Cheltenham.

Baumgärtner, S., Klauer, B., Petersen, T., Quaas, M.F., Schiller, J. and Stumpf, K.H. (2010). Sustainability under uncertainty as responsibility. Paper presented at the *Biennial Conference of the International Society for Ecological Economics* (ISEE), Oldenburg/Bremen, August 22–25, 2010.

Brekke, K.A., Kverndokk, S. and Nyborg, K. (2003). An economic model of moral motivation. *Journal of Public Economics* 87: 1967–1983.

Fleurbaey, M. (2008). *Fairness, Responsibility, and Welfare*. Oxford University Press, Oxford.

Frey, B.S. (1997). *Not Just for the Money. An Economic Theory of Personal Motivation*. Edward Elgar, Cheltenham, UK.

Gertler, B. (2010). *Self-Knowledge*. Routledge, New York.

Gollier, C. (2010). Ecological discounting. *Journal of Economic Theory* 145: 812–829.

Griffin, J. (1992). The human good and the ambitions of consequentialism. *Social Philosophy and Policy* 9(2): 118–132.

Heyes, A. and Kapur, S. (2011). Regulating altruistic agents. *Canadian Journal of Economics/Revue canadienne d'économique* 44(1): 227–246.

[ILO] International Labour Organization (1976). *Employment, Growth and Basic Needs: A One-World Problem*. *Report of the Director-General of the International Labour Office*. Geneva.

Jonas, H. (1979). *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation*. Insel-Verlag, Frankfurt am Main. English translation: *The Imperative of Responsibility: In Search of an Ethics for the Technological Age*. Chicago University Press, Chicago. 1984.

Kant, I. (1992 [1793]). *Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis*. 5th Edition. Klostermann, Frankfurt am Main. English translation: On the common saying: This may be true in theory, but it does not apply in practice. In: *Political Writings*. Cambridge University Press, Cambridge. 1991: 61–92.

Krysiak, F.C. (2009). Sustainability and its relation to efficiency under uncertainty. *Economic Theory* 41: 297–315.

Locke, J. (1959 [1690]). *An Essay Concerning Human Understanding*. Dover, New York.

Max-Neef, M. (1991). *Human Scale Development: Conception, Application, and Further Reflections.* Apex Press, London.

Nyborg, K. and Rege, M. (2003). Does public policy crowd out private contributions to public goods? *Public Choice* 115: 397–418.

Partridge, E. (2003). In search of sustainable values. *International Journal of Sustainable Development* 6(1): 25–41.

Pigou, A.C. (1920). *The Economics of Welfare*. Macmillan, London.

Ramsey, F.P. (1928). A mathematical theory of saving. *The Economic Journal* 38: 543–559.

Roemer, J. (2011). The ethics of intertemporal distribution in a warming planet. *Environmental and Resource Economics* 48(3): 363–390.

Singer, P. (1972). Famine, affluence, and morality. *Philosophy and Public Affairs* 1(3): 229–243.

Singer, P. (1993). *Practical Ethics*. Cambridge, New York.

Vallentyne, P. (2008). Brute luck and responsibility. *Politics, Philosophy & Economics* 7: 57–80.

[WCED] World Commission on Environment and Development (1987). *Our Common Future*. Oxford University Press, Oxford.

Williams, G. (2008). Responsibility as a virtue. *Ethical Theory and Moral Practice* 11(4): 455–470.

# Paper 2: Verantwortung für Nachhaltigkeit von Konsumenten

JOACHIM FÜNFGELT*

Department of Sustainability Sciences and Department of Economics, Leuphana Universität Lüneburg, Deutschland

**Abstract**: Während weitestgehend Einigkeit darüber herrscht, dass Konsumenten sich für politische Entscheidungen im Sinne der Nachhaltigkeit einsetzen sollten, besteht eine Kontroverse hinsichtlich ihrer Verantwortung, nachhaltig zu konsumieren. In diesem Artikel untersuche ich daher die Frage, welche Verantwortung für Nachhaltigkeit Konsumenten in ihren Konsumhandlungen tragen. Dafür konzeptionalisiere ich Nachhaltigkeit als Gerechtigkeit und leite legitime Ansprüche ab, die heute und zukünftig lebende Menschen geltend machen können. Die Erfüllung dieser Ansprüche sicher zu stellen, ist das Ziel der Verantwortung für Nachhaltigkeit. Zu diesem Zweck gehören zur Verantwortung für Nachhaltigkeit drei Verpflichtungen: die Verpflichtungen zu *vermeiden*, zu *schützen* und zu *helfen* (Shue 1996). Bezüglich ihrer Konsumhandlungen sind Konsumenten dafür verantwortlich, Verletzungen von Nachhaltigkeitsansprüchen zu *vermeiden*. Die Verpflichtungen zu *schützen* und zu *helfen* können Konsumenten durch ihre Konsumhandlungen oder durch politisches Engagement erfüllen.

**JEL Classification**: P46, Q56

**Keywords:** Moral responsibility, consumer responsibility, sustainability, sustainable consumption, political responsibility

---

* Sustainability Economics Group, Leuphana University of Lüneburg, P.O. Box 2440, D-21314 Lüneburg, Germany, phone: +49.4131.677-2668, fax: +49.4131.677-1381, email: fuenfgelt@uni.leuphana.de

# 1 Einleitung

Konsum ist ein alltäglicher Bestandteil unseres Lebens. Jeder Mensch muss konsumieren, selbst wenn er lediglich um die Befriedigung von Grundbedürfnissen wie Ernährung, Bekleidung oder Unterkunft bemüht ist (König 2008). Im Rahmen der öffentlich und wissenschaftlich geführten Nachhaltigkeitsdebatte stellt sich daher die Frage, ob und inwiefern Menschen in ihrer Rolle als Konsument Verantwortung für Nachhaltigkeit tragen.

In der öffentlich geführten Debatte gehen die Meinungen dabei weit auseinander. Während eine Seite von „Zwangsernährung mit ökologisch korrekten Produkten“ (von Weizsäcker 2005) und „Ökotyrannei“ (von Petersdorf 2011) spricht und somit eine Verantwortung für Nachhaltigkeit von Konsumenten ablehnt, betont die andere Seite die Möglichkeit und damit Pflicht der Konsumenten, zur Nachhaltigkeit beizutragen (Langrock-Kögel 2012). Der jüngst zum Deutschen Bundespräsidenten gewählte Joachim Gauck sieht in der Verantwortung dabei gar keine Last, sondern im Gegenteil „eine der schönsten und größten Möglichkeiten des menschlichen Daseins“ (Gauck 2012).

Auch die wissenschaftliche Diskussion wird polemisch geführt. Insbesondere eine seit 2010 in der Zeitschrift *GAIA* geführte Debatte zeigt, wie umstritten die Verantwortung von Konsumenten ist. Grunwald (2010) kritisiert die Zuschreibung von Verantwortung an Konsumenten, da dies erstens nicht zielführend sei, zweitens Konsumenten aufgrund ihres begrenzten Wissens überfordere und drittens die individuelle Freiheit der Konsumenten gefährde. Nachhaltigkeit sei nach Grunwald primär eine politische Aufgabe, die sich nicht auf private Konsumhandlungen erstrecke. Petersen und Schiller (2011) greifen Grunwalds Argumentation auf. Sie sehen Nachhaltigkeit als Gegenstand politischer Verantwortung, weil es ein Ziel der politischen Gemeinschaft sei (Petersen und Schiller 2011: 160). Eine moralische Verantwortung, eindeutig nachhaltigkeitsschädigende Folgen zu unterlassen, schreiben Petersen und Schiller den Konsumenten allerdings zu. Eine gegenläufige Position nehmen Billharz et al. (2011) ein. Sie argumentieren, dass durch die Erweiterung der Handlungsspielräume für Konsumenten, nachhaltiger Konsum effektiv zur Erreichung der Nachhaltigkeit beitragen kann und sollte. Siebenhüner (2011) sieht zudem in nachhaltigem Konsum eine Möglichkeit, politische Entscheidungen zu beeinflussen.

Der einzige Konsens in der wissenschaftlichen Debatte besteht darin, dass Konsumenten politisch aktiv werden und sich für politische Entscheidungen im Sinne der Nachhaltigkeit einsetzen sollten. Doch worin die Verantwortung von Konsumenten bezüglich des Konsums besteht wird in der Debatte nicht eindeutig beantwortet. Der Grund ist, dass Nachhaltigkeit als Rechtfertigung von Verantwortung unklar bleibt und damit auch eine klare Analyse der Verantwortung für Nachhaltigkeit nicht ohne Weiteres möglich ist.

Dieser Aufsatz verfolgt daher das Ziel die Verantwortung von Konsumenten bezüglich ihrer Konsumhandlungen zu klären. Dafür konzeptionalisiere ich zunächst Nachhaltigkeit als intra- und intergenerationelle Gerechtigkeit. Als solche begründet Nachhaltigkeit legitime Ansprüche von heute und zukünftig lebenden Menschen (Ott und Döring 2008). Diese Ansprüche haben die Verantwortung zur Folge, ihre Erfüllung sicher zu stellen (Shue 1988).

Um die Erfüllung der Ansprüche wirklich sicher stellen zu können, gehören nach Shue (1996: 52) zu dieser Verantwortung drei Verpflichtungen: Erstens, die Verpflichtung die Verletzung legitimer Ansprüche zu *vermeiden*. Zweitens, die Verpflichtung vor der Verletzung legitimer Ansprüche zu *schützen*. Drittens, die Verpflichtung denjenigen, deren legitimen Ansprüche verletzt wurden, zu *helfen*. Anhand dieser Konzeption von Verantwortung für Nachhaltigkeit werde ich untersuchen, welche der genannten Verpflichtungen zur Verantwortung von Konsumenten bezüglich ihrer Konsumhandlungen gehören.

Dieser Aufsatz ist dafür folgendermaßen gegliedert. In Kapitel 2 konzeptionalisiere ich Nachhaltigkeit als Gerechtigkeit und begründe damit eine Verantwortung für Nachhaltigkeit. In Kapitel 3 spezifiziere ich diese Verantwortung für Nachhaltigkeit allgemein. In Kapitel 4 verwende ich diese Spezifikation zu Analyse der Verantwortung von Konsumenten. In Kapitel 5 ziehe ich ein Fazit.

## 2 Nachhaltigkeit als Gerechtigkeit

Wenn von einer Verantwortung für Nachhaltigkeit die Rede ist, liegt eine handlungsleitende, normative Idee der Nachhaltigkeit zugrunde. Diesen normativen Charakter erhält die Nachhaltigkeit in Anlehnung an die Idee der Gerechtigkeit. Daher ist es sinnvoll, Nachhaltigkeit als Gerechtigkeit zu spezifizieren, um daraus eine Verantwortung für Nachhaltigkeit abzuleiten.

Regeln und Grundsätze der Gerechtigkeit gehen auf legitime Ansprüche („claims”) zurück, die jemand („Anspruchsberechtigter“) gegenüber anderen geltend gemacht hat und die von anderen als berechtigt angesehen wurden (Ott und Döring 2008: 59). So können zum Beispiel zukünftige Generationen gegenüber der heutigen Generation den Anspruch geltend machen, ihnen ein stabiles Klima und funktionierende Ökosysteme zu hinterlassen.[1] Aus solchen Ansprüchen ergibt sich dann die heutige Generation die Verantwortung, diese legitimen Ansprüche zu erfüllen.

Je nachdem, wie Anspruchsberechtigte und deren legitimen Ansprüche spezifiziert werden, ergibt sich daraus eine andere Verantwortung. Es ist also zunächst zu klären, welche Akteure welche legitimen Ansprüche aufgrund der Nachhaltigkeit als Gerechtigkeit geltend machen können. Hier bietet die Nachhaltigkeits-Literatur ein breites Spektrum an Spezifikationen. Da es den Rahmen dieses Artikels sprengen würde, diese zu diskutieren, nehme ich eine Spezifikation im Sinne eines Minimalkonsenses über Nachhaltigkeit vor. Ein solcher Minimalkonsens sollte alle Anspruchsberechtigten und deren legitimen Ansprüche („Nachhaltigkeitsansprüche“) beinhalten, die jede Theorie der Nachhaltigkeit als Gerechtigkeit mindestens beinhaltet.

Zu den Anspruchsberechtigten gehören mindestens alle heute und zukünftig lebenden Menschen, da es „in jeder Nachhaltigkeitstheorie […] zentral um intra- und intergenerationale

[1] Es ist dabei nicht relevant, dass künftige Generationen ihre Ansprüche nicht selbst geltend machen können. Dies kann durch Fürsprecher aus der heutigen Generation geschehen (Ott und Döring 2008: 62).

Gerechtigkeit [geht]“ (Ott und Döring 2008: 45).[2] Ein Minimalkonsens der Nachhaltigkeit ist also anthropozentrisch und lässt der Natur eine instrumentelle Rolle zur Befriedigung der Nachhaltigkeitsansprüche von Menschen der heutigen und der zukünftigen Generationen. Gleichwohl ist selbst ein Minimalkonsens der Nachhaltigkeit ein sehr umfassendes Gerechtigkeitsverständnis, da alle heute und zukünftig lebenden Menschen zu den Anspruchsberechtigten gehören.

Als Minimalkonsens über Nachhaltigkeitsansprüche kann die Definition aus dem Brundtland-Report von 1987 gelten (Petersen 2009). Nach dieser ist eine Entwicklung nachhaltig, wenn sie “[…] die Bedürfnisse der Gegenwart befriedigt, ohne zu riskieren, dass künftige Generationen ihre eigenen Bedürfnisse nicht befriedigen können“ (Hauff 1987: 46). Mit Bedürfnissen sind „insbesondere die Grundbedürfnisse der Ärmsten der Welt“ (Hauff 1987: 46) gemeint. Heute und zukünftig lebende Menschen, insbesondere die Ärmsten, haben also einen Nachhaltigkeitsanspruch darauf, dass die eigenen Grundbedürfnisse befriedigt werden können. Dies ist ein absoluter Standard, der jeder Person und jeder Generation die Bedingungen eines menschenwürdigen Lebens sichert (Ott und Döring 2008: 80).[3] Der Begriff der Grundbedürfnisse kann dabei „basic needs“-Ansatzes (ILO 1976) oder anhand des Fähigkeitenansatzes (Sen 1986) eingeführt werden. Der wichtigste Unterschied liegt in einer unterschiedlichen Auffassung darüber, wie ein menschenwürdiges Leben gesichert wird. Der „basic needs“-Ansatz setzt voraus, dass die Sicherung eines menschenwürdigen Lebens direkt von Nutzen und Konsumgütern (wie Trinkwasser, Nahrung oder medizinischer Versorgung) ausgeht (ILO 1976). Der Fähigkeitenansatz hingegen nimmt an, dass die Sicherung eines menschenwürdigen Lebens als eine Angelegenheit von Funktionen und Fähigkeiten (wie die Fähigkeit, eine gute Gesundheit zu haben oder die Fähigkeit zur sozialen Interaktion) betrachtet werden muss (siehe z.B. Sen 1986 oder Nussbaum 2007). Die weitere Analyse dieses Artikels gilt für beide Ansätze gleichermaßen.

# 3 Nachhaltigkeit als Verantwortung

Nachhaltigkeitsansprüche gewinnen dann an Bedeutung, wenn eine Verantwortung besteht, sie zu erfüllen (Shue 1988). Wenn alle heute und zukünftig lebenden Menschen einen Nachhaltigkeitsanspruch auf die Erfüllung ihrer Grundbedürfnisse haben, so muss für aktuell lebende Menschen die Verantwortung bestehen, die Erfüllung der Grundbedürfnisse sicher zu stellen. Das bedeutet nicht, dass jeder einzelne Mensch in gleichem Maße für die Erfüllung der Nachhaltigkeitsansprüche verantwortlich ist. Die bestehende Verantwortung, wie auch immer sie verteilt ist, muss lediglich sicher stellen können, dass die Nachhaltigkeitsansprüche

---

[2] Weiter reichende Nachhaltigkeitstheorien schließen Ansprüche der Natur mit ein, geben der Natur einen intrinsischen Wert (Baumgärtner und Quaas 2010). Danach können z.B. Tiere einen Anspruch haben, nicht zu leiden. Nichtsdestotrotz kommt der dauerhaften Erhaltung der natürlichen Lebensgrundlage eine besondere Rolle zu, um legitime Ansprüche erfüllen zu können (Ewringman et al. 2012).

[3] Im Gegensatz dazu stehen komparative Standards, die Ansprüche über Vergleiche zwischen Menschen oder Menschengruppen festlegen.

erfüllt werden. Woraus sich eine solche Verantwortung genau zusammensetzt, wird im Folgenden behandelt.

Zunächst gehe ich hierfür auf drei Aspekte des Begriffs der Verantwortung ein. Verantwortung wird erstens rein deskriptiv als *Zurechnung* von Handlungen und Handlungsfolgen verwendet. Person *A* ist in diesem Sinne für *X* verantwortlich, wenn *A* der Täter oder Verursacher von *X* ist. Für Verantwortung als Zurechnung geht es also nur um die Frage, ob einer Person eine Handlung oder Handlungsfolge zuzurechnen ist (Petersen und Schiller 2011). Zweitens wird Verantwortung häufig im Sinne einer *Verpflichtung* verwendet (Baumgärtner et al. 2006). Als Verpflichtung bedeutet Verantwortung, dass eine Person *A* gewisse Handlungen ausführen oder unterlassen *sollte*, um ein erwünschtes Ergebnis *X* zu erzielen. In diesem Sinn erhält Verantwortung normativen Charakter. Als dritte Bedeutung der Verantwortung definiert Williams (2008) die Bereitschaft, auf eine *Vielzahl an Verpflichtungen* einzugehen. Verantwortung trägt in diesem Sinne also, wer sich einer Vielzahl an Verpflichtung ausgesetzt sieht. Person *A* trägt die Verantwortung, verschiedene Handlungen auszuführen oder zu unterlassen, um etwa die Ergebnisse *X*, *Y* und *Z* zu erzielen.

Eine wichtige Eigenschaft von Verantwortung im Sinne einer Vielzahl an Verpflichtungen ist, dass Verantwortung begrenzt ist. Dafür gibt es zwei Gründe. Der erste Grund sind die legitimen Ansprüche des verantwortlichen Menschen. Da jeder Mensch aufgrund der Nachhaltigkeit legitime Ansprüche hat, und da die Handlungsmöglichkeiten und Ressourcen jedes Menschen begrenzt sind, muss eine Verantwortung jedes Menschen begrenzt sein, weil ein Mensch zu einem gewissen Punkt Ressourcen für sich selbst verwenden muss, um die eigenen Bedürfnisse zu befriedigen (Shue 1988: 690). Den zweiten Grund liefert das meta-ethische „Sollen impliziert Können" Prinzip (Huber 2008). Es besagt, dass Person *A*, die etwas tun soll, dazu auch prinzipiell in der Lage sein muss, da das „sollen" ansonsten sinnlos wäre (Griffin 1992). Prinzipiell in der Lage zu sein, bedeutet dabei, dass *A* „ […] zur Ausführung oder Unterlassung körperlich, geistig, psychisch, zeitlich, örtlich fähig sein und auch die Gelegenheit dazu haben [muss]" (Huber 2008). Zu der geistigen Fähigkeit der Ausführung oder Unterlassung gehört dabei insbesondere das Wissen über die Konsequenzen der eigenen Handlungen. Wenn *A* nicht wissen kann, dass eine Handlung ihre Verpflichtung etwas zu unterlassen verletzt, so ist *A* dafür nicht moralisch verantwortlich. Diese Begrenzung spielt im Falle der Verantwortung von Konsumenten eine wichtige Rolle.

Nachdem nun die Bedeutung des Begriffs der Verantwortung, sowie die Grenzen jeder Verantwortung geklärt sind, stellt sich die Frage, welche Verpflichtungen zur Verantwortung für Nachhaltigkeit gehören. Dabei ist entscheidend, welche Verpflichtungen bestehen müssen, damit die Erfüllung aller Nachhaltigkeitsansprüche gewährleistet ist. Grundsätzlich müssen für diese Gewährleistung die folgenden Verpflichtungen bestehen (Shue 1996: 52):

I: Verpflichtung die Verletzung der Nachhaltigkeitsansprüche zu *vermeiden*,

II: Verpflichtung vor der Verletzung der Nachhaltigkeitsansprüche zu *schützen*,

III: Verpflichtung, denjenigen, deren Nachhaltigkeitsansprüche verletzt wurden, zu *helfen*.

*Verpflichtung zu vermeiden*

Die Verpflichtung zu *vermeiden* besteht darin, alle Handlungen, die die Nachhaltigkeitsansprüche anderer Menschen verletzen würden, ohne dabei der Befriedigung von eigenen Nachhaltigkeitsansprüchen zu dienen, zu unterlassen. Diese Verpflichtung ist negativer Art, da sie eine Unterlassung fordert und keine aktive Handlung. Sie ist zudem universell und gilt für alle Menschen, sofern ihre Ausübung möglich ist („Sollen impliziert können“).

Um die Erfüllung von Nachhaltigkeitsansprüchen garantieren zu können, sind zwei weitere Verpflichtungen notwendig. Denn erstens ist nicht davon auszugehen, dass jemals alle Menschen von sich aus die Verpflichtung zu *vermeiden* erfüllen. Zweitens besteht jederzeit die Möglichkeit, dass die Nachhaltigkeitsansprüche mancher Menschen z.B. durch Naturkatastrophen verletzt werden, ohne dass auch nur ein Mensch seiner Verpflichtung zu *vermeiden* nicht nachgekommen wäre.

Da nicht davon auszugehen ist, dass jemals alle Menschen der Verpflichtung zu *schützen* nachkommen, muss also die Verpflichtung bestehen, Nachhaltigkeitsansprüche zu *schützen*. Die Verpflichtung zu *schützen* wird erfüllt, in dem entweder die Verpflichtung zu *vermeiden* erzwungen wird, oder in dem Institutionen so gestaltet werden, dass es keine starken Anreize gibt, die Verpflichtung zu *vermeiden* zu verletzen (Shue 1996: 60). Die Einhaltung der Verpflichtung zu *vermeiden* zu erzwingen, wird aus praktischen Gründen meist von Institutionen durch den Erlass von Gesetzen wahrgenommen. So ist beispielsweise die Polizei gesetzlich ermächtigt, den Anspruch der Menschen auf körperliche Unversehrtheit durchzusetzen und zu erzwingen. Auch im Wirtschaftsprozess können Regierungsorgane durch Gesetze erzwingen, dass durch die Produktion und den Gebrauch von Gütern Nachhaltigkeitsansprüche nicht verletzt werden.

In vielen Fällen ist es aber weder wünschenswert noch möglich, jede Verletzung der Verpflichtung zu *vermeiden* gesetzlich zu verbieten bzw. zu erzwingen. Nicht jede nicht nachhaltige Produktionsmethode und nicht jedes nicht nachhaltige Gut kann und sollte sofort verboten werden. Jedoch können Institutionen so gestaltet werden, dass sie es gewöhnlichen Menschen, also weder Heiligen noch Genies, ermöglichen, sich gegenseitig nur ein Minimum an ernsthaftem Schaden zuzufügen (Shue 1996: 60). Beispielsweise können durch das Steuersystem oder durch Gesetze, wie dem deutschen Erneuerbaren Energien Gesetz, Anreize so gesetzt werden, dass es Menschen leicht gemacht wird, ihre Verpflichtung zu *vermeiden* zu erfüllen. So besteht die Aufgabe von Regierungen nicht nur darin, die Interessen des Landes wahrzunehmen, sondern auch die Menschen dabei zu unterstützen, ihren Verpflichtungen nachzukommen (Nihlén Fahlquist 2009).

*Verpflichtung zu schützen*

Die Verpflichtung zu *schützen* ist besonders dann wichtig, wenn die Ursachen für die Verletzung von Nachhaltigkeitsansprüchen im Zusammenwirken aus individuellen und institutionellen Handlungen, wie im Falle des Klimawandels, zu suchen sind. Da in solchen Fällen kein einzelner Akteur das Problem verursacht, verschwimmt die Grenze aus den

Verpflichtungen zu *vermeiden* und zu *schützen*, so dass die Gestaltung von Institutionen zum Schutz vor der Verletzung von Nachhaltigkeitsansprüchen besonders wichtig ist (Shue 1996: 59).

Für einzelne Menschen ist die Verpflichtung zu *schützen* eine positive und indirekte Verpflichtung. Denn im Allgemeinen besteht sie darin, Institutionen zu schaffen bzw. zu gestalten. Es gibt mindestens zwei Gründe dafür, die Verpflichtung zu *schützen* mit Hilfe von Institutionen zu erfüllen. Erstens ist es meist effizienter. Institutionen können die Einhaltung der Verpflichtung zu vermeiden mit wesentlich geringerem Aufwand erzwingen oder erleichtern, als dies durch unkoordiniertes individuelles Handeln möglich wäre (Nihlén Fahlquist 2009). Zweitens würde die Verpflichtung zu *schützen*, wenn sie nicht durch Institutionen ausgeübt würde, einzelne Menschen überfordern (Shue 1988: 697). Der Sinn dieser Verpflichtung liegt ja nicht darin, den Verantwortlichen eine Last aufzuerlegen, sondern die Erfüllung legitimer Ansprüche zu garantieren. Da Institutionen dies im Fall der Verpflichtung zu *schützen* meist besser können, sollte diese Verpflichtung auch durch Institutionen ausgeübt werden. Insofern ist die Verpflichtung zu *schützen* für einzelne Menschen indirekt.

Wie viel ein Mensch zur Gestaltung und zur Schaffung von Institutionen beizutragen hat, ist eine schwierige Frage. Einerseits gilt die Verpflichtung zu *schützen* für alle Menschen, die etwas beitragen können. Andererseits können nicht alle Menschen gleichermaßen und in gleicher Weise an der Gestaltung und der Schaffung von Institutionen mitwirken. Richardson (1999) schlägt daher vor, dass Menschen, die die Verpflichtung zu *schützen* besonders gut ausfüllen können, eine besondere Verantwortung tragen, dies zu tun. Wie weit diese Verantwortung allerdings reicht, ist in hier nicht endgültig zu klären.

*Verpflichtung zu helfen*

Die Verpflichtung zu *helfen* gewinnt an Bedeutung, wenn die Verpflichtungen zu *vermeiden* und zu *schützen* nicht oder unzureichend erfüllt werden, und wenn Phänomene wie Naturkatastrophen häufig auftreten. Die Verpflichtung zu *helfen* bedeutet, dass Ressourcen jedweder Art an Personen transferiert werden, deren Nachhaltigkeitsansprüche nicht erfüllt werden. Insofern ist diese Verpflichtung ist positiv. Sie kann entweder bestehen, wenn Hilfe notwendig wird, weil die Verpflichtungen zu *vermeiden* und zu *schützen* nicht ausreichend erfüllt wurden oder wenn Hilfe aufgrund natürlicher Umstände, wie z.B. Naturkatastrophen, benötigt wird. In den meisten Fällen ist auch die Verpflichtung zu *helfen* eine indirekte Verpflichtung, die darin besteht, Institutionen wie NGOs, Stiftungen oder Regierungsorgane zu unterstützen.

Zusammenfassend lässt sich sagen, dass um die Erfüllung von Nachhaltigkeitsansprüchen zu garantieren, die Verantwortung für Nachhaltigkeit aus den drei genannten Verpflichtungen bestehen muss. Sich auf die Verpflichtung zu *vermeiden* zu verlassen, wäre naiv, da kaum alle Menschen sie je gänzlich erfüllen werden. Daher ist die Verpflichtung zu *schützen* notwendig. Sich allerdings gänzlich auf die Verpflichtung zu *schützen* zu verlassen, würde eine starke Abhängigkeit von staatlicher Regulierung bedeuten, was kaum wünschenswert erscheint.

Daher müssen beide Verpflichtungen bestehen. Da es auch aufgrund von Naturkatastrophen zur Verletzung von Nachhaltigkeitsansprüchen kommen kann, ist zudem die Verpflichtung zu *helfen* notwendig, um die Erfüllung von Nachhaltigkeitsansprüchen zu garantieren.

# 4 Verantwortung für Nachhaltigkeit von Konsumenten

Nachhaltigkeit als Gerechtigkeit ernst zu nehmen, bedeutet also, dass eine Vielzahl an Verpflichtungen übernommen werden müssen. Aus welchen dieser Verpflichtungen die Verantwortung für Nachhaltigkeit von Konsumenten bezüglich ihrer Konsumhandlungen besteht, wird in diesem Kapitel behandelt.

Als Konsumenten lassen sich im weitesten Sinne alle Menschen bezeichnen, „die am Wirtschaftsprozess teilnehmen und um Rahmen dieses Prozesses nach der Befriedigung ihrer Bedürfnisse suchen“ (Heidbrink und Schmidt 2011: 35). Die vornehmliche Handlung von Konsumenten ist damit der Konsum, also der Kauf und Gebrauch von Gütern. Welche der drei Verpflichtungen (*vermeiden*, *schützen*, *helfen*) sind nun Teil der Verantwortung von Konsumenten bezüglich ihres Konsums?

Die Verpflichtung zu *vermeiden* gilt für alle Menschen und damit auch für Konsumenten und deren Konsumhandlungen, sofern diese nicht der Befriedigung eigener Nachhaltigkeitsansprüche dienen und sofern ausreichend Handlungsmöglichkeiten und Wissen vorhanden sind, der Verpflichtung nachzukommen. Meyer et al. (2010) bestätigen zudem, dass die Wahrnehmung dieser Verpflichtung einen erheblichen Beitrag zur Nachhaltigkeit leisten könnte. Sie schätzen, dass bei einer moderaten Verhaltensänderung in den Feldern Ernährung, Bauen und Wohnen und Mobilität bis zum Jahr 2020 bei gegebener Mobilität und Technologie bis zu 8,7% der $CO_2$ Emissionen im Vergleich zum Business as usual eingespart werden könnten (Meyer et al. 2010). Trotzdem ist es wichtig, die Grenzen dieser Verpflichtung – Handlungsmöglichkeiten und Wissen – genauer zu beleuchten.

Die Handlungsmöglichkeiten von Konsumenten, der Verpflichtung zu *vermeiden* nachzukommen, unterliegen zwar Einschränkungen, doch es bestehen zahlreiche Handlungsspielräume, die sie ausschöpfen können. Diese Handlungsspielräume sind einer starken gesellschaftlichen Dynamik unterworfen und können sich je nach kulturellen Gegebenheiten, rechtlicher Rahmenbedingungen und individuellen Kapazitäten unterscheiden (Heidbrink und Schmidt 2011: 42). Daher gilt die Verpflichtung zu *vermeiden* zwar grundsätzlich für alle Konsumenten, aber in unterschiedlichem Maße. So bestehen beispielsweise für Konsumenten, die in größeren Städten leben, ganz andere Handlungsspielräume als für Konsumenten, die in ländlichen Regionen wohnen. Stadtbewohner können zum Beispiel auf öffentliche Verkehrsmittel zurückgreifen, während Landbewohner häufiger auf das eigene Auto angewiesen sind.

Auch das Wissen von Konsumenten zu den Folgen ihres Konsums ist heterogen verteilt und es ist nicht möglich, alle Folgen des eigenen Konsums zu kennen (Petersen und Schiller 2011). Doch das bedeutet nicht, dass Konsumenten keine Verpflichtung zu *vermeiden* tragen. Denn erstens sind zumindest viele Folgen der eigenen Konsumhandlungen bekannt und zweitens können und sollten Konsumenten sich bemühen, Wissen über die Folgen der eigenen Konsumhandlungen zu erlangen (Baumgärtner et al. 2006). Dafür stehen etwa Informationsmaterialien von Verbraucherzentralen und Nichtregierungsorganisationen, Warentests, Labels und Kennzeichnungen zur Verfügung (Heidbrink und Schmidt 2011: 42).

Daher besteht für Konsumenten bei ihren Konsumhandlungen, die Verpflichtung zu *vermeiden*. Je nach Handlungsmöglichkeiten und Wissen, gilt diese Verpflichtung für manche Konsumenten stärker als für andere. Konsumenten, die dieser Verpflichtung nicht oder kaum nachkommen können, sind dafür in besonderem Maße dazu verpflichtet, Einfluss auf Institutionen zu nehmen, die ihnen die Erfüllung der Verpflichtung zu *vermeiden* ermöglichen. So können beispielsweise Konsumenten aus ländlichen Gegenden öffentliche Verkehrsmittel kaum nutzen und Schadstoffemissionen, die zur Verletzung von Nachhaltigkeitsansprüchen führen, nicht vermeiden. Sie können sich aber für den Ausbau des öffentlichen Nahverkehrs einsetzen.

Ein solcher Einsatz ist bereits Teil der Verpflichtung zu *schützen*. Die Erfüllung dieser Verpflichtung wird in der Debatte meist unter dem Begriff der politischen Verantwortung geführt (zum Beispiel in Petersen und Schiller 2011). Konsumenten können diese Verpflichtung auf verschiedene Weise erfüllen – indem sie im Rahmen ihres persönlichen Konsumhandelns bleiben, in dem sie sich in zivilgesellschaftlichen Organisation engagieren, an Demonstrationen teilnehmen, sich an Unterschriftenaktionen- oder Bundestagspetitionen beteiligen, oder das Gespräch mit Händlern oder Herstellern suchen. (Heidbrink und Schmidt 2011: 41). Individuelle Konsumhandlungen sind also nur eine Möglichkeit, die Verpflichtung zu erfüllen. Insofern weisen Petersen und Schiller (2011) zu Recht auf die Konsumentensouveränität hin und es den Konsumenten selbst überlassen, wie sie die Verpflichtung zu *schützen* erfüllen mögen. Sie sind nicht dazu verpflichtet, dies über ihre Konsumhandlungen zu tun. Gleichzeitig betonen aber auch Billharz et al. (2011) zurecht, dass auch durch individuelle Konsumhandlungen, die Verpflichtung zu *schützen* erfüllt werden kann. Denn auch hier sind die Möglichkeiten der Menschen unterschiedlich. Manchen Menschen ist es ein leichtes, sich politisch zu engagieren, während andere Menschen eher die Möglichkeit haben, ihre Konsumhandlungen zu verändern.

Die Verpflichtung zu *helfen* besteht für Konsumenten, wenn Nachhaltigkeitsansprüche durch ihren Konsum oder das marktwirtschaftliche System verletzt wurden (Heidbrink und Schmidt 2011: 36). So sind Konsumenten auch dann dazu verpflichtet zu helfen, wenn Verletzungen von Nachhaltigkeitsansprüchen nicht direkt von ihnen aber durch das marktwirtschaftliche System verursacht wurden (Heidbrink und Schmidt 2011: 47). Allerdings gelten für die Verpflichtung zu *helfen* dieselben Überlegungen wie für die Verpflichtung zu *schützen*: Konsumenten können die Verpflichtung zu *helfen* durch ihre Konsumhandlungen erfüllen (zum Beispiel durch den Kauf von Fair Trade Produkten), können aber auch direkt an Hilfsorganisationen spenden. Insofern sind Konsumenten nicht dazu verpflichtet, ihre Verpflichtung zu *helfen* über ihre Konsumhandlungen zu erfüllen.

Zusammenfassend lässt sich also feststellen, dass die Verantwortung für Nachhaltigkeit von Konsumenten aus den drei Verpflichtungen zu *vermeiden*, zu *schützen* und zu *helfen* besteht. Bezüglich ihrer Konsumhandlungen sind Konsumenten dafür verantwortlich, Verletzungen von Nachhaltigkeitsansprüchen zu *vermeiden*. Die Verpflichtungen zu *schützen* und zu *helfen* können Konsumenten durch ihre Konsumhandlungen erfüllen, aber auch durch andere Handlungen wie politischem Engagement. Um Nachhaltigkeitsansprüche garantieren zu können, ist es erforderlich, dass Konsumenten alle drei Verpflichtungen wahr nehmen (Shue 1996: 54). Denn es scheint einerseits unwahrscheinlich, dass Konsumenten alle nachhaltigkeitsschädigenden Konsumhandlungen *vermeiden* können, da ihnen dazu das notwendige Wissen fehlt. Andererseits können auch politische Institutionen nicht alle Nachhaltigkeitsansprüche schützen, da immer wieder neue Produkte und Produktionsmethoden entstehen, auf die politische Institutionen erst reagieren müssen. Ob nun die Verpflichtung zu *vermeiden*, oder die Verpflichtungen zu *schützen* oder zu *helfen* stärker betont werden sollten, bleibt eine empirische Frage, die sich in unterschiedlichen sozialen und kulturellen Kontexten jeweils anders stellt und daher anders beantwortet werden muss.

# 5 Fazit

Die Überlegungen in diesem Artikel bestätigen den Konsens in der in GAIA geführten Debatte, dass Konsumenten sich politisch für Nachhaltigkeit einsetzen sollten. Sie zeigen aber auch, dass man der Verantwortung für Nachhaltigkeit nicht gerecht wird, in dem man sie zu einem rein politischen Ziel erklärt und damit eine Verantwortung bezüglich privater Konsumhandlungen ablehnt. Vielmehr bedarf die Klärung einer Verantwortung für Nachhaltigkeit zunächst einer Spezifizierung des Konzepts der Nachhaltigkeit. Denn das Ziel jeder Verantwortung für Nachhaltigkeit ist es, die Erfüllung der legitimen Ansprüche, die sich aus dem Konzept der Nachhaltigkeit ableiten, zu gewährleisten.

Um dieses Ziel zu erreichen, sollte eine Verantwortung für Nachhaltigkeit aus den Verpflichtungen zu *vermeiden*, zu *schützen* und zu *helfen* bestehen. Konsumenten tragen für ihre Konsumhandlungen eine Verantwortung für Nachhaltigkeit bestehend aus der Verpflichtung zu *vermeiden*. Doch über ihre Konsumhandlungen hinaus beinhaltet ihre Verantwortung für Nachhaltigkeit mehr, nämlich die Verpflichtungen zu *schützen* und zu *helfen*. Insofern betonen Grunwald (2010) und Petersen und Schiller (2011) zu Recht die Bedeutung der politischen Verantwortung für Nachhaltigkeit. Es wäre allerdings ein Fehler, die private Verantwortung bezüglich der Konsumhandlungen zu negieren oder zu vernachlässigen. Welche Verpflichtungen innerhalb der Verantwortung für Nachhaltigkeit von Konsumenten nun gerade im öffentlichen Diskurs stärker zu betonen ist, bleibt eine deskriptive Frage. Aus normativer Sicht muss in jedem Fall die Notwendigkeit aller drei Verpflichtungen betont werden.

# 6 Literatur

Baumgärtner, S., Faber, M., und Schiller. J. (2006). *Joint production and responsibility in Ecological Economics. On the foundations of environmental policy*. Cheltenham, Edward Elgar.

Baumgärtner, S., und Quaas, M. (2010). What is sustainability economics? *Ecological Economics* 69: 445–450.

Bilharz, M., Fricke, V., und Schrader, U. (2011). Wider die Bagatellisierung der Konsumentenverantwortung. *GAIA* 20(1): 9–13.

Ewringmann, D., Faber, M., Petersen, T., und Zahrnt, A. (2012). Schluss mit der Harmonie! Warum das Nachhaltigkeitskonzept zur Umweltgefahr geworden ist. *Die Zeit*, 30.01.2012.

Gauck, J. (2012). *Rede nach der Vereidigung zum Bundespräsidenten*. Retrieved from: http://www.bundespraesident.de (April 6, 2012). Bundespraesidialamt, Berlin.

Gosselin, A. (2006). Global poverty and responsibility: identifying the duty-bearers of human rights. *Human Rights Review* 8(1): 35–52.

Griffin, J. (1992). The human good and the ambitions of consequentialism. *Social Philosophy and Policy* 9(2): 118–132.

Grunwald, A. (2010). Wider die Privatisierung der Nachhaltigkeit. *GAIA* 19(3): 178–182.

Grunwald, A. (2011). Statt Privatisierung: Politisierung der Nachhaltigkeit. *GAIA* 20(1): 17–19.

Hampicke, U. (2005). Gedanken zu minimaler Gerechtigkeit. In: *Das Ethische in der Ökonomie*. Bschorner, T., Eger, T. (Eds.). Marburg, Metropolis Verlag. 475–490.

Hauff, V. (Ed.). (1987). *Unsere gemeinsame Zukunft. Der Brundtland-Bericht der Weltkommission für Umwelt und Entwicklung*. Greven, Eggenkamp.

Heidbrink, L., und Schmidt, I. (2011). Das Prinzip der Konsumentenverantwortung – Grundlagen, Bedingungen und Umsetzungen verantwortlichen Konsums. In: *Die Verantwortung des Konsumenten*. Heidbrink, L., Schmidt, I., Ahaus, B. (Eds.). Frankfurt am Main, Campus Verlag.

Huber, H. (2008). *Sollen impliziert Können*. Retrieved from: http://www.gavagai.de/themen/HHPT22.htm (April 6, 2012).

[ILO] International Labour Organization. (1976). *Employment, Growth and Basic Needs: A One-World Problem*. Report of the Director-General of the International Labour Office. Geneva.

König, W. (2008). *Kleine Geschichte der Konsumgesellschaft. Konsum als Lebensform der Moderne*. Stuttgart, Franz Steiner.

Langrock-Kögel, C. (2012). Wir Teilzeit-Helden. *Enorm Magazin* 1(2012): 16–26.

Meyer, B., Meyer, M., und Meyer zu Holte, I. (2010). *Die ökologischen und ökonomischen Wirkungen eines nachhaltigeren Konsums in Deutschland*. Studienreihe Mut zur Nachhaltigkeit. Nr. 1. Retrieved from: http://www.mut-zur-nachhaltigkeit.de/images/dokumente/Infopool/studienreihe%20mzn%201_nachhaltiger%20konsumieren.pdf (April 3, 2012)

Nihlén Fahlquist, J. (2009). Moral responsibility for environmental problems—individual or institutional? *Journal of Agricultural and Environmental Ethics* 22(2): 109–124.

Nussbaum, M. (2007). *Frontiers of Justice*. Cambridge/Mass., Harvard University Press.

Ott, K., und Döring, R. (2008). *Theorie und Praxis starker Nachhaltigkeit*. 2. Edition. Marburg, Metropolis Verlag.

Petersen, T. (2009). Nachhaltigkeit und Verteilungsgerechtigkeit. In: *Die Greifswalder Theorie starker Nachhaltigkeit*. Egan-Krieger, T., Schultz, J., Pratap Thapa, P., Voget, L. (Eds.). Marburg, Metropolis Verlag. 81–96.

Petersen, T. und Schiller, J. (2011). Politische Verantwortung für Nachhaltigkeit und Konsumentensouveränität. *GAIA* 20(3): 157–161.

Richardson, H. (1999). Institutionally divided moral responsibility. In: *Responsibility*. Paul, E., Miller, F., Paul, J. (Eds.). Cambridge, Cambridge University Press. 218–249.

Sen, A. (1986). The standard of living. In: *The Tanner lectures on human values*. McMurrin, M. (Ed.). Salt Lake City, University of Utah Press. 1–51.

Shue, H. (1988). Mediating Duties. *Ethics* 98: 687–704.

Shue, H. (1996). *Basic rights: Subsistence, affluence and U.S. foreign policy*. 2. Edition. Princeton, NJ, Princeton University Press.

Siebenhüner, B. (2011). Kann die Politik es richten? Konsument(inn)en als politische Akteure. *GAIA* 20(1): 14–16.

von Petersdorff, W. (2011). Grüne Revolution – Die herzliche Ökodiktatur. *Frankfurter Allgemeine Zeitung*. 16.05.2011.

Williams, G. (2008). Responsibility as a virtue. *Ethical Theory and Moral Practice* 11(4): 455–470.

# Paper 3: Regulation of morally responsible agents with motivation crowding

JOACHIM FÜNFGELT* AND STEFAN BAUMGÄRTNER

Department of Sustainability Sciences and Department of Economics,
Leuphana University of Lüneburg, Germany

**Abstract:** We study the regulation of a morally responsible agent in the context of a negative consumption externality and motivation crowding. In particular, we analyze how various governmental interventions affect the agent's motivation to assume moral responsibility. Employing a motivation-crowding model, we find that morally motivated behavior will, in general, not ensure Pareto efficiency without intervention. A Pigouvian tax may be efficient under motivation crowding. But the efficient taxe rate needs to be higher, which may lead to a full crowding-out of moral motivation. By contrast, an inefficiently low taxe rate may increase the market failure due to motivation crowding. Provision of information is efficient only in very specific cases but may be effective in reducing the extent of market failure. A complementary tax-and-information policy approach is superior to a tax as single instrument if its aim is to reduce consumption and if provision of information raises moral motivation.

**JEL Classification:** D03, D11, D62, H23, Q58

**Keywords:** Altruism, externality, moral motivation, motivation crowding, Pareto efficiency, regulation, responsibility, taxes, provision of information

* Corresponding author: Sustainability Economics Group, Leuphana University of Lüneburg, P.O. Box 2440, D-21314 Lüneburg, Germany, phone: +49.4131.677-2668, fax: +49.4131.677-1381, email: fuenfgelt@uni.leuphana.de

# 1 Introduction

Many environmental problems, such as climate change or the loss of biodiversity, are driven by negative externalities. Essentially, such externalities cause market failure for which conventional economic wisdom suggests introducing governmental policies in the form of taxes or subsidies (e.g. Pigou 1932, Baumol 1972). These suggestions rely on the assumption of purely self-centered human behavior. However, this assumption is not generally justified since human beings often assume moral responsibility, that is in their actual behavior they respond to moral obligations (e.g. Sen 1977, Brekke et al. 2011, Perino et al. 2011). Furthermore, Motivation Crowding Theory (e.g. Deci 1971 or Frey 1997, 2001) suggests that extrinsic interventions, such as governmental policies, severely affect individuals' motivation to assume moral responsibility. In this paper, we study the regulation of a morally responsible individual with motivation crowding in the context of a negative consumption externality.

In the case of environmental policies, command and control instruments, but also incentive-based instruments such as tradable emission rights or Pigouvian taxes, tend to undermine moral motivation, while information, appeals and participation enhance moral motivation (Frey and Jegen 2001). Empirical evidence is plentiful,[1] but there are few theoretical studies on the issue and these do not simultaneously consider negative externalities and motivation crowding. Heyes and Kapur (2011) analyze how moral motivation, in the context of negative externalities, affects the optimal specification of particular policy instruments. Their focus, however, is on motivational heterogeneity and they do not consider the case of motivation crowding. Further literature on moral motivation has mainly focused on the voluntary provision of public goods by morally motivated individuals (e.g. Andreoni 1988, 1990, Brekke et al. 2003, Nyborg and Rege 2003). Moral motivation is generally modeled as a warm-glow, based on a utilitarian norm by which an optimal level of giving is defined. Something like motivation crowding occurs in those models when environmental policies influence the optimal level of giving.

We contribute to the literature in three ways: First, we consider the case of externalities which is more general than the case of public goods. Second, we focus on responsible behavior rather than behavior driven by a warm-glow or self-image. Third, we model motivation crowding as a psychological phenomenon (in the sense of the self-determination theory of Deci and Ryan 1985), and not as a purely economic phenomenon driven by changes in the optimal allocation. Altogether, this allows us to identify fundamental psychological determinants for the efficiency of taxes, provision of information and a policy mix of the two instruments.

More specifically, we analyze the regulation of a morally responsible individual in the context of a negative consumption externality and motivation crowding, focusing on the moral principle: 'You ought not to consciously harm others against their will'. Against this

[1] For an economic survey of the issue see Bowles and Hwang (2008) or Gneezy et al. (2011).

background, we focus on two questions: (1) Is responsibility – understood as moral motivation of individual actors – sufficient for Pareto efficiency in a decentralized economy when individual action causes negative externalities? (2) Can a Pigouvian tax, provision of perfect information, or a complementary policy combining both instruments lead to Pareto efficiency when moral motivation is subject to motivation crowding?

For this analysis, we use a simple model: there are two goods, one numeraire good and one polluting good, and two individuals, *A* and *B*. *A* derives utility from private consumption of both goods and a morally weighted disutility from her knowledge about her causing the externality. *B* derives utility from the numeraire good and disutility from *A*'s consumption of the polluting good. We thus have an asymmetric, unidirectional power structure, as *A* is responsible for the harm inflicted on *B*.[2] The moral weight in *A*'s utility function reflects the personal desirability of responsible behavior and is affected by policy measures (motivation crowding): it decreases with a tax, and increases with provision of information. The model thus allows us to study the effects of regulatory policies with respect to Pareto efficiency: price regulation through a Pigouvian tax on the polluting good, descriptive information provisioning as lowering uncertainty about the externality, and a complementary policy combining both instruments.

Our results show that morally responsible behavior will in general not lead to Pareto efficiency without governmental intervention as it may diminish or exacerbate market failure. Intervention through taxation leads to crowding-out of moral motivation, but there always exists a tax rate so that the equilibrium allocation is efficient. However, such a tax-only policy has three weaknesses due to motivation crowding: First, crowding requires a higher tax rate which may be difficult to implement due to political pressure. Second, setting the tax rate inefficiently low may exacerbate the market failure. And third, an efficient tax rate may fully crowd-out moral motivation if there are motivational spill-overs. Intervention through provision of information is only efficient for very restrictive assumptions, but can be effective in reducing the market failure. Intervention through a complementary tax and information approach is an efficient instrument just as a tax-only policy, and may overcome the weaknesses of a tax-only policy for some (but not for all) parameter values. Altogether, our study highlights the need for the development of new policy instruments in the face of externalities and motivational crowding.

The paper is organized as follows. Section 2 prepares the conceptual basis for the analysis. Section 3 introduces the model. Section 4 presents our results. Section 5 concludes.

---

[2] Such structures are particularly important when future generations (*B*) are affected by the behavior of the present generation (*A*).

# 2 Conceptual foundations: moral responsibility and motivation crowding

In this section, we prepare the conceptual basis of the paper by first defining the concepts of moral responsibility and motivation crowding. Second, we link both concepts.

Responsibility is a multifarious notion. In the philosophical discussion of responsibility, at least three different aspects of the notion have been distinguished. (1) The primary meaning of responsibility is being the perpetrator of one's own actions, that is, "[…] one ascribes an action to oneself and allows for it to be thus ascribed" (Baumgärtner et al. 2006: 227). The primary meaning is purely descriptive and has no moral relevance by itself. It simply states that *A* is responsible for *X* if and only if *A* is the perpetrator of *X*. This is a precondition of morality, as one can only be morally praised or blamed for an action that can be ascribed to oneself. (2) When we speak of 'responsibility', we often use 'responsibility' as a synonym for 'obligation' (Williams 2008: 458). This is what Baumgärtner et al. (2006) call the secondary meaning of responsibility. In this meaning, responsibility attains a moral significance when obligations exist which a person morally has to accept, that is, *A* ought to do *X* or ought not to do *X* for moral reasons. (3) Williams (2008) defines a third meaning of responsibility: "Responsibility represents the readiness to respond to a plurality of normative demands" (Williams 2008: 459). In other words, responsibility is important whenever individuals are facing a plurality of normative obligations[3]. One specific suggestion as to how to ethically balance two rivaling normative obligations is due to the utilitarianist Peter Singer (1972). He suggests that two obligations ought to be balanced to the point of marginal utility at which both obligations are equally met at the margin.

In line with the above reflection, we consider the responsibility of an agent for (the consequences of) her actions [aspect 1], as facing a moral obligation [aspect 2] while also striving [aspect 3] for personal happiness. An individual assuming responsibility for her actions is self-negotiating two aims: the obligation to herself to have a good life, and the moral obligation not to harm others against their will. This act of assuming responsibility requires that an individual is motivated to act responsibly.

To be motivated means to be moved to do something. An individual who feels no inspiration to act is characterized as unmotivated, whereas an individual who is activated toward an end is considered motivated (Ryan and Deci 2000: 54). Individuals may have different levels, but also different kinds of motivation. The psychological literature distinguishes between intrinsic and extrinsic motivation. One is said to be *intrinsically* motivated to perform an activity when one receives no apparent rewards except the activity itself (Deci 1971). Kunda and Schwartz (1983) consider the will to fulfill a moral obligation and to assume responsibility as a special type of intrinsic motivation. Such intrinsic motivation might be either innate or learned (White 1959), and may thus change. *Extrinsic* motivation comes from outside the individual. All forms of monetary reward or threat (e.g.

[3] With "obligation" we here refer to what Williams (2008) describes as "normative demand".

taxes, subsidies, fines) are examples of extrinsic motivation. Such extrinsic rewards or threats can lead to overjustification and a subsequent reduction of intrinsic motivation (Kunda and Schwartz 1983). For example, Titmuss (1971) finds that paying individuals for donating blood might decrease the willingness to donate blood.[4] The reason simply is that individuals wish to donate blood because they are intrinsically motivated to do so. If they are offered a monetary reward, this intrinsic motivation is replaced, or crowded-out, by the extrinsic motivation to receive money. If the intrinsic motivation was stronger than the subsequent extrinsic motivation, the willingness to donate blood decreases.

In the late 1990's, the work of Frey (1997, 2001) put motivation crowding on the research agenda of economics. By now, there exists plenty of empirical evidence for economic instruments crowding-out[5] intrinsic motivation in the economic[6] and in the psychological[7] literature. Still open is the question of how the extent of motivation crowding depends on the quantity of monetary compensation or taxes. Frey and Oberholzer-Gee (1997) find that individuals' willingness to accept a nuclear waste facility in their neighborhood does not increase with monetary compensation levels. This suggests that the crowding effect of monetary compensation increases with the compensation offered. In contrast, Gneezy and Rustichini (2000b) find that higher compensations for previously unpaid tasks increase effort levels, which suggests that higher compensation levels do not have stronger crowding effects. Therefore, we leave the relationship between the quantity of the extrinsic intervention and the extent of the crowding effects open. Shedding more light on this relation remains an interesting task for empirical research.

To summarize, individuals want to assume moral responsibility and their intrinsic motivation is the key to understand how they react to governmental policies. Yet, this intrinsic motivation is prone to crowding – both positive and negative – from regulatory intervention. This is the starting point for the analysis in this paper. In the following section, we set up a model of motivation crowing which allows us to study the relationship between people's intrinsic motivation and different policy instruments.

[4] Mellström and Johannesson (2008) recently confirmed the Titmuss result, but only for women. For men they did not find crowding effects.

[5] Charness and Gneezy (2009) is one of the few studies finding crowding-in through monetary incentives. However, they do not analyze morally motivated behavior, but the motivation to exercise in a gym.

[6] See e.g. Frey (1999, 2001), Nyborg (2003b), Goeschl and Perino (2012) and Perino et al. (2011), d'Adda (2011), Ariely et al. (2009), or Falk and Kosfeld (2006). See Gneezy et al. (2011) for a recent overview of the literature.

[7] See e.g. Deci et al. (1999), or Heyman and Ariely (2004).

# 3 Model

There are two individuals, *A* and *B*, and two goods, *X* and *Y*, where *Y* is a numeraire good that is consumed by both individuals. Let $y^j \geq 0$ denote the consumption of *Y* by individual *j* $(j = A,B)$. In contrast, *X* is only consumed by individual *A*. *A*'s consumption of *X*, denoted by $x \geq 0$, causes a negative externality on *B*'s utility, $d(x)$, with $d(0) = 0$, $d'(x) > 0$, and $d''(x) \geq 0$ for all $x \geq 0$.

Government may intervene to regulate the externality through either one, or both, of the following two policy instruments: (1) a Pigouvian tax with tax rate *t* on the polluting good *X*, where *t* may be greater or smaller than zero, i.e. it may be a tax or a subsidy[8]; (2) provision of perfect information *i* on the actual extent of damage *d*(*x*).

In this unidirectional power structure, in which *A*'s behavior has consequences for *B*'s well-being, we assume that *A* is morally motivated to act responsibly. In the utility function that determines her actual behavior, she is thus self-negotiating two obligations: the moral obligation not to harm *B*, and the obligation to maximize her self-directed well-being:[9]

$$U^A\left(x,y^A\right) = y^A + u(x) - m(t,i)k(i)d(x). \tag{1}$$

The first part of this additively separable utility function, $y^A + u(x)$, denotes *A*'s quasi-linear self-directed utility[10] from private consumption of both goods, with $u'(x) > 0$, and $u''(x) < 0$ for all $x \geq 0$. We further assume that $\lim_{x \to 0} u'(x) \to +\infty$, which ensures that *A* always consumes a strictly positive amount of the polluting good *X*.

This self-directed utility is reduced by $m(t,i)k(i)d(x)$, which represents *A*'s moral motivation not to harm *B*. This second part depends on consumption of the polluting good *X*, and on the level of government intervention through taxes, *t*, and information, *i*. $k(i)d(x)$ denotes *A*'s expectation of the externality on *B*'s utility from her consumption of *x*. The term differs from the actual harm to *B*, $d(x)$, by a factor of $k(i)$, which measures *A*'s knowledge about the externality. For $k(i) < 1$, *A* underestimates the externality, while for $k(i) > 1$, she overestimates the externality. We assume that without any provision of information, *A* has

[8] We assume that the tax income is lump-sum redistributed by the government, and that subsidies are paid from government funds that are raised in a non-distortionary manner.

[9] We thus apply the dual preferences model of Brekke et al. (2003) and extend it for the case of externalities and for the idea of motivation crowding from the model of Frey and Oberholzer-Gee (1997).

[10] One may consider the numeraire good *Y* as a composite good, such as money left for all other goods. *Y* will thus account for the major part of *A*'s utility.

some knowledge of the externality, $k(0)=\kappa \geq 0$. The government can influence *A*'s knowledge by providing perfect information $i=i^*$, such that individual *A* is perfectly informed about the externality, i.e. $k(i^*)=1$. The government's aim is thus to fully inform individual *A*, so that *A* becomes fully aware of the harm her consumption of *X* inflicts on *B*.[11] When *A* is perfectly informed, the expected damage equals the actual damage. Altogether, it is the knowingly inflicted harm on *B* which reduces *A*'s utility.

Assuming responsibility for herself and for *B*, person *A* self-negotiates her self-directed utility and the known externality with a moral-motivation factor $m(t,i)$ [12] that expresses her intrinsic motivation to act responsibly:

$$m(t,i)=\max\{n(t,i),0\}. \tag{2}$$

*A* has a basic level $m(0,0)=\mu \geq 0$ of intrinsic motivation not to harm *B* when there is no government intervention. If the government imposes extrinsic incentives, the total intrinsic motivation $m(t,i)$ is affected. A Pigouvian tax *t* on the polluting activity *X* reduces the intrinsic motivation, $n_t<0$[13] for $t>0$ and $n_t>0$ for $t<0$.[14,15] For the provision of perfect descriptive information $i^*$, we follow Reeson and Tisdell (2008) and assume crowding-in of intrinsic motivation, i.e. $n(t,i^*)>n(t,0)$, because such a policy acknowledges the freedom

[11] In order to maximize welfare, a government could use information strategically to reach a Pareto-efficient state in the short-run (Asheim 2010). We do not consider this possibility, because information cannot be used strategically in the long-run, as Abraham Lincoln stated: "You may fool all the people some of the time; you can even fool some of the people all the time; but you can't fool all of the people all the time."

[12] Frey and Oberholzer-Gee (1997) introduced a somewhat similar model. Our motivation-crowding model for *m*(*t*,*i*) extends theirs as, firstly, they analyzed the effect of only one instrument, which secondly was not a tax or information, but monetary compensation for the willingness to accept a hazardous facility in one's neighborhood.

[13] Subscripts denote partial derivatives, in this case $n_t=\partial n(t,i)/\partial t$.

[14] Note that, given the evidence from Perino et al. (2011), we assume crowding-out effects not only for positive tax rates but also for subsidies ($t<0$).

[15] Frey and Oberholzer-Gee (1997) find evidence for linear crowding of rewards to accept nuclear facilities in one's neighborhood, i.e. ($n_t=\text{const.}$), assuming that there is a constant elasticity of income. To our knowledge, this is the only evidence for the curvature of motivation crowding.

and self-determination of individuals.[16] The government provides $i^*$ at no cost. Lacking empirical evidence, we leave it open whether there are cross effects between the two policy instruments on motivational crowding, $n_{ti} \gtreqless 0$. Furthermore, we assume that $m(t,i)$ cannot become negative. To sum up, the term $m(t,i)$ expresses that *A* is intrinsically motivated not to harm *B*. A tax *t* crowds-out this motivation while provision of information *i* crowds-in this motivation.

Individual *B*'s utility function also has two parts: the linear utility he derives from his own consumption of *Y* and the harm caused by *A*'s consumption of *X*:

$$U^B\left(y^B, x\right) = y^B - d\left(x\right). \tag{3}$$

Let us further assume that both individuals have exogenous income $I^j > 0$ ($j = A,B$). By choice of units, the market price of the numeraire good *Y* equals one, while the market price of *X* is *p*. As individual *B* only consumes the numeraire good, he maximizes his utility (Eq. (3)) spending all his income for it: $I^B = y^B$. Individual *A* maximizes her utility function (Eqs. (1) and (2)) subject to the budget constraint: $I^A = px + y^A$.

# 4 Analysis and Results

Let us start with the equilibrium conditions with and without government intervention:

**Lemma 1**

For every government intervention $(t,i) \in \mathbb{R} \times \left\{0, i^*\right\}$ there uniquely exists an equilibrium allocation of good *X*, $x^*(t,i) > 0$, which is characterized by the following first order condition:

$$u'\left(x^*\right) = p + t + m(t,i)k(i)d'\left(x^*\right). \tag{4}$$

**Proof:** See Appendix A.1

Lemma 1 shows that, in equilibrium, *A*'s marginal utility of consuming *X* equals *A*'s marginal opportunity cost of consumption, *p*+*t*, plus *A*'s marginal moral costs. The latter are the product of *A*'s moral motivation and the expected marginal damage. Both terms of this product are contingent on government intervention. First, *A*'s moral motivation $m(t,i)$ takes

[16] Nyborg (2011) argues that it can be rational for moral agents to remain ignorant or even pay for not being provided with information, as information may reduce their utility. In contrast, we focus on genuinely responsible agents, who also have a responsibility to actively seek information (see Baumgärtner et al. 2007: 240ff).

on a different value for every level of government intervention. Second, the expected marginal damage $k(i)d'(x^*)$ is contingent on the government's information policy *i*.

To assess individual behavior and government policies from a societal perspective, we employ the criterion of Pareto-efficiency. An allocation is called *Pareto-efficient* if and only if it is not feasible to improve the well-being of one person without lowering the well-being of the other person. We do not use a social welfare function to assess social optimality, but rather stay with the weaker efficiency criterion, because any welfare function implies some position on distributive justice, which we do not study here. A second reason for employing the Pareto-efficiency criterion is that our basic concept of moral obligation is that it is wrong to consciously harm others against their will, or in other words, it is wrong to benefit in terms of well-being from doing harm to, that is reducing the well-being of, others. The Pareto-efficiency criterion captures this moral obligation very well.[17] The criterion of Pareto-efficiency, as a criterion of societal choice, is thus in line with the moral responsibility that individual agents feel obliged to comply with.

There has been a discussion as to whether the moral-motivation term in person *A*'s utility function should be included in the Pareto-efficiency criterion. We follow the predominant view expressed by Hammond (1978) and Diamond (2006) who argue against including this term for a number of reasons.[18] All taken together, Diamond (2006) advocates using the moral-motivation model for positive (i.e. descriptive) purposes only, while staying with the standard model of self-directed well-being for evaluating Pareto-efficiency.[19]

**Lemma 2**

There uniquely exists a Pareto-efficient allocation of good *X*, $\hat{x} > 0$, which is characterized by the following first order condition:

$$u'(\hat{x}) = p + d'(\hat{x}). \tag{5}$$

**Proof:** See Appendix A.2.

The first-order condition for Pareto-efficiency requires that *A*'s marginal utility of consuming *X* equals *A*'s marginal opportunity cost of consumption plus the marginal costs of the consumption of *X*, that is, the marginal harm on *B*.

---

[17] This is in contrast to, for example, the utilitarian welfare criterion, according to which it may well be socially desirable to increase one person's individual well-being at the cost of reducing someone else's individual well-being.

[18] First, the analysis of moral motivation would always be incomplete and thus misleading. Second, the outcome of moral motivation would be very sensitive to the framing, since moral motivation is highly context dependent. Third, including the moral-motivation term leads to double counting of the externality which was not justified.

[19] Based on these arguments, Heyes and Kapur (2011) do not include the moral-motivation term in their welfare analysis of how to regulate altruistic agents.

An equilibrium allocation $x^*$ is Pareto-efficient if and only if it equals $\hat{x}$: $x^* = \hat{x}$. Any deviation of $x^*$ from $\hat{x}$ indicates a market failure. Large (small) deviations yield large (small) market failure in the following sense.

**Definition 1**

The *extent of the market failure* under government intervention $(t, i)$ is measured by

$$\Phi(t,i) = \left| x^*(t,i) - \hat{x} \right|. \tag{6}$$

The extent of the market failure is thus defined as the absolute deviation of the equilibrium allocation $x^*$ from the efficient allocation $\hat{x}$. This definition allows comparing the extent of the market failure induced by any two government interventions *t* and *i*.

For future reference, we define one special case. A shift from one government intervention $(t_1, i_1)$ to another one $(t_2, i_2)$ that shifts the equilibrium allocation from $x_1^*(t_1, i_1)$ to $x_2^*(t_2, i_2)$ with $x_1^* > \hat{x} > x_2^*$ and $\Phi(t_1, i_1) < \Phi(t_2, i_2)$ (or likewise the other way round) is called a *strong reversal of market failure*. Strong reversals of the market failure increase the extent of the market failure.

In light of the first-order conditions for the equilibrium and the Pareto-efficient allocations, we now study four different policy scenarios: (1) a "laissez-faire" scenario in which government does not intervene at all; (2) a "tax policy" scenario in which government levies a Pigouvian tax *t* on the consumption of good *X* that causes the negative externality, but does not provide any information *i* on the actual damage caused by the consumption of *X*; (3) an "information policy" scenario in which government provides perfect information $i^*$ about the negative externality caused by good *X*, but does not levy a tax *t*; and (4) a "complementary policy" scenario in which government levies a tax *t* on the consumption of good *X* that causes the negative externality and also provides perfect information $i^*$.

### 4.1 Laissez-faire

To start with, we consider the laissez-faire scenario without government intervention, i.e. $t = 0$ and $i = 0$.

**Proposition 1**

The laissez-faire equilibrium allocation, $x^{\text{LF}} = x^*(0,0)$, is Pareto efficient if and only if $\mu = 1/\kappa$. Thus, morally motivated behavior alone is, in general, not sufficient for Pareto efficiency in the presence of externalities, and government intervention remains necessary to achieve Pareto efficiency. However, moral motivation may diminish or exacerbate the extent of the market failure:

$$\frac{\mathrm{d}\Phi(0,0)}{\mathrm{d}\mu}\begin{Bmatrix}<\\>\end{Bmatrix}0 \text{ for } \mu\begin{Bmatrix}<\\>\end{Bmatrix}\frac{1}{\kappa}. \tag{7}$$

**Proof**: See Appendix A.3

Whenever individual *A* faces a moral obligation, she has to self-negotiate it with her desire for personal consumption. In our model, *A* self-negotiates her moral obligation not to harm *B* with her personal consumption desire, by having a certain level of moral motivation $m(t,i)$. It is, however, purely coincidental whether her level of moral motivation ($\mu$) in combination with her knowledge ($\kappa$) lead her to consume an efficient amount of *X*. Hence, for all combinations of basic moral motivation and knowledge but one ($\mu = 1/\kappa$), individual *A*'s basic moral motivation is either too low in relation to her knowledge ($\mu < 1/\kappa$) or too high ($\mu > 1/\kappa$), so that the outcome is not Pareto-efficient. In the following, we refer to the case of $\mu < 1/\kappa$ as individual *A* being "undermotivated" and to the case of $\mu > 1/\kappa$ as her being "overmotivated".

Given that moral motivation alone does not preclude the existence of a market failure, Proposition 1 also makes a statement about the effect of moral motivation on the extent of the market failure. The extent of market failure decreases with the level of basic moral motivation if moral motivation is inefficiently small, that is for $\mu < 1/\kappa$, and increases otherwise. In other words, if individual *A* is undermotivated, every increase in her basic moral motivation shifts the Laissez-faire equilibrium level $x^{\mathrm{LF}}$ closer to the efficient level $\hat{x}$ as $\mathrm{d}x^{\mathrm{LF}}/\mathrm{d}\mu < 0$. If individual *A* is overmotivated, further moral motivation shifts the equilibrium away from the efficient level. Moral motivation may thus diminish or exacerbate the market failure in the presence of externalities.

## 4.2 Tax policy

In this scenario, government introduces a Pigouvian tax *t* on good *X*, but provides no information ($i = 0$). The consumer price of *X* becomes $p + t$. Besides the relative price effect, we have motivation crowding-out, as the tax reduces *A*'s moral motivation.

**Proposition 2**

There exists at least one tax rate $\hat{t} \leq d'(\hat{x})$, so that the equilibrium allocation $x^t = x^*(t,0)$ is Pareto-efficient. All efficient tax rates are characterized by the following first order condition:

$$\hat{t} + \left(m(\hat{t},0)\kappa - 1\right)d'\left(x^t\right) = 0 . \tag{8}$$

If $m\left(d'(\hat{x}),0\right) > 0$, $\hat{t}$ is unique and positive (negative) for $\mu < (>) 1/\kappa$.

If and only if $m\left(d'(\hat{x}),0\right)=0$, Eq. (8) holds for $\hat{t}=d'(\hat{x})$. There exist two further solutions if $\mu>1/\kappa$ or if $\mu<1/\kappa$ and $m_{tt}\ll 0$.

Increasing the tax rate at inefficiently low levels may exacerbate the market failure:

$$\frac{\mathrm{d}\Phi(t,0)}{\mathrm{d}t}>0 \qquad \text{iff (1) } m_t<-1/\kappa d'(\hat{x}) \text{ and} \qquad \text{(2) } \mu<1/\kappa . \tag{9}$$

**Proof**: see Appendix A.4

The first order condition given by Eq. (8) reveals that without moral motivation ($m(t,0)\equiv 0$) we obtain the standard result: there exists a Pigouvian tax rate which must equal marginal damage $d'(\hat{x})$. As we include moral motivation in the analysis, there exists at least one efficient tax rate, which may, however, differ considerably from $d'(\hat{x})$.

Furthermore, Eq. (8) reveals that there cannot exist an efficient tax rate larger than $d'(\hat{x})$. This is intuitive, as there are only two possibilities: either a tax rate $t=d'(\hat{x})$ does not crowd-out all intrinsic motivation ($m\left(d'(\hat{x}),0\right)>0$), from which it follows that $t=d'(\hat{x})$ is inefficiently high; or a tax rate $t=d'(\hat{x})$ fully crowds-out all intrinsic motivation ( $m\left(d'(\hat{x}),0\right)=0$), which renders $t=d'(\hat{x})$ an efficient tax rate.

We keep differentiating these two cases for the further discussion. $m\left(d'(\hat{x}),0\right)>0$ implies relatively weak crowding effects as $t=d'(\hat{x})$ does not crowd-out all intrinsic motivation. In this case, we find that there uniquely exists an efficient tax rate $\hat{t}$ . Intuitively, $\hat{t}$ is positive if and only if individual *A* is undermotivated ($\mu<1/\kappa$). $\hat{t}$ is negative if and only if *A* is undermotivated ($\mu>1/\kappa$).

$m\left(d'(\hat{x}),0\right)=0$ implies that at a tax rate $t=d'(\hat{x})$ all intrinsic motivation is crowded-out and the crowding effect is relatively strong. In this case, it follows that $t=d'(\hat{x})$ is an efficient tax rate because individual *A* reacts on it as if she was not morally motivated at all. But there are more possible solutions in this case.

First, if and only if individual *A* is undermotivated ($\mu<1/\kappa$) and the crowding effect is highly concave ($m_{tt}\ll 0$), there further exist one or two positive, efficient tax rates smaller than $d'(\hat{x})$. The intuition is that highly concave crowding implies that the crowding effect is very weak for low taxes which allows for the possibility of one or two low efficient tax rates.

Second, if and only if individual *A* is overmotivated ($\mu>1/\kappa$), there also exists a negative tax rate, so that the equilibrium allocation is efficient. More surprisingly, there also exists an efficient, positive tax rate smaller than $d'(\hat{x})$. In this case the crowding effects are much stronger than the price effect of the tax, such that *A*'s consumption increases with the tax to the efficient level. Hence, strong crowding effects may allow for low tax rates instead of subsidies if *A* is overmotivated.

Motivation crowding does, in principle, not jeopardize the efficiency of Pigouvian taxes. With motivation crowding, Pigouvian taxation has some side effects, though, which deserve further attention:

First, motivation crowding may lead to a higher efficient tax rate than if individual *A* was not prone to motivation crowding ($m(t,0) = \mu$ if and only if $\mu < 1/\kappa$). This may be problematic if the government faces political pressure by, for example, industrial lobby groups which lobby for low taxes (see e.g. Aidt 1998 or Fredriksson 1997).

Second, if taxes are set inefficiently low, they may exacerbate the market failure (Eq. (9)). Standard theory suggests that even an inefficiently low Pigouvian tax is an improvement compared to no taxation. If there are crowding effects however, inefficiently low taxation may actually increase the extent of the market failure. This is a serious problem as in reality Pigouvian taxes are frequently set too low.

Third, the efficient tax rate may completely crowd out moral motivation ($m(d'(\hat{x}),0) = 0$). This is a problem if there are motivational spill-over effects such that the crowding-out effect spreads to unregulated areas of behavior.[20]

We thus conclude at this point that despite the efficiency of taxes, it remains necessary to investigate alternative policy instruments which are superior to taxes or complement them such that the described side effects are mitigated.

### 4.3 Information policy

We analyze the effect of the provision of perfect information *i* as an alternative policy instrument. We now assume that rather than levying a tax, the government provides perfect descriptive information $i^*$, such that $k(i^*) = 1$ and individual *A* is perfectly informed of the externality. The aim of the government is thus to enable *A* to consume responsibly based on all available information. We now examine whether this policy can be Pareto-efficient.

**Proposition 3**

The equilibrium allocation under provision of perfect information, $x^i = x^*(0, i^*)$, is Pareto-efficient if and only if

[20] Such motivational spill-over effects are described in Frey (1999). He states that when intrinsic motivations are linked across areas, an instrument may work efficiently in the area where it is applied but at the same time reduce the positive effect of moral motivation at other areas of behavior.

$$m\left(0,i^{*}\right)=1. \tag{10}$$

Thus, perfectly informing morally motivated individuals is, in general, neither necessary nor sufficient for Pareto-efficiency.

Perfect information may reduce the extent of the market failure compared to the Laissez-faire:

$$\Phi\left(0,i^{*}\right)<\Phi\left(0,0\right), \text{ iff (1) } \mu<(>)\frac{1}{\kappa} \text{ and}$$
$$\text{(2) } \kappa<(>)\frac{m\left(0,i^{*}\right)}{\mu} \text{ and} \tag{11}$$
$$\text{(3) there is no strong reversal of market failure.}$$

**Proof**: see Appendix A.5

The provision of information has two effects on moral motivation: a direct crowding effect and an indirect information effect. The direct crowding effect raises person *A*'s moral motivation. The indirect information effect occurs because the provision of information changes *A*'s knowledge of the externality, which may either increase or decrease. If *A* underestimates the externality in the Laissez-faire scenario, $\kappa<1$, information provision increases *A*'s knowledge of the externality and hence the impact of *A*'s moral motivation increases. If *A* overestimates the externality in the Laissez-faire scenario, $\kappa>1$, the indirect information effect weakens the effect of *A*'s moral motivation as *A*'s knowledge decreases. The direct and the indirect effect of the provision of information are hence additive for $\kappa<1$ and cause a net increase of moral motivation. For $\kappa>1$, they are countervailing and cause a net increase (reduction) of moral motivation if the crowding effect is stronger (weaker) than the information effect.

Furthermore, Eq. (10) holds under very specific conditions only, as the crowding-in effect must be of a given extent. For strong (weak) crowding-in, that is for $m\left(0,i^{*}\right)>(<)1$, provision of information does not lead to efficiency. Since all variables in Eq. (10) are exogenous to the government, it would be purely coincidental for Eq. (10) to hold. Therefore, the provision of perfect information is, in general, not an efficient single instrument.

However, perfect information may be effective in reducing the market failure as compared to the Laissez-faire ($\Phi\left(0,i^{*}\right)<\Phi\left(0,0\right)$, Eq. (11)). This requires in any case that perfect information does not lead to a strong reversal of the market failure. Further, for *A* being undermotivated ($\mu<1/\kappa$), perfect information reduces the market failure if its motivation effect is larger than its information effect ($m\left(0,i^{*}\right)/\mu>\kappa$). For A being overmotivated ($\mu>1/\kappa$), perfect information reduces the market failure if its motivation effect is smaller than its information effect ($m\left(0,i^{*}\right)/\mu<\kappa$).

### 4.4 Complementary policy

Frey (1999) proposes a third policy option as efficient alternative: a complementary policy approach. He argues that "[…] where an instrument tends to crowd out the intrinsic motivation […], an instrument tending to crowd in environmental morale should be used" (Frey 1999: 412). His argumentation remains intuitive, lacking a clear analytical or empirical proof or test. In this section, we use our model to test his hypothesis analytically.

More specifically, we analyze whether a complementary policy, that is a tax $t$ complemented with provision of perfect information $i = i^*$, leads to efficiency and overcomes the problems discussed in Section 4.2 which tax policies may cause when there are motivation crowding effects.

**Proposition 4**

There exists at least one complementary tax rate $\hat{t}^{\text{c}} \leq d'(\hat{x})$ so that the equilibrium allocation $x^{\text{c}} = x^*\left(\hat{t}^{\text{c}}, i^*\right) > 0$ is Pareto-efficient. The corresponding first order condition is given by

$$\hat{t}^{\text{c}} + \left(m\left(\hat{t}^{\text{c}}, i^*\right) - 1\right) d'(\hat{x}) = 0 \ . \tag{12}$$

Compared to $\hat{t}$ (discussed in Proposition 2), $\hat{t}^{\text{c}}$ has the following properties:

(1) $\hat{t}^{\text{c}}$ is smaller than $\hat{t}$, if and only if the motivation effect of perfect information is larger than its information effect:

$$\hat{t}^{\text{c}} < \hat{t} \quad \text{iff} \ \frac{m\left(\hat{t}^{\text{c}}, i^*\right)}{m\left(\hat{t}, 0\right)} > \kappa \ . \tag{13}$$

(2) An inefficiently low tax rate $t^{\text{low}} < \hat{t}, \hat{t}^{\text{c}}$ yields a smaller extent of market failure in the complementary setting than in the tax-only setting, if and only if the motivation effect of perfect information is larger than its information effect and there is no strong reversal of market failure:

$$\Phi\left(t^{\text{low}}, i^*\right) < \Phi\left(t^{\text{low}}, 0\right) \quad \text{iff} \quad (1) \ \frac{m\left(t^{\text{low}}, i^*\right)}{m\left(t^{\text{low}}, 0\right)} > \kappa \ \text{and} \tag{14}$$

(2) there is no strong reversal of market failure.

(3) $\hat{t}^{\text{c}}$ does not fully crowd-out moral motivation while $\hat{t}$ does, if the motivation effect from the complementary information is strong:

$$m\left(\hat{t}^{\text{c}}, i^*\right) > m\left(\hat{t}^{\text{c}}, 0\right) = 0 \quad \text{iff} \ \frac{m\left(\hat{t}^{\text{c}}, i^*\right)}{m\left(\hat{t}, 0\right)} > 0 \, . \tag{15}$$

**Proof**: see Appendix A.6

Proposition 4 shows that combining a tax with the provision of perfect information leads to Pareto efficiency for all parameter values. But more interestingly, a complementary policy may be superior or inferior to a tax-only policy with respect to the three weaknesses discussed in Section 4.2.

First, an efficient complementary policy may require a lower, equal or higher tax rate than a tax-only policy (Eq. (13)). Consider the case when individual $A$ is overmotivated ($\mu > 1/\kappa$). In this case, we find that a complementary policy requires a higher subsidy rate or a higher taxe rate than a tax only policy except when the provision of information causes a net reduction of moral motivation.[21] In other words, in a situation in which $A$ is overmotivated, complementing a tax with an instrument which further crowds-in moral motivation, does not make sense as this requires an even higher taxe rate. Now, consider the case when individual $A$ is undermotivated ($\mu > 1/\kappa$). In this case, we find that a complementary policy allows for a lower or equal tax rate than a tax-only policy except, again, the provision of information causes a net reduction of moral motivation. The equality of the tax rates occurs if and only if a tax rate at the level of marginal damage crowds-out all moral motivation with and without provision of moral motivation. This result shows that, even in this case, a complementary policy may or may not be an improvement over a tax-only policy with respect to allowing for a lower tax rate. Yet, it allows for a lower tax rate in the special case in which $A$ is undermotivated, underestimates the externality and in which the marginal crowding effect of the tax is smaller than its relative price effect.

Second, at the same inefficiently low tax rate, a complementary policy may yield a smaller extent of market failure than a tax-only policy (Eq. (14)). The intuition is the following: if $A$ is undermotivated and provision of information causes a net increase in moral motivation for every given tax rate, then complementing any tax with perfect information must reduce $x^*$. This reduction in $x^*$ is also a reduction of the market failure if there is no strong reversal of market failure. In other words, an inefficiently low complementary tax may still exacerbate the market failure but less than a tax-only policy.

Third, there exists an efficient complementary policy which does not fully crowd out moral motivation while an efficient tax-only policy would do so, if the provision of information causes a strong net increase in moral motivation and/or a strong cross reduction in the marginal crowding of taxes such that $m\left(d'(\hat{x}),0\right)=0$ and $m\left(d'(\hat{x}),i^*\right)>0$ (Eq. (15)).

The superiority of a complementary policy approach as hypothesized by Frey (1999) can thus only be confirmed for specific parameter constellations. Our results suggest that, first, a tax should be complemented with the provision of perfect information if and only if $A$ is undermotivated. Second, such a complementary policy reduces the risk of exacerbating the market failure by inefficiently low taxes. Third, for certain parameter constellations, an

[21] Remember that this requires that individuals overestimate the externality and that the marginal information effect of the provision of information is stronger than its marginal crowding effect.

efficient complementary policy does not fully crowd-out moral motivation while the efficient tax-only policy does.

# 5 Conclusion

We have studied motivation-crowding to analyze the influence of governmental policies on individual responsibility in a situation of negative consumption externalities and motivation crowding. To this end, we have formulated a model where we model motivation crowding as a preference change due to extrinsic intervention, namely taxes and provision of information. We have shown that in the absence of government regulation, responsible behavior will, in general, not lead to Pareto efficiency. Only if the individuals' basic moral motivation and knowledge meet a very restrictive condition, responsible behavior leads to Pareto-efficiency. It is much more likely that individuals' are either under-motivated or over-motivated. If individuals are under-motivated, moral motivation diminishes the extent of the market failure. The necessity for governmental intervention then remains, but becomes less urgent than if there was no moral motivation. If individuals are over-motivated, moral motivation increases the market failure.

Further, we have shown that a Pigouvian tax as a single instrument is Pareto efficient in all situations. There may exist more than one efficient tax rate. Motivation crowding thus does not question the efficiency of taxes. But it creates three problems with taxation: first, crowding requires a higher tax rate which may be difficult for a government to implement due to political pressure. Second, setting a tax rate inefficiently low may exacerbate the market failure. And third, an efficient tax rate may fully crowd-out moral motivation which is harmful if there are motivational spill-overs.

For the provision of descriptive information, our analysis shows that it might lead to Pareto efficiency under very restrictive assumptions. But it may be well suited to diminish the extent of the market failure. The instrument should be used with caution since its effectiveness is contingent on several parameter values: individuals' knowledge, their basic moral motivation and the extent the information crowding effect. For example, when individuals consume excessive amounts of a polluting good and underestimate the externality, provision of information diminishes the market failure if the crowding effects are not too strong.

Since both instruments, taxes and provision of information, have serious weaknesses when applied on their own, we considered a third policy option: a complementary policy, consisting of both instruments (as e.g. proposed by Frey 1999). Such a complementary policy may require a lower taxe rate, may reduce the risk of exacerbating the market failure by inefficiently low taxes, and may lead to efficiency without fully crowding-out moral motivation. The drawback is that these effects are highly contingent on parameter values. We thus can recommend a complementary policy for some but not for all cases.

For decision makers facing externalities, our study shows that the extent of crowding effects should be tested before implementing a policy regime. It is necessary to find out if there are one or more efficient tax rates to be able to choose the one with the most desired side effects, such as tax income level or incentive distortions. Further, governments should know if provision of information is at least effective in diminishing the market failure or if it is counterproductive. If there are motivational spill-overs to be expected, governments should consider a complementary policy if crowding effects from information are strong enough. Lastly, governments should be aware that they should not implement a tax at all rather than implementing an inefficiently low tax which may increase the problem.

For economists, our study has two major implications. First, empirical research needs to further investigate in how far higher taxes or levels of information cause stronger crowding than lower taxes or levels of information, and in how far complementary instruments affect the crowding effects of taxes. These insights will be crucial in understanding the efficiency and effectiveness of taxes and other instruments. Second, our analysis suggests that economists should re-think existing market based instruments. One seemingly fruitful starting point is a paper by Mellström and Johannesson (2008). They show that crowding effects of taxes are contingent on the redistribution regime. The full effects of ecological tax reforms may thus depend on whether the tax income is e.g. spent for environmental innovation or for pension funds. Still, it may even be necessary to think of new instruments or draw more attention to the use of command and control instruments, since their effectiveness is not contingent on crowding effects.

# 6 Appendix

### A.1 Proof of Lemma 1

*Definition*

For given income distribution $\left(I^A, I^B\right)$ and government policy $\left(t, i\right)$, an allocation $\left(y^{A*}, y^{B*}, x^*\right)$ and price system $\left(1, p\right)$ is an *equilibrium* if and only if it has the following properties:

Both individuals *A* and *B* take prices $\left(1, p\right)$ and income $\left(I^A, I^B\right)$ as given.

For both individuals, the equilibrium allocation is a utility maximum s.t. the respective budget constraint:

a. $$\max_{y^A, x} \; U^A\left(x^*, y^{A*}\right) \text{ s.t. } I^A = y^A + \left(p+t\right)x \tag{16}$$

b. $$\max_{y^B} \; U^B\left(y^{B*}\right) \text{ s.t. } I^B = y^B \tag{17}$$

Supply equals demand in the markets for both goods:

$$y^{A*}\left(I^A, p\right) + y^{B*}\left(I^B, p\right) = y^S \tag{18}$$

$$x^*\left(I^A, p\right) = x^S \tag{19}$$

*Equilibrium conditions*

Utility maximization of individual *A* leads to the following Lagrangian:

$$L\left(y^A, x, \lambda^A\right) = y^A + u(x) - m(t,i)k(i)d(x) + \lambda^A\left(I^A - y^A - (p+t)x\right). \tag{20}$$

Differentiating with respect to $y^A$, *x* and $\lambda^A$ yields three first-order conditions, from which it is apparent that $\lambda^A = 1$. With this, the two remaining first-order conditions are:

(A) $$u'\left(x^*\right) = p + t + m(t,i)k(i)d'\left(x^*\right) \quad , \tag{21}$$

(B) $$I^A = y^{A*} + (p+t)x^*. \tag{22}$$

Utility maximization of individual *B* leads to the following Lagrangian:

$$L\left(y^B, \lambda^B\right) = y^B - d(x) + \lambda^B\left(I^B - y^B\right). \tag{23}$$

The resulting first-order condition requires that B spends all his income on $y^B$:

(C) $$I^B = y^{B*}. \tag{24}$$

As above, let $x^S$ and $y^S$ denote total supply of good *X* and *Y*. Market clearing conditions are given by:

(D) $$y^{A*} + y^{B*} = y^S \tag{25}$$

(E) $$x^*(p) = x^S \tag{26}$$

Conditions (A) – (E) characterize the equilibrium.

*Solution*

We now show that conditions (A) – (E) hold for the assumptions of our model.

*Condition (A):*

The left-hand side of condition (A) is positive and decreasing. Per assumption, for all $I^A, t, i, p$, it is characterized by:

$$u'(x) > 0,\ u''(x) < 0,\ \lim_{x\to 0} u'(x) = +\infty,\ \lim_{x\to\infty} u'(x) = 0 \tag{27}$$

The right-hand side of condition (A) is positive and increasing, given that for all $I^A, t, i, p$:

$$d'(x) > 0,\ d'(0) = 0,\ d''(x) > 0 \tag{28}$$

It follows that there exists a $x^* > 0$ for which condition (A) holds.

*Condition (B) and (C):*

Since it is possible to consume infinitively small amounts of both goods *Y* and *X*, the income of each individual must be large enough to fulfill condition (B) and (C).

*Condition (D) and (E):*

Per assumption, prices of both goods are exogenously given and fulfill the market clearing conditions.

Since conditions (A) – (E) are fulfilled by one $x^* > 0$, we conclude:

A unique and stable interior equilibrium with $x^* > 0$ exists for all $I^A, t, i$. The Equilibrium is characterized by Equation (4). □

### A.2 Proof of Lemma 2

We find the necessary first-order conditions for Pareto-efficiency by solving the following maximization problem:

$$\max_{y^A, y^B, x} U^A\left(x, y^A\right) \text{ s.t. } \bar{U}^B = y^B - d\left(x\right) \text{ and } I^A + I^B + tx = y^A + y^B + \left(p+t\right)x . \tag{29}$$

The Lagrangian is given by:

$$\mathrm{L}\left(y^A, y^B, x, \lambda, \delta\right) = y^A + u\left(x\right) + \lambda\left(y^B - d\left(x\right) - \bar{U}^B\right) + \delta\left(I^A + I^B - y^A - y^B - px\right) . \tag{30}$$

Differentiating with respect to $y^A$, $y^B$ and $x$ yields three first-order conditions, from which it is apparent that $\lambda = 1$ and $\delta = 1$. The remaining first-order condition is:

$$u'\left(\hat{x}\right) = p + d'\left(\hat{x}\right) . \tag{31}$$

Since $\lim_{x \to 0} u'\left(x\right) \to +\infty$, $u'' < 0$, and $d'' > 0$ there exists a unique $\hat{x} > 0$ solving Eq. (31).

□

### A.3 Proof of Proposition 1

Using Eq. (4) with $t, i = 0$, the equilibrium allocation in the laissez-faire scenario is characterized by

$$u'\left(x^{\mathrm{LF}}\right) = p + \mu\kappa d'\left(x^{\mathrm{LF}}\right) . \tag{32}$$

Comparison with Eq. (5) shows that the equilibrium allocation is Pareto-efficient, i.e. $x^{\mathrm{LF}} = \hat{x}$, if and only if:

$$p+\mu\kappa d'\left(x^{\mathrm{LF}}\right)=p+d'\left(x^{\mathrm{LF}}\right) \ . \tag{33}$$

Simple rearrangement yields:

$$\mu=\frac{1}{\kappa} \ . \tag{34}$$

Market failure

Using Eq. (32), the total derivative of the equilibrium level $x^{\mathrm{LF}}$ with respect to $\mu$ is given by:

$$\frac{\mathrm{d}x^{\mathrm{LF}}}{\mathrm{d}\mu}=\frac{\kappa d'\left(x^{\mathrm{LF}}\right)}{u''\left(x^{\mathrm{LF}}\right)-\mu\kappa d''\left(x^{\mathrm{LF}}\right)} \ . \tag{35}$$

The right hand side of Eq. (35) is clearly negative and $x^{\mathrm{LF}}$ decreases with $\mu$. It follows that $x^{\mathrm{LF}}>(<)\hat{x}$ for all $\mu<(>)1/\kappa$. Since further $\mathrm{d}\hat{x}/\mathrm{d}\mu=0$ (from Eq. (5)), it follows that $\mathrm{d}\Phi/\mathrm{d}\mu<(>)0$ for $\mu<(>)1/\kappa$. Hence, the extent of the market failure decreases with $\mu$ if $\mu<1/\kappa$ and increases otherwise. □

**A.4 Proof of Proposition 2**

Using Eq. (4) with $i=0$ and $t\in\mathbb{R}$, the equilibrium allocation in the tax-policy scenario is characterized by:

$$u'\left(x^{t}\right)=p+t+m(t,0)\kappa d'\left(x^{t}\right) \ . \tag{36}$$

Comparison with Eq. (5) shows that the equilibrium allocation is Pareto-efficient, i.e. $x^{t}=\hat{x}$, if and only if:

$$p+\hat{t}+m\left(\hat{t},0\right)\kappa d'(\hat{x})=p+d'(\hat{x}) \ . \tag{37}$$

Simple rearrangements yield:

$$\hat{t}+\left(m\left(\hat{t},0\right)\kappa-1\right)d'(\hat{x})=0 \ . \tag{38}$$

Call $\Omega(t,0)=\Omega(t)=t+\left(m(t,0)\kappa-1\right)d'(\hat{x})$, which is the left hand side of Eq. (38). We now show that there exists at least one $\hat{t}$ such that $\Omega\left(\hat{t}\right)=0$. We know from the Laissez-faire solution that $\Omega(0)=0$ only holds for $\mu=1/\kappa$.

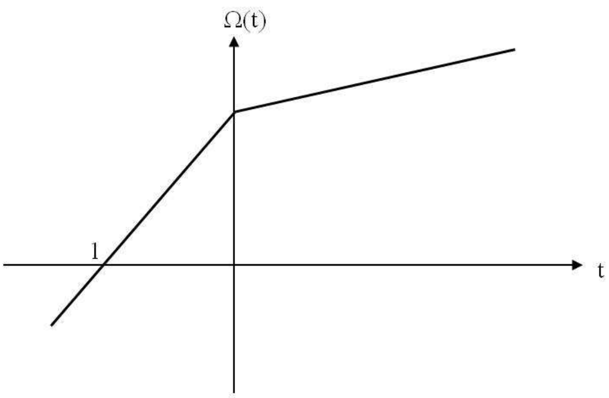
Suppose that $\Omega(0)>0$, which requires that $\mu>1/\kappa$. There are three possible solutions.

The first solution is a subsidy and exists for all parameter values. It is straightforward that at $\underline{t} = -m(t,0)\kappa d'(\hat{x})$, $\Omega(\underline{t}) < 0$. Due to monothonicity of $m(t,0)$, it follows from the Intermediate Value Theorem that there exists a $\hat{t}$, with $\underline{t} < \hat{t} < 0$, such that $\Omega(\hat{t}) = 0$.

The second solution is a positive tax and exists if and only if $m(d'(\hat{x}),0) = 0$, from what immediately follows that there exists a $\hat{t} = d'(\hat{x})$ with $\Omega(\hat{t}) = 0$.

The third solution is also a positive tax and also exists if and only if $m(d'(\hat{x}),0) = 0$. If and only if $m(d'(\hat{x}),0) = 0$, there must exist a $t_0$, with $0 < t_0 < d'(\hat{x})$, such that $n(t_0,0) = 0$ and $\Omega(t_0) < 0$. Since $\Omega(0) > 0$, there must also exist a $\tilde{t}$, with $0 < \tilde{t} < t_0$, such that $m(\tilde{t},0) = 1/\kappa$ and $\Omega(\tilde{t}) > 0$. Due to monothonicity of $m(t,0)$, it follows from the Intermediate Value Theorem that there exists a $\hat{t}$, with $\tilde{t} < \hat{t} < t_0$, such that $\Omega(\hat{t}) = 0$.

Note that solution 2 and 3 require that the marginal crowding effect of the tax is larger than its relative price effect, that is $m_t \kappa d'(\hat{x}) < -1$. Figure 1 shows the possible shapes of $\Omega(\hat{t})$ for linear crowding and the respective solutions $\Omega(t) = 0$ for $\Omega(0) > 0$:



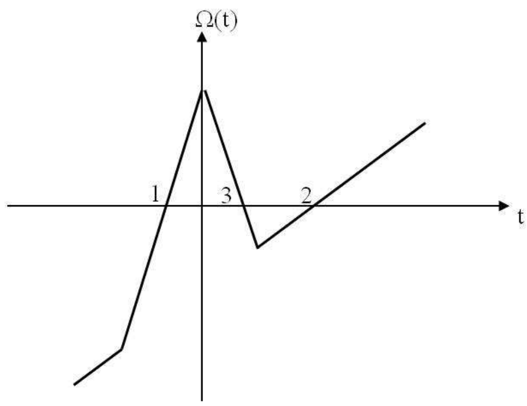




Figure 1: Possible efficient tax rates for $\Omega(0) > 0$

Now suppose that $\Omega(0) < 0$, which requires that $\mu < 1/\kappa$. There are one or two solutions:

Suppose that $m(d'(\hat{x}),0) > 0$. It follows that $\Omega(d'(\hat{x})) > 0$. Due to monothonicity of $m(t,0)$, it follows from the Intermediate Value Theorem that there exists a unique $\hat{t}$, with $0 < \hat{t} < d'(\hat{x})$, such that $\Omega(\hat{t}) = 0$.

Suppose that $m(d'(\hat{x}),0) = 0$. It follows immediately that there exists a $\hat{t} = d'(\hat{x})$ with $\Omega(\hat{t}) = 0$.

Suppose again that $m\left(d'(\hat{x}),0\right)=0$. It follows that there exists a $t_0$, with $0<t_0<d'(\hat{x})$, such that $n(t_0,0)=0$ and $\Omega(t_0)<0$. Now further suppose that the crowding term $m(t,0)$ is concave in *t*, such that $m_{tt}<0$. It follows that there exists a a $t_{\max}$, with $0<t_{\max}<t_0$, such that $\Omega_{t_{\max}}=0$. This implies that $\Omega(t_{\max})$ is a local maximum. If and only if $\Omega(t_{\max})>0$ there exist two $\hat{t}$, with $0<\hat{t}<t_{\max}$, such that $\Omega(\hat{t})=0$. If and only if $\Omega(t_{\max})=0$ there exists one $\hat{t}$, with $0<\hat{t}<t_{\max}$, such that $\Omega(\hat{t})=0$. Both requires a minimum degree of concavity of $m(t,0)$.

Figure 2 shows possible shapes of $\Omega(\hat{t})$ for and the respective solutions $\Omega(t)=0$ for $\Omega(0)<0$:

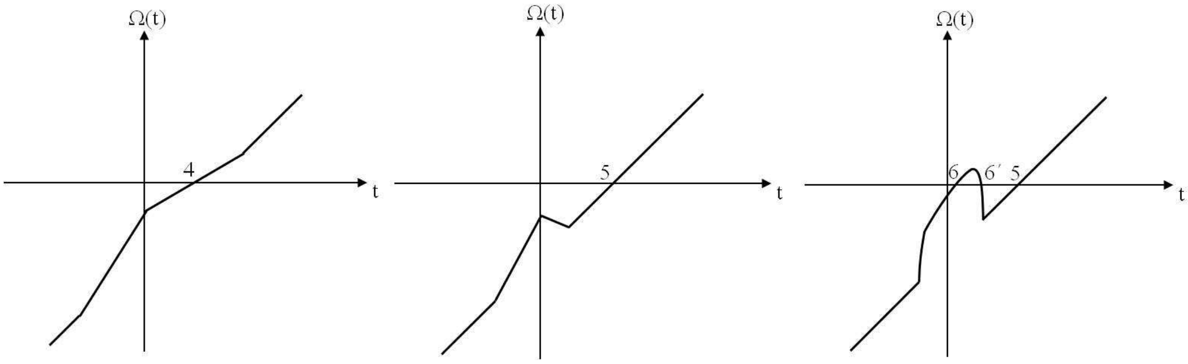



Figure 2: Possible efficient tax rates for $\Omega(0)<0$

Market failure:

Applying Eq. (6) to the tax only scenario, the market failure is given by $\Phi\left(x^t,\hat{x}\right)=\left|x^t-\hat{x}\right|$. Since $\hat{x}$ is not contingent on *t*, the total derivative of $\Phi\left(x^*,\hat{x}\right)$ with respect to *t* equals the total derivative of $\hat{x}$ with respect to *t*: $\mathrm{d}\Phi\left(x^t,\hat{x}\right)/\mathrm{d}t=\mathrm{d}x^t/\mathrm{d}t$.

Using Eq. (36), we derive the total derivative of $\hat{x}$ with respect to *t*. Rearrangements yield:

$$\frac{\mathrm{d}x^t}{\mathrm{d}t}=\frac{1+m'(t,0)\kappa d'\left(x^t\right)}{u''\left(x^t\right)-m(t,0)\kappa d''\left(x^t\right)}\ . \tag{39}$$

Eq. (39) is positive for $m'(t,0)\kappa d'\left(x^t\right)<-1$ and negative otherwise.

Inefficiently low tax rates require that $\mu<1/\kappa$ and yield $x^t>\hat{x}$. From Eq. (39) follows that increasing an inefficiently low tax rate increases $x^t$ and thus also the market failure $\Phi\left(x^*,\hat{x}\right)$.

□

### A.5 Proof of Proposition 3

Using Eq. (4) with $i = i^*$ and $t = 0$, the equilibrium allocation in the information-policy scenario is characterized by

$$u'\left(x^i\right) = p + m\left(0, i^*\right) d'\left(x^i\right) \ . \tag{40}$$

Comparison with Eq. (5) shows that the equilibrium allocation is Pareto-efficient, i.e. $x^i = \hat{x}$, if and only if:

$$p + m\left(0, i^*\right) d'\left(\hat{x}\right) = p + d'\left(\hat{x}\right), \tag{41}$$

which can be rearranged to:

$$m\left(0, i^*\right) = 1 \ . \tag{42}$$

Market failure

Information policy reduces the market failure $\Phi\left(x^*, \hat{x}\right)$ if its equilibrium allocation $x^{\mathrm{i}}$ deviates less from the efficient allocation $\hat{x}$ than the Laissez-faire allocation $x^{\mathrm{LF}}$:

$$\left|x^{\mathrm{i}} - \hat{x}\right| < \left|x^{\mathrm{LF}} - \hat{x}\right|. \tag{43}$$

As there are 4 distinguished cases, Eq. (43) has 4 solutions as shown in Table 1:

| | $x^{\mathrm{LF}} > \hat{x}$ | $x^{\mathrm{LF}} < \hat{x}$ |
|---|---|---|
| $x^{\mathrm{i}} > \hat{x}$ | a) $x^{\mathrm{i}} < x^{\mathrm{LF}}$ | b) $x^{\mathrm{i}} < 2\hat{x} - x^{\mathrm{LF}}$ |
| $x^{\mathrm{i}} < \hat{x}$ | c) $x^{\mathrm{i}} > 2\hat{x} - x^{\mathrm{LF}}$ | d) $x^{\mathrm{i}} > x^{\mathrm{LF}}$ |

Table 1: Solutions of Eq. (43)

Information policies have two effects, a crowding and an information effect. If and only if $\kappa < 1 = k\left(i^*\right)$ both effects have the same direction and $x^{\mathrm{i}} < x^{\mathrm{LF}}$. However, if and only if $\kappa > 1$, the two effects are countervailing. It follows that $x^{\mathrm{i}} < (>) x^{\mathrm{LF}}$ if $\frac{m\left(0, i^*\right)}{\mu} > (<) \kappa$, i.e. if the motivation effect $\frac{m\left(0, i^*\right)}{\mu}$ is larger (smaller) than the information effect $\kappa$ [22]. Case a) ( $x^{\mathrm{i}} < x^{\mathrm{LF}}$) is thus solved for all $\kappa < \frac{m\left(0, i^*\right)}{\mu}$. Case d) ( $x^{\mathrm{i}} > x^{\mathrm{LF}}$) is solved for all $\kappa > \frac{m\left(0, i^*\right)}{\mu}$.

[22] The full denotation of the information effect is $\kappa / k\left(i^*\right)$, which equals $\kappa$ as $k\left(i^*\right) = 1$.

The cases b) and c) in Table 1 indicate that the impact of information policies on the consumption levels must not be too large in order to mitigate the market failure. If, e.g. in case c), condition $x^{\mathrm{i}} > 2\hat{x} - x^{\mathrm{LF}}$ is violated, the consumption level decreases from a inefficiently high level $x^{\mathrm{LF}}$ to an inefficiently low level $x^{\mathrm{i}}$ such that the resulting deviation from the efficient consumption level is larger than in the laissez-faire scenario. This is a strong reversal of the market failure. These cases cannot be solved analytically.

□

**A.6 Proof of Proposition 4**

The proof of the existence of one or several efficient complementary tax rates and their respective sign is the same as in A.4 if you set $\kappa = 1$ and if you exchange all $m(t,0)$ with $m(t,i^*)$.

Using Eq. (4) $i = i^*$ and $t \in \mathbb{R}$, the equilibrium allocation in the complementary policy scenario is characterized by

$$u'(x^{\mathrm{c}}) = p + t^{\mathrm{c}} + m(t^{\mathrm{c}}, i^*) d'(x^{\mathrm{c}}) \; . \tag{44}$$

Comparison with Eq. (5) shows that $x^{\mathrm{c}}$ is Pareto-efficient, i.e. $x^{\mathrm{c}} = \hat{x}$, if and only if:

$$\hat{t}^{\mathrm{c}} + \left( m(\hat{t}^{\mathrm{c}}, i^*) - 1 \right) d'(\hat{x}) = 0 \, . \tag{45}$$

Property (1)

Comparison of Eq. (45) with Eq. (38) shows that $\hat{t}^{\mathrm{c}}$ is smaller than $\hat{t}$, if and only if the motivation effect of complementary perfect information is larger than its information effect:

$$\hat{t}^{\mathrm{c}} < \hat{t} \text{ iff } \frac{m(\hat{t}, i^*)}{m(\hat{t}, 0)} > \kappa \, . \tag{46}$$

This condition is fulfilled for all $\kappa < 1$ but only for some $\kappa > 1$.

Property (2)

Suppose that $\mu < 1/\kappa$ from which follows that $x^*(0,0) > \hat{x}$. Suppose further that $m'(t,0)\kappa d'(x^t) < -1$ such that $\frac{\mathrm{d}x^t}{\mathrm{d}t} > 0$ which implies that marginal increases in the tax rate yield larger market failure, until *t* reaches a critical level $t^0$ such that $m(t^0, 0) = 0$. Further increases in the tax rate will lead to a decrease in *x* as all moral motivation has been crowded out.

Now suppose that e.g. due to lobby pressure the government chooses a tax rate $t^{\text{low}} < t^0$ and $t^{\text{low}} < \hat{t}, \hat{t}^{\text{c}}$. Given the assumptions, this leads to a larger market failure in a tax-only scenario than in the laissez-faire scenario if there is no strong reversal of the market failure.

Equating Eq. (36) and (44) shows that $t^{\text{low}}$ yields the same market failure in the tax-only scenario and in the complementary policy scenario, if and only if $m\left(t^{\text{low}}, i^*\right) = m\left(t^{\text{low}}, 0\right)\kappa$. Since $x^{\text{c}}$ decreases with $m\left(t^{\text{low}}, i^*\right)$ (from Eq. (44)), it follows that the market failure at $t^{\text{low}}$ is larger in the tax-only scenario than in the complementary policy scenario if and only if $m\left(t^{\text{low}}, i^*\right) > m\left(t^{\text{low}}, 0\right)\kappa$, or rearranged:

$$\frac{m\left(t^{\text{low}}, i^*\right)}{m\left(t^{\text{low}}, 0\right)} > \kappa \ , \tag{47}$$

and if there is no strong reversal of the market failure.

Property (3)

From Eq. (45) follows that there exists a $\hat{t}^{\text{c}}$ which does not fully crowd out moral motivation while $\hat{t}$ would if and only if:

$$m\left(\hat{t}^{\text{c}}, i^*\right) > m\left(\hat{t}, 0\right) = 0 . \tag{48}$$

□

# 7 References


Aidt, T. (1998). Political internalization of economic externalities and environmental policy. *Journal of Public Economics* 69: 1–16.

Andreoni, J. (1988). Privately provided public goods in a large economy: the limits of altruism. *Journal of Public Economics* 35: 57–73.

Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *The Economic Journal* 100: 464–477.

Ariely, D., Bracha, A. and Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review* 99(1): 544–555.

Asheim, G. (2010). Strategic use of environmental information. *Environmental and Resource Economics* 46(2): 207–216.

Baumgärtner, S., Faber, M. and Schiller, J. (2006). *Joint Production and Responsibility in Ecological Economics. On the Foundations of Environmental Policy*. Edward Elgar, Cheltenham.

Baumol, W. (1972). On taxation and the control of externalities. *American Economic Review* 62(3): 307–322.

Bowles, S. and Hwang, S. (2008). Social preferences and public economics: mechanism design when social preferences depend on incentives. *Journal of Public Economics* 92(8–9): 1811–1820

Brekke, K., Kverndokk, S. and Nyborg, K. (2003). An economic model of moral motivation. *Journal of Public Economics* 87: 1967–1983.

Charness, G. and Gneezy, U. (2009). Incentives to exercise. *Econometrica* 77(3): 909–931.

d'Adda, G. (2011). Motivation crowding in environmental protection: evidence from an artefactual field experiment. *Ecological Economics* 70(11): 2083–2097.

Deci, E. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology* 18(1): 105–115.

Deci, E. and Ryan, R. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. New York: Plenum.

Deci, E., Koestner, R. and Ryan, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 125(6): 627–668.

Diamond, P. (2006). Optimal tax treatment of private contributions for public goods with and without warm glow preferences. *Journal of Public Economics* 90(4-5): 897–919.

Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *American Economic Review* 96(5): 1611–1630.

Fredriksson, P. (1997). The political economy of pollution taxes in a small open economy. *Journal of Environmental Economics and Management* 33: 44–58.

Frey, B.S. (1997). *Not Just for the Money. An Economic Theory of Personal Motivation*. Cheltenham, UK: Edward Elgar.

Frey, B. and Oberholzer-Gee, F. (1997). The cost of price incentives: an empirical analysis of motivation crowding-out. *American Economic Review* 87(4): 746–755.

Frey, B. (1999). Morality and rationality in environmental policy. *Journal of Consumer Policy* 22: 395–417.

Frey, B. and Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys* 15(5): 589–611.

Gneezy, U., and Rustichini, A. (2000a). A fine is a price. *Journal of Legal Studies* 29(1): 1–17.

Gneezy, U. and Rustichini, A. (2000b). Pay enough or don't pay at all. *Quarterly Journal of Economics* 115(3): 791–810.

Gneezy, U., Meier, S. and Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives* 25(4): 191–210.

Goeschl, T. and Perino, G. (2012). Instrument choice and motivation: evidence from a climate change experiment. *Environmental and Resource Economics* 52(2): 195–212.

Hammond, P. (1978). Altruism. In: Eaton, J., Milgate, M. and Newman, P. (Eds.). *The New Palgrave: a Dictionary of Economics*. Macmillan Press, London.

Heyes, A. and Kapur, S. (2011). Regulating altruistic agents. *Canadian Journal of Economics/Revue canadienne d'économique* 44(1): 227–246.

Heyman, J. and Ariely, D. (2004) Effort for Payment: A tale of two markets. *Psychological Science* 15(11): 787 –793.

Kallbekken, S., Westskoga, H., and Mideksaa, T. (2009). Appeals to social norms as policy instruments to address consumption externalities. *Journal of Socio-Economics* 39(4): 447–454.

Kunda, Z. and Schwartz, S. (1983). Undermining intrinsic moral motivation: external reward and self-presentation. *Journal of Personality and Social Psychology* 45(4): 763–771.

Mellström, C. and Johannesson, M. (2008). Crowding out in blood donation: was Titmuss right? *Journal of the European Economic Association* 6(4): 845–863.

Nyborg, K. and Rege, M. (2003a). Does public policy crowd out private contributions to public goods? *Public Choice* 115: 397–418.

Nyborg, K. and Rege, M. (2003b). On social norms: the evolution of considerate smoking behavior. *Journal of Economic Behavior & Organization* 52: 323–340.

Nyborg, K. (2011). I don't want to hear about it: Rational ignorance among duty-oriented consumers. *Journal of Economic Behavior &* Organization 79: 263–274. .

Perino, G., Panzone, L., Swanson, T. (2011). Crowding-in, crowding-out and over-crowding: The interaction between price and quantity based instruments and intrinsic motivation. *CBESS Discussion Paper 11-10*. Norwich, UK.

Pigou, A. (1932)[1920]. *The Economics of Welfare*. 4th edition. London.

Reeson, A. and Tisdell, J. (2008). Institutions, motivations and public goods: an experimental test of motivational crowding. *Journal of Economic Behavior & Organization* 68(1): 273–281.

Ryan, R. and Deci, E. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology* 25: 54–67.

Sen, A. (1977). Rational fools: A critique of the behavioural foundations of economic theory. *Philosophy and Public Affairs* 6: 317–344.

Singer, P. (1972). Famine, affluence, and morality. *Philosophy and Public Affairs* 1(3): 229–243.

Titmuss, R. (1971). *The Gift Relationship: from Human Blood to Social Policy*. New York, Pantheon Books.

White, R. (1959). Motivation reconsidered: the concept of competence. *Psychological Review* 66: 297–333.

Williams, G. (2008). Responsibility as a virtue. *Ethical Theory and Moral Practice* 11(4): 455–470.

# Paper 4: Endogenous Environmental Policy when Pollution is Transboundary

JOACHIM FÜNFGELT[A] AND GÜNTHER G. SCHULZE[B]

Leuphana University of Lüneburg University of Freiburg

**Abstract:** We analyze the formation of environmental policy to regulate transboundary pollution if governments are self-interested. In a common agency framework, we portray the environmental policy calculus of two political support-maximizing governments that interact strategically with respect to their environmental policies, but too small to affect world market prices. Our model shows how distortions created by the strategic interaction of national governments interact with distortions created by the political processes in both countries. Relatively strong industry lobbies may lead to subsidization of polluting goods. Relatively strong environmental lobbies may improve welfare as they counteract the distortion created by the international externality. A marginal increase in the size of environmentally concerned individuals increases aggregate welfare up to a point due to increased environmental lobbying. However, very strong political distortions create instability and lead to corner solutions.

**JEL classification***:* Q 58, F 5

**Keywords**: Political economy, environmental policy, transboundary pollution, common agency, strategic interaction

[a] Sustainability Economics Group, Leuphana University of Lüneburg, P.O. Box 2440, D-21335 Lüneburg, Germany, fuenfgelt@uni.leuphana.de

[b] Dept. of International Economic Policy, University of Freiburg, 79085 Freiburg i.Br., Germany, Guenther.Schulze@vwl.uni-freiburg.de

# 1. Introduction

This paper examines how lobby groups in pluralistic societies affect the determination of environmental policy when countries are linked through transboundary pollution and their political support maximizing governments are unable to alter prices on the global goods markets.

It is widely recognized that environmental policy formation is influenced by lobby groups. Such lobby groups are present at international conferences for instance at Kyoto, Copenhagen or Cancún; they also affect the formulation of national policies. While environmental lobby groups advocate stricter environmental standards, industry associations often lobby for lower standards in order to retain competitiveness in international markets. Governments seeking to maximize political support respond systematically to such lobbying.[1] The resulting equilibrium regulation differs considerably from the Pigouvian rule, thus creating a politically motivated distortion of environmental policy (Aidt 1998).

Due to the scale of economic activity and the properties of ecological systems, pollution often spills over to neighboring countries making national environmental policies relevant for adjacent countries as well. Transboundary pollution has become a serious challenge over the past decades, especially in East Asia: In China sulphur oxide emissions increased by 53% between 2000 and 2006 and spilled over to Southwest Japan (Lu et al. 2010). Ichikawa and Fujita (1995) estimate that China's contribution to wet sulphate deposition in Japan represent 50% of the total. Furthermore, anthropogenic $NO_x$ emissions over Asia have more than doubled since 1985 (Akimoto 2003). This increase in emissions has created atmospheric brown clouds. They are fuelled by emissions of two or more countries in the region and affect those countries negatively. Brown clouds "start as indoor and outdoor air pollution consisting of particles and pollutant gases, such as nitrogen oxides ($NO_x$), carbon monoxide (CO), sulphur dioxide ($SO_2$), ammonia ($NH_3$), and hundreds of organic gases and acids" (Ramanathan et al. 2008). They affect many small countries. Their hotspots are in East Asia, Indo-Gangetic Plain in South Asia, Southeast Asia, Southern Africa, and the Amazon Basin. They have severe environmental impacts as, for instance, they accelerate the meltdown of Himalayan glaciers, decrease crop yields by as much as 20%, or result in over 330.000 deaths per year in China and India alone, as their particles cause pulmonary illnesses and chronic respiratory problems (Ramanathan et al. 2008).

Another example for multi-directional transboundary pollution affecting small countries is Scandinavian $SO_2$ depositions, which are dependent to a high degree on the emission activity in neighboring countries. While all Scandinavian countries apply emission taxes, actual tax rates differ very strongly. Cansier and Krumm (1997) find that tax rates in Sweden are three

[1] Cf. Binder and Neumayer (2005) and Fredriksson et al. (2005) for empirical evidence on the political influence of environmental lobby groups and List and Sturm (2006) on the relative importance of voters and lobby groups for environmental policies of US states.

times higher than in Denmark, which is only partly due to abatement cost differentials and therefore is hardly explained by welfare-maximizing behavior alone. [2]

Such transboundary pollution gives rise to a second distortion (in addition to the political distortion described above), if national environmental policies remain non-cooperative: even welfare maximizing governments would internalize the externalities only to the extent that they affect their own country (Markusen 1975). Yet, how do these two distortions interact? How do politically-motivated, self-interested governments set environmental policies in the presence of transboundary pollution? This is the concern of our paper. Governments respond to lobbying efforts of opposing lobby groups and at the same time are in a situation of strategic interaction with neighboring governments that are likewise seeking to maximize their political support.

Our study adds to the literature on endogenous environmental policy. Fredriksson (1997) analyzes the effects of world price changes and lobbying on the politically optimal environmental tax rate. He shows that pollution may increase in presence of an abatement subsidy because the pollution tax is reduced due to a change in lobbying influence. Schleich (1997) introduces a second policy instrument and analyzes the choice between domestic taxes and tariffs when the externality is in production or consumption.[3] Aidt (1998) assumes that pollution stems from the use of an input rather than production and demonstrates how a politically optimizing government deviates from the social optimum in deciding on its environmental policy.[4] Fredriksson and Svensson (2003) analyze the effects of interaction of corruption and political instability on endogenous environmental policy. They show that political instability has a negative effect on the stringency of environmental policy if corruption is low and a positive effect if corruption is high. Damania et al. (2003) investigate how the effect of trade liberalization on environmental regulation is affected by corruption levels.

These papers use a common agency model to portray the political game that determines environmental policy. Yet, they do not take into account the strategic interaction that governments are exposed to in the international arena, when deciding on their environmental policies. Thus, the environmental policies within such a framework are determined by domestic considerations alone.[5] To our knowledge the only exception is Conconi (2003) who

---

[2] A related example is the environmental degradation of the Baltic Sea: It is affected by fishing, riverine pollution, and atmospheric nitrogen deposition from the neighboring states (Helcom 2010). Gren (2001) demonstrates the inefficiency of uncoordinated environmental policy for the Baltic Sea.

[3] Schleich and Orden (1999) generalize the small economy case to the large economy setting.

[4] Hillman and Ursprung (1994) analyze the influence of environmental concerns on endogenous trade policy, but they do not study environmental policy formation. Bommer and Schulze (1997) consider the effect of trade liberalization on endogenous trade policy.

[5] Strategic interaction in the determination of environmental policy is analyzed in the literature on transboundary pollution (e.g. Markusen 1975) and the literature on strategic environmental policy (e.g. Barrett 1994). Both strands of literature, however, do not take into account the political-economic rationale in environmental policy

portrays two large open economies, which jointly determine their trade and environmental policies. In her model, strategic interaction occurs as environmental policies alter the world market prices for the traded goods. When a large country taxes the production of a polluting good, the world market price rises and as a consequence foreign production and foreign emissions increase (thus giving rise to 'emission leakages'). Conconi shows that under free trade and in the presence of strong emission leakages, environmental lobbying might actually lower emission taxes as unilaterally formulated taxes will tend to increase degradation. Yet, in her setup strategic interaction occurs only because countries are large on commodity markets, an assumption that does not hold for most countries that are exposed to cross border pollution.

Our paper deviates from her approach by assuming that the economies are small on the globalized world markets and cannot affect world market prices through their environmental or trade policies. While there are countries that may affect world market prices for certain goods, or even a range of goods, we believe that the majority of countries do not have the capacity to influence their terms of trade through a choice of policies. Nonetheless, transboundary pollution remains to be an important policy issue for a number of countries. We thus model two small open economies which produce a pollution intensive good with pollution spilling over to the other country.[6] National governments set their environmental policies in order to maximize their political support, which is composed of voter support and lobbying contributions. We employ a common agency model developed by Bernheim and Whinston (1986) and introduced by Grossman and Helpman (1994) in the literature on endogenous policy formation and assume functionally specified interest groups (environmentalists and industrialists). The strategic interaction between the countries results from transboundary pollution – foreign environmental regulation is a substitute for domestic policy for environmental quality, but it places the burden on foreign rather than domestic producers. Countries may be structurally different in their preferences for environmental quality and their political process, that is, in the strength of lobby groups and in the importance of social welfare for the governments' support.

We show that politically optimal tax rates will exacerbate the environmental degradation compared welfare-maximizing governments if industrial lobbying groups are relatively strong compared to environmental lobby groups. If the industries' political influence is very strong tax rates can even be negative in equilibrium, for one country or for both; a situation that cannot occur in the benevolent dictators' equilibrium.

The effect of politically influential environmental concerns is quite different. We demonstrate that high relative political power of environmental groups may improve welfare, especially if the marginal environmental damage that is caused by production is high, as their lobbying offsets the inefficiency created by strategic interaction of the two governments. In

---

formation. For a comprehensive analysis of the interaction between trade and environmental policy cf. Rauscher (1997), for surveys of the literature see Rauscher (2005).

[6] We exclude thus environmental regulation of *global* pollutants which can be analyzed only in a multi-country setting (cf. Barrett 2003).

that case the political game leads to a higher welfare than non-cooperative social planners would be able to achieve. Even a marginal increase in the size of the environmentalist group increases aggregate welfare up to a point and declines thereafter. If either of the lobby groups becomes too powerful, however, any interior equilibrium is unstable which leads to a corner solution with one country setting a tax prohibitively high. Our paper is the first to study the political economy of environmental policy formation for small open economies in the presence of transboundary pollution and thus fills an important gap in the literature on endogenous environmental policy.

The paper is organized as follows. Section 2 introduces the two country model with transboundary pollution. Section 3 derives the social welfare maximum for non-cooperative governments, which serve as a reference point. Section 4 introduces the common agency approach, derives the politically optimal tax rate, characterizes the equilibrium and simulates it for various parameter constellations, and derives comparative-static results. Section 5 concludes.

## 2. Transboundary Pollution in a Two Country Model

The model consists of two countries, which produce a good that creates environmental pollution. They are small open economies on the goods market but are nevertheless in a situation of strategic interaction, as their pollution spills over to the other country.

### 2.1 The economy

Each economy has two sectors. The first sector produces the non-polluting numeràire good $z$ by labor alone. Units are chosen so that the world and domestic price for the numeràire good equal one. Free trade prevails in both markets; goods prices are determined on the world markets. By choice of units, wage rate is normalized to unity. The second sector produces the polluting good $x$ with labor and a sector-specific factor, which is non-tradable and inelastically supplied. $S$ denotes environmental pollution, which is assumed to affect both countries equally and to be quadratic in total production:

$$S = \beta\left(X + X^{*}\right)^{2} \tag{1}$$

The variable $\beta$ is an exogenously given damage coefficient and $X$ $(X^*)$ is the home (foreign) production of $x$. Foreign country variables are denominated with a “*”. The government levies a tax on each domestically produced unit of $x$ on the producer (if home production of $x$ is positive). The production costs are assumed to be quadratic in the produced quantity. Sector-specific income from the production of $x$ is hence defined as:

$$\Pi(X) = (p - t)X - X^{2} \tag{2}$$

where $p$ is the exogenously given world market price of $x$. Technology exhibits diminishing returns to scale. We assume that in both countries $x$ is produced by only one firm, which chooses $X$ to maximize Eq. (2):

$$X = \frac{1}{2}(p - t) \tag{3}$$

Foreign production is obtained symmetrically. Obviously, the countries only produce positive amounts of *x*, if the respective production tax does not exceed the world market price of good *x*. Countries are in a situation of strategic interaction with respect to their production tax rates as they affect production and thereby pollution in both countries.

As we substitute *X* from Eq. (3), and symmetrically $X^*$, into Eq. (1), we obtain pollution contingent on the tax rates:

$$S(t, t^*) = \frac{\beta}{4}(2p - t - t^*)^2 \tag{4}$$

Pollution increases with the world market price, and decreases with the tax rates. Sector specific income is derived by substituting Eq. (3) in Eq. (2):

$$\Pi(t) = \frac{1}{4}(p - t)^2 \tag{5}$$

Sector specific income decreases with *t* while it increases with the world market price *p*. Total domestic revenue from production taxes, $\tau(t) = tX$, is:

$$\tau(t) = \frac{1}{2}t(p - t) \tag{6}$$

$\tau$ is redistributed uniformly to all citizens of the respective country.[7] Since an increase in the world market price leads to a higher production of *X*, tax revenue increases with the world market price. However, the effect of an increase of the tax rate is ambiguous. On the one hand, a higher tax rate leads to more tax income per unit produced. On the other hand, it leads to fewer units produced as the production of *X* becomes less profitable. Hence, tax revenue $\tau$ increases (decreases) with *t*, iff $p/2 > (<) t$.

### 2.2 Population and Utility Functions

The home country is populated by *N* heterogeneous citizens of three different types: environmentalists, industrialists, and workers. *N* is normalized to one. All citizens have labor income. The total amount of labor in each country equals $l$. Each individual has the same share of $l$. Let $\alpha_E$ be the exogenously given share of environmentalists in the population and $\alpha_I$ ($\alpha_W$) be the share of industrialists (workers). Environmentalists have disutility from pollution while industrialists and workers are not concerned with pollution. Environmentalists and workers derive income from labor only; industrialists also obtain specific factor's income from production of good *x*.

Individual maximization problems are defined as follows:

---

[7] If taxes are negative, all individuals are taxed uniformly. The assumption of uniform redistribution of the tax revenue is in line with the literature and made for simplicity reasons. Cf. Aidt (2010) for an analysis of different refunding schemes.

Each environmentalist solves:

$$\max_{c^z,c^x} U_\mathrm{E} = c^z + u(c^x) - S$$
$$\text{s.t.} \quad l + \tau = c^z + pc^x \tag{7}$$

$c^z$ is consumption of the numeràire good $z$ and $c^x$ is consumption of good $x$. $u(c^x)$ is the concave, differentiable utility function from consumption of $x$. The utility of all environmentalists is equally affected by total pollution.

Each industrialist solves:

$$\max_{c^z,c^x} U_\mathrm{I} = c^z + u(c^x)$$
$$\text{s.t.} \quad l + \tau + \frac{\Pi}{\alpha_\mathrm{I}} = c^z + pc^x \tag{8}$$

The term $\frac{\Pi}{\alpha_\mathrm{I}}$ in Eq. (8) expresses that sector-specific income is equally proportioned to all industrialists. Finally, workers solve:

$$\max_{c^z,c^x} U_\mathrm{W} = c^z + u(c^x)$$
$$\text{s.t.} \quad l + \tau = c^z + pc^x \tag{9}$$

Since prices are given by the world markets, we obtain the following aggregate utility functions of environmentalists, industrialists, and workers:

$$\Omega_\mathrm{E}\left(t,t^*\right) = \alpha_\mathrm{E}[-S + l + \tau] \tag{10}$$

$$\Omega_\mathrm{I}\left(t,t^*\right) = \Pi + \alpha_\mathrm{I}\,[l + \tau] \tag{11}$$

$$\Omega_\mathrm{W}\left(t,t^*\right) = \alpha_\mathrm{W}[l + \tau] \tag{12}$$

The sum of the aggregate utility functions of each country is defined as gross aggregate welfare:

$$\Omega_\mathrm{A}\left(t,t^*\right) \equiv \Omega_\mathrm{E} + \Omega_\mathrm{I} + \Omega_\mathrm{W} = l + \tau + \Pi - \alpha_\mathrm{E} S \tag{13}$$

The term $\alpha_\mathrm{E} S$ represents aggregate disutility of the environmentalists from pollution and thus to the society as a whole. It is the product of total pollution and the share of the environmentalists. Sector specific income, by contrast, is independent of the relative size of industrialists, since $\alpha_\mathrm{I}$ merely defines among how many industrialists the sector-specific income is divided. To obtain gross aggregate welfare –contingent on the tax rates – we substitute Eqs. (4), (5), and (6) in Eq. (13). Rearrangements yield:

$$\Omega_\mathrm{A}\left(t,t^*\right) = \frac{p^2 - t^2}{4} - \alpha_\mathrm{E}\beta\left[\frac{\left(t + t^*\right)^2}{4} + p\left(p - t - t^*\right)\right] + l \tag{14}$$

# 3. Benevolent Dictators' Solution

As a reference point for our further analysis, we derive the benevolent dictators' solution for tax rates that are set non-cooperatively.[8] Each government seeks to maximize its country's aggregate welfare.

### 3.1 Governments' Reaction Functions

The domestic government chooses $t$ in order to maximize Eq. (14), taking the foreign tax rate as given:

$$\frac{\partial \Omega_{\mathrm{A}}(t,t^*)}{\partial t} = \alpha_{\mathrm{E}}\beta(2p - t - t^*) - \frac{t}{2} \overset{!}{=} 0 \tag{15}$$

Solving Eq. (15) for $t$ gives the domestic government's reaction function for positive production of both firms:

$$\tilde{t}^{\mathrm{BD}} = \frac{\alpha_{\mathrm{E}}\beta(2p - t^*)}{1 + \alpha_{\mathrm{E}}\beta} \tag{16}$$

The reaction function is linear in $t^*$, and it is downward sloping. Intuitively, if the foreign tax rate increases, the home country is less affected by the negative externality of foreign pollution and can thus reduce its own tax rate. The foreign country's reaction function is isomorphous. For positive values of foreign production , $t^* < p$, and thus $\tilde{t}^{\mathrm{BD}}$ is positive.

Eq. (16) defines the reaction function only for positive production of both firms, hence for $p > t$ and $p > t^*$. However, we cannot exclude corner solutions. They occur if one country produces so large an amount of the polluting good (thereby producing large quantities of pollution in both countries) that it is optimal for the other country not to add to this pollution by setting a prohibitive tax rate. Any further pollution damage created by own production would exceed the welfare gains from the profits of its firm.

For instance, if the foreign country sets its tax rate equal or below a lower threshold, $t_{\mathrm{lb}}^{\mathrm{BD}*}$, it is optimal for the home country to set its own tax rate prohibitively high so that its production becomes zero. Algebraically, $t_{\mathrm{lb}}^{\mathrm{BD}*}$ is derived by setting Eq. (16) equal to $p$ and solving for $t^*$: $t_{\mathrm{lb}}^{\mathrm{BD}*} = \left(1 - \frac{1}{\alpha_{\mathrm{E}}\beta}\right)p$.[9] Conversely, if the foreign country imposes a prohibitive tax, $t^* \geq p$, the home country will set a tax $t_{\mathrm{low}}^{\mathrm{BD}} = \frac{\alpha_{\mathrm{E}}\beta p}{1+\alpha_{\mathrm{E}}\beta}$, which is derived by setting $t^* = p$ in Eq. (16). In other words, if foreign competition is absent, the home country will optimize its own production and pollution by setting a strictly positive tax rate $t_{\mathrm{low}}^{\mathrm{BD}}$.

---

[8] We index this solution with 'BD' for **b**enevolent **d**ictator.

[9] As $t$* is strictly positive (cf. the foreign equivalent of Eq. (16)) such a situation can only occur if $\alpha_{\mathrm{E}}\beta > 1$.

The domestic reaction function is thus defined by:

$$t^{\mathrm{BD}} = \begin{cases} t_{\mathrm{low}}^{\mathrm{BD}}, & \text{for } t^* \geq p \\ \tilde{t}^{\mathrm{BD}}, & \text{for } t_{\mathrm{lb}}^{\mathrm{BD}*} < t^* < p \\ p, & \text{for } t^* \leq t_{\mathrm{lb}}^{\mathrm{BD}*} \end{cases} \tag{17}$$

The foreign reaction function is isomorphous.

### 3.2 Equilibrium

The three types of possible equilibria are depicted in Figure 1.[10] Either one country sets a prohibitive tax rate and the other country sets its best response tax rate ($t_{\mathrm{low}}^{\mathrm{BD}}$ or $t_{\mathrm{low}}^{\mathrm{BD}*}$, respectively) or both countries set non-prohibitive tax rates thereby creating an inner solution with both countries producing the polluting good.

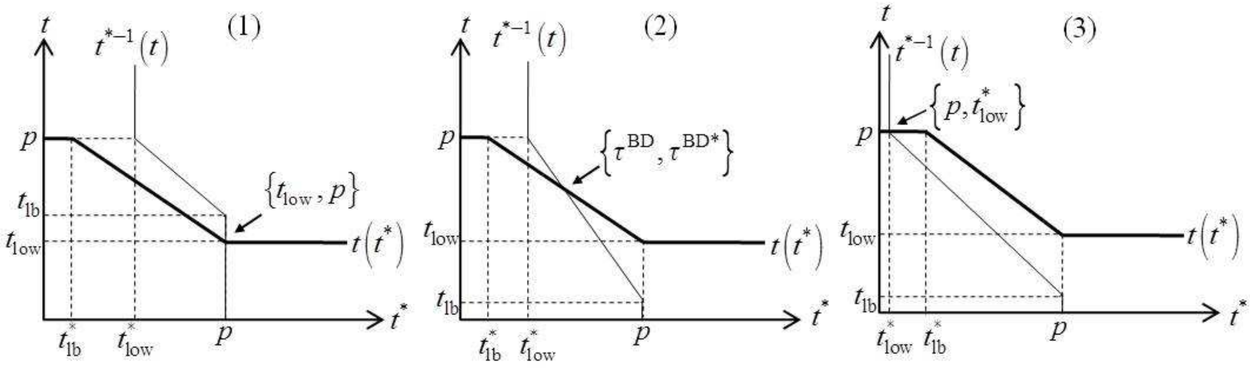



Figure 1: The Equilibrium with Benevolent Dictators

Panel (1) depicts the situation in which the foreign country sets its prohibitive tax rate and the home country reacts by setting its tax rate equal to $t_{\mathrm{low}}^{\mathrm{BD}}$. This situation requires that $t_{\mathrm{lb}}^{\mathrm{BD}*} > t_{\mathrm{low}}^{\mathrm{BD}}$,[11] which amounts to the condition $\alpha_{\mathrm{E}}{}^* - \alpha_{\mathrm{E}} > \frac{1}{\beta}$. Panel (3) depicts the opposite corner solution with the home country setting the prohibitive tax with $t_{\mathrm{lb}}^{\mathrm{BD}*} > t_{\mathrm{low}}^{\mathrm{BD}*}$ and $\alpha_{\mathrm{E}}{}^* - \alpha_{\mathrm{E}} > \frac{1}{\beta}$. In other words, corner solutions occur if the marginal damages from pollution differ strongly between countries – the country with the higher valuation of environmental quality introduces

[10] In Figure 1, we have assumed that $\alpha_{\mathrm{E}}\beta > 1$ and $\alpha_{\mathrm{E}}^*\beta > 1$. If for instance $\alpha_{\mathrm{E}}^*\beta < 1$, $t_{\mathrm{lb}}^{\mathrm{BD}}$ would be negative and therefore the foreign government would never set a prohibitive tax rate. Graphically, the foreign reaction curve would not have a kink at $t_{\mathrm{lb}}^{\mathrm{BD}*} = p$, but $t_{\mathrm{lb}}^{\mathrm{BD}*}(t^* = 0) < p$. The equilibrium depicted in panel (1) could not exist.

[11] This condition is intuitive and follows from the definition of these threshold values: Only if the foreign country sets a prohibitive tax rate, which requires $t < t_{\mathrm{lb}}^{\mathrm{BD}}$, will the home country set its best response at $t_{\mathrm{low}}^{\mathrm{BD}}$. For a corner solution to exist and the reaction functions to intersect at $t^* = p$, it is required that $t_{\mathrm{low}}^{\mathrm{BD}} < t_{\mathrm{lb}}^{\mathrm{BD}}$. Otherwise an interior solution would result.

a prohibitive tax while the other keeps producing with a tax rate equal to the marginal damage from production.

Panel (2) shows the case in which both countries produce, create externalities from pollution, and tax their production. It requires that the two countries do not differ too much in their marginal damage from pollution: $|\alpha_{\mathrm{E}} - \alpha_{\mathrm{E}}{}^{*}| < \frac{1}{\beta}$. The more polluting production is (i.e. the larger $\beta$), the more similar the valuations of environment need to be for an interior solution.

The interior equilibrium $\left\{\tilde{T}^{\mathrm{BD}}, \tilde{T}^{\mathrm{BD}*}\right\}$ is given by the intersection of the reaction functions $\tilde{t}^{\mathrm{BD}}$ (from Eq. (16)) and $\tilde{t}^{\mathrm{BD}*}$, as shown in Figure 1, Panel (2). The domestic tax rate amounts to

$$\tilde{T}^{\mathrm{BD}} = \frac{2\alpha_{\mathrm{E}}\beta p}{1 + \beta(\alpha_{\mathrm{E}} + \alpha_{\mathrm{E}}{}^{*})}. \tag{18}$$

$\tilde{T}^{\mathrm{BD}*}$ is calculated accordingly.

We can now define the equilibrium for the benevolent dictator setting in the following Proposition.

**Proposition 1**

The equilibrium tax rates on production for two welfare-maximizing governments, $T^{\mathrm{BD}}, T^{\mathrm{BD}*}$, are given by

$$\left\{T^{\mathrm{BD}}, T^{\mathrm{BD}*}\right\} = \begin{cases} (1)\left\{t_{\mathrm{low}}^{\mathrm{BD}}, p\right\}, & \text{for } \alpha_{\mathrm{E}}{}^{*} - \alpha_{\mathrm{E}} > \dfrac{1}{\beta} \\ (2)\left\{\tilde{T}^{\mathrm{BD}}, \tilde{T}^{\mathrm{BD}*}\right\}, & \text{for } |\alpha_{\mathrm{E}} - \alpha_{\mathrm{E}}{}^{*}| < \dfrac{1}{\beta} \\ (3)\left\{p, t_{\mathrm{low}}^{\mathrm{BD}*}\right\}, & \text{for } \alpha_{\mathrm{E}} - \alpha_{\mathrm{E}}{}^{*} > \dfrac{1}{\beta}\ , \end{cases} \tag{19}$$

with $t_{\mathrm{low}}^{\mathrm{BD}} = \frac{\alpha_{\mathrm{E}}\beta p}{1+\alpha_{\mathrm{E}}\beta}$ and $\tilde{T}^{\mathrm{BD}}$ defined by Eq. (18).

Equilibrium tax rates are strictly positive.

The equilibrium is unique and stable.

*Proof*: See Appendix 1. □

### 3.3 Comparative Statics

Comparative static effects of variations in all exogenous variables ($\alpha_{\mathrm{E}}$, $\alpha_{\mathrm{E}}{}^{*}$, $\beta$, and $p$) are straightforward. An increase in $\alpha_{\mathrm{E}}$ raises $t_{\mathrm{lb}}^{\mathrm{BD}*}$ and $t_{\mathrm{low}}^{\mathrm{BD}}$, and it shifts the domestic reaction curve (bold line) upwards while the foreign reaction curve (thin line) is unaffected. [12] This is depicted in Panel 1 of Figure 2 for an inner solution. The equilibrium shifts from A to B.

The equilibrium shifts to the Northwest: as a consequence, for large enough variations in $\alpha_E$, an interior equilibrium may change into a corner equilibrium with the home country setting a prohibitive tax rate. Conversely a corner solution with the foreign country setting a prohibitive tax rate may turn into an inner solution. This follows immediately from differentiating Eq. (18) w.r.t. $\alpha_E$.[13] An equilibrium in which the domestic tax rate was already prohibitive remains unaffected.

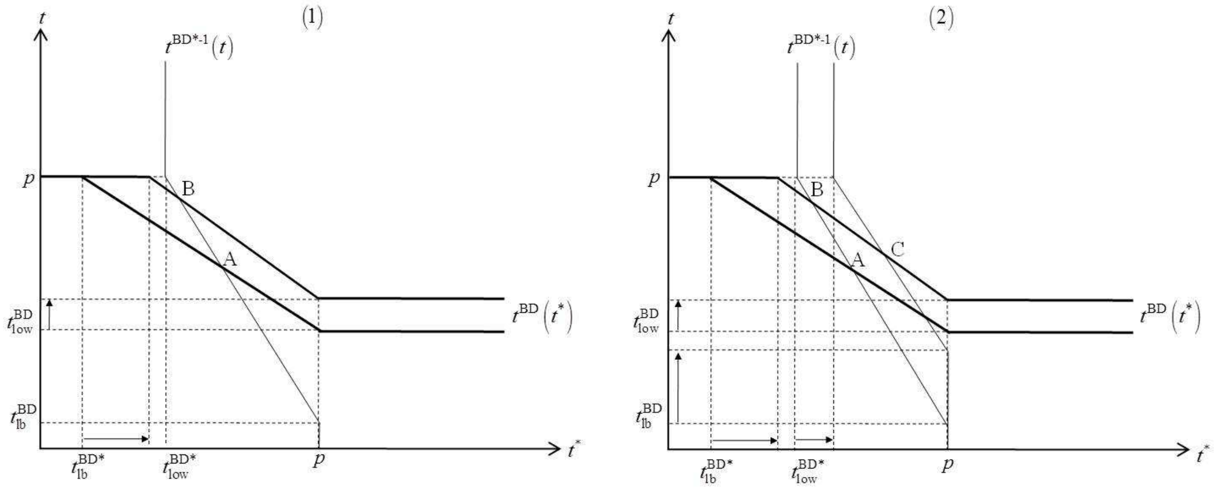



Figure 2: An increase in $\alpha_E$ (panel 1) and in $\beta$ (panel 2)

An increase in the damage parameter $\beta$ raises $t_{lb}^{BD*}, t_{lb}^{BD}, t_{low}^{BD*}, t_{low}^{BD}$ .[14] This is shown in Panel 2 of Figure 2. Qualitatively, the domestic reaction curve shifts, as depicted in Panel 1. The main difference is that the foreign reaction curve shifts in the same way. This is intuitive: As the marginal environmental damage increases in both countries, countries raise their best response tax rates. As a result, the range of the corner solutions is enlarged on both ends; for the interior solution the resulting new equilibrium (at point C) is characterized by higher foreign and domestic tax rates. Analytically, this can be seen by differentiating (18) w.r.t. $\beta$.

# 4. Interest Based Approach

We now employ a more realistic setting and assume that governments are self-interested. More specifically, we assume a common agency framework (Bernheim and Whinston 1986,

---

[12] This is seen from $\frac{\partial t_{lb}^{BD*}}{\partial \alpha_E} = \frac{p}{\alpha_E{}^2 \beta} > 0$ and $\frac{\partial t_{low}^{BD}}{\partial \alpha_E} = \frac{p\beta}{[1+\alpha_E \beta]^2} > 0$. The increase of $t_{lb}^{BD*}$ is not proportional to that of $t_{low}^{BD}$, which implies that the slope of the reaction function changes.

[13] $\frac{\partial \tilde{T}^{BD}}{\partial \alpha_E} = \frac{2p\beta(1+\alpha_E{}^*\beta)}{\left(1+\beta(\alpha_E+\alpha_E{}^*)\right)^2} > 0$. Analogously, $\frac{\partial \tilde{T}^{BD}}{\partial \beta} > 0$, $\frac{\partial \tilde{T}^{BD}}{\partial p} > 0$, and $\frac{\partial \tilde{T}^{BD}}{\partial \alpha_E{}^*} < 0$.

[14] For instance, $\frac{\partial t_{lb}^{BD*}}{\partial \beta} = \frac{p}{\alpha_E(\beta)^2} > 0$ and $\frac{\partial t_{low}^{BD}}{\partial \beta} = \frac{\alpha_E p}{[1+\alpha_E \beta]^2} > 0$.

Grossman and Helpman 1994), in which governments maximize a political support function. The political support function is a weighted sum of social welfare and contributions offered by political interest groups.

### 4.1. The Political Setting

We assume that individuals with similar interests form national lobby groups in both countries and offer campaign contributions to their governments. Environmentalists form environmental lobby groups, industrialists form industry lobby groups while workers do not organize.[15] The underlying assumption is that workers are large in number and cannot overcome the free-riding problem described by Olson (1965).

Let *i* denote the type of lobby group, *E* for environmental and *I* for industry. $\alpha_i$ defines the fraction of the population that are members of lobby group *i*. Each lobby group offers campaign contribution schedules to their country's government denoted by $\Lambda_i(t)$.[16] Their intention is to influence the government's choice of environmental policy: These contribution schedules are contingent on the pollution tax rate selected by the government and reward the policy choice. Each lobby group's strategy consists of a continuous function $\Lambda_i: T \rightarrow \mathbb{R}_+$. Lobby groups offer a monetary payment $\Lambda_i$ to the government for choosing the tax rate $t \in T, T \in \mathbb{R}$. All contribution schedules are assumed to be non-negative and differentiable around the equilibrium point.[17] Lobby groups at home and abroad act independently from each other. The foreign pollution tax rate will be taken as given when lobby groups decide on their lobby schedules.[18]

Faced with the lobby contribution offers, the incumbent government selects a pollution tax rate with the objective to maximize its own political welfare, i.e. the probability of re-election. The government's objective function is a weighted sum of average welfare and lobby contributions. Average welfare is important to the government because chances for re-election depend on the well-being of the general voter or citizen. Contributions matter as they can be

[15] Note that if workers also formed a lobby group, and hence all individuals were organized in lobby groups, the tax rates of the political game would equal the benevolent dictator tax rates.

[16] The offers of campaign contributions are neither formal contracts nor do they have to be explicitly announced. We only assume that governments know that there is an implicit relationship between their chosen tax rates and the contributions from lobby groups which they expect to receive. Campaign contributions should be interpreted broadly as campaign funds, support demonstrations, or bribes, since lobby groups employ different strategies to influence governments, see Conconi (2003).

[17] Contribution schedules are not differentiable if the assumption of non-negativity becomes binding, that is, when the government chooses a tax rate from which follows that $\Lambda_i = 0$.

[18] We follow Grossman and Helpman (1995) who argue that contribution schedules cannot be observed from abroad and thus have no influence on the decisions made abroad. We may then assume that lobby groups take foreign policies as given, and decide upon their contribution schedules before the actual foreign tax rate is set. We also disregard the possibility that interest groups lobby across the border. For such an analysis cf. Hillman and Ursprung (1988), Aidt (2005).

used to influence imperfectly informed voters, e.g. through political advertising (Grossman and Helpman 1995). The home government's objective function is defined as:

$$\upsilon = \sum_{i \in L} \Lambda_i + a\Omega_{\mathrm{A}} \tag{20}$$

where $L$ is the set of lobby groups, and $a \geq 0$ is the exogenously given weight that the government places on aggregate social welfare relative to campaign contributions.[19] The government weighs the political value of lobbying funds (in terms of votes gained) against their political cost associated with the loss of welfare in the determination of the weighting parameter $a$.

### 4.2. The Formation of Environmental Policy

The game between the incumbent government and the lobby groups has two stages. In the first stage, the lobby groups simultaneously offer their campaign contribution schedules, taking the other lobby group's strategy as given. In the second stage, the two governments select their tax rates, which maximize their objective functions $v$ and $v^*$ given the strategic interaction with the other government, and collect the corresponding contribution from the lobby groups in their country.[20] The lobby groups offer contribution schedules anticipating the optimization calculus of their governments in the second stage.

### General Characterization of the Political Equilibrium

In the two country common agency setup, the equilibrium is characterized by governments setting tax rates that maximize their respective political support functions, taking the other country's policy and their national lobby groups' contribution schedules as given. The lobby groups maximize their respective utilities, contingent on national policies, by offering feasible contribution schedules to their governments. They take the other government's policy and the contributions of the other lobby groups in their respective countries as given.

Applying Lemma 2 of Bernheim and Whinston (1986), or Proposition 1 in Grossman and Helpman (1994) to our setup, the equilibrium is characterized as follows.

---

[19] For an analysis that endogenizes the weight of social welfare for the political objective function cf. Fredriksson et al. (2005).

[20] It is assumed that lobby groups keep their promises and thus make the announced payments.

**Proposition 2**

$\left(\{\Lambda_i\}_{i\in L},\{\Lambda_i*\}_{i\in L},\left\{T^{\mathrm{PG}},T^{\mathrm{PG}*}\right\}\right)$ is a Subgame Perfect Nash Equilibrium of the pollution tax game, if and only if:

(a) $\Lambda_i,\Lambda_i^*$ are feasible for all $i\in L$;

(b) $T^{\mathrm{PG}}$ maximizes $v$ on $T$, and $T^{\mathrm{PG}*}$ maximizes $v^*$ on $T^*$;

(c) $T^{\mathrm{PG}}$ maximizes $\Omega_j(t)-\Lambda_j(t)+v$ on $T$ for all $j\in L$, and $T^{\mathrm{PG}*}$ maximizes $\Omega_j{}^*(t^*)-\Lambda_j{}^*(t^*)+v^*$ on $T^*$ for all $j\in L^*$;

(d) For all $j\in L$ there exists a $t^{-j}\in T$ that maximizes $v$ on $T$ such that $\Lambda_j(t^{-j})=0$, and for all $j\in L^*$ there exists a $t^{-j*}\in T^*$ that maximizes $v^*$ on $T^*$ such that $\Lambda_j{}^*(t^{-j*})=0$.

A set of policies $\left\{T^{\mathrm{PG}},T^{\mathrm{PG}*}\right\}$ and the sets of contribution schedules $\{\Lambda_i\}_{i\in L},\{\Lambda_i*\}_{i\in L}$ are a subgame perfect Nash equilibrium if conditions (a) to (d) hold. Condition (a) stipulates that contribution schedules must be feasible, that is, they must be non-negative and no greater than the aggregate income available to the lobby group's members. Condition (b) ascertains that the governments set their pollution taxes $T^{\mathrm{PG}},T^{\mathrm{PG}*}$ to maximize their respective objective functions $\{v\}$ and $\{v^*\}$ taking the contribution schedules offered by their lobby groups and the other country's policy as given. Condition (c) stipulates that the equilibrium tax rate must maximize the joint welfare of the government and each of the national lobby groups, given the contribution schedule offered by the other lobby group. In other words, no lobby group $j$ has a feasible strategy other than the equilibrium strategy that would lead to an increase in the joint surplus of the government and the lobby group, of which it could appropriate a share. Condition (d) requires that for every lobby group $j$, a tax policy $t^{-j}$ exists that gives the government the same utility as the equilibrium tax rate $T^{\mathrm{PG}}$, if the lobby group $j$ does not contribute. If no such $t^{-j}$ existed, lobby group $j$ could increase its welfare by lowering its campaign bid without changing the government's choice of tax policy. This would leave lobby group $j$ better off and can thus not be possible in equilibrium (Bernheim and Whinston 1986). Conditions (c) and (d) ensure that the lobbying schedule is optimal.

**Political-economic Reaction Functions**

Next we derive the home government's reaction function from Proposition 2. Conditions (b) and (c) characterize the optimization calculus of the government.

Condition (b)

$$\sum_{i\in L}\frac{\partial\Lambda_i}{\partial t}+a\frac{\partial\Omega_A}{\partial t}\overset{!}{=}0 \qquad (21)$$

and Condition (c)

$$\frac{\partial \Omega_j}{\partial t} - \frac{\partial \Lambda_j}{\partial t} + \frac{\partial v}{\partial t} \overset{!}{=} 0 \quad \text{for all } j \in L \tag{22}$$

Eqs. (21) and (22) imply that, in equilibrium, each lobby group sets its contribution schedule such that, the marginal utility from a change in the tax rate equals its marginal change in contribution. Thus each lobby group's marginal net utility is zero in equilibrium.

$$\frac{\partial \Omega_i}{\partial t} = \frac{\partial \Lambda_i}{\partial t} \qquad \text{for all } i \in L \tag{23}$$

Substituting Eq. (23) into Eq. (21) we obtain the equilibrium characterization:

$$\sum_{i \in L} \frac{\partial \Omega_i}{\partial t} + a \frac{\partial \Omega_A}{\partial t} \overset{!}{=} 0 \tag{24}$$

Next we calculate $\frac{\partial \Omega_{\mathrm{E}}}{\partial t}$ and $\frac{\partial \Omega_{\mathrm{I}}}{\partial t}$ to derive the politically optimal tax rate as a function of the other country's tax rate and the parameters of the model. $\frac{\partial \Omega_{\mathrm{A}}}{\partial t}$ is given by Eq. (15). The lobby groups' marginal utilities w.r.t. the tax rate in Eq. (24) are calculated by substituting Eqs. (4), (5), and (6) in Eqs. (10) and (11), and differentiating them with respect to *t*:

$$\frac{\partial \Omega_{\mathrm{E}}}{\partial t} = \frac{1}{2} \alpha_{\mathrm{E}} \left[ \beta \left( 2p - t - t^* \right) + 2 \left( \frac{p}{2} - t \right) \right] \tag{25}$$

and

$$\frac{\partial \Omega_{\mathrm{I}}}{\partial t} = \alpha_{\mathrm{I}} \left( \frac{p}{2} - t \right) - \frac{1}{2} (p - t) \tag{26}$$

Environmentalists' marginal utility with respect to the home tax rate can have either sign (cf. Eq. (25)). There are two relevant effects. First, when the domestic tax rate increases, home production of *x* decreases, and hence pollution decreases. Second, total tax revenue changes with *t* and thus the share redistributed to environmentalists. It increases if $\frac{p}{2} - t > 0$ and decreases otherwise (cf. Eq. (6)). If the environmentalists' revenue share increases, their marginal utility with respect to the home tax rate is unambiguously positive. Otherwise, the loss in tax revenue may outweigh the effect of reduced pollution – making $\frac{\partial \Omega_{\mathrm{E}}}{\partial t}$ negative.[21]

Eq. (26) shows that industrialists' marginal utility from an increase in *t* is strictly negative: sector specific income $\Pi$ decreases; tax revenue may increase or decrease (see above), but an increase can only partially compensate industrialists for the decline in profits as tax revenue is distributed among all members of the society.

We calculate the reaction function of the home country by substituting Eqs. (15), (25), and (26) in Eq. (24), and solving it for *t*. This yields:

---

[21] This may happen only for small $\beta$ because a further increase in the tax rate reduces pollution only negligibly but may reduce the tax revenue significantly as the tax base diminishes.

$$\tilde{t}^{\mathrm{PG}} = \frac{\alpha_{\mathrm{E}}\beta(a+1)(2p-t^{*}) - \alpha_{\mathrm{W}}p}{(a+1)(\alpha_{\mathrm{E}}\beta+1) - 2\alpha_{\mathrm{W}}} \tag{27}$$

### 4.3 The Political-Economic Equilibrium

Eq. (27) is derived from the first order condition for a maximum of the political support function conditional on the value of the foreign tax rate. As in the benevolent dictator case, the reaction function is linear in the foreign tax rate. The second order condition for an interior maximum, i.e. $\frac{\partial^2 v}{\partial t^2} < 0$, requires that

$$(a+1)(\alpha_{\mathrm{E}}\beta+1) - 2\alpha_{\mathrm{W}} > 0\,. \tag{28}$$

If condition (28) was violated, the interior solution given by (27) would characterize a minimum and hence a corner solution would be optimal. Industrialists' marginal utility would increase faster as $t$ is lowered below $\tilde{t}^{\mathrm{PG}}$ than the sum of the weighted marginal welfare, and the environmentalists' marginal utility would decrease.[22] It would be optimal for the government to reduce the tax rate to the minimal amount possible.[23]

This corner solution is a degenerate case as it is hard to conceive that society directs all its resources from all members of the society by an 'infinitely' negative tax on production towards the industrialist sector only to increase output, profits and environmental degradation to the maximum extent possible.[24] This is not what we observe and it would require unrealistic parameter values. We thus exclude this uninteresting case by the following assumption:

**Assumption 1:** $\frac{1}{2}(a+1)(\alpha_{\mathrm{E}}\beta+1) > \alpha_{\mathrm{W}}$ and $\frac{1}{2}(a^{*}+1)(\alpha_{\mathrm{E}}{}^{*}\beta+1) > \alpha_{\mathrm{W}}{}^{*}$

Assumption 1 guarantees a stable interior maximum of the political support function conditional on the value of the foreign tax rate and stable Nash equilibria as shown below.

The sign of the tax rate in Eq. (27) is ambiguous — while the denominator is positive under Assumption 1, the numerator can be positive or negative. Thus in contrast to the

[22] This is intuitive as Eq. (28) states that the absolute values of the second derivative of the environmentalists' utility function $\Omega_{\mathrm{E},\,tt} = -\alpha_{\mathrm{E}} - \frac{1}{2}\alpha_{\mathrm{E}}\beta$ plus the weighted second derivative of the welfare function $a\Omega_{\mathrm{A},\,tt} = -\,a(\frac{1}{2}+\frac{1}{2}\alpha_{\mathrm{E}}\beta)$ exceed the value of the second derivative of the industrialists' utility function $\Omega_{\mathrm{I},\,tt} = \frac{1}{2} - \alpha_{\mathrm{I}}$.

[23] While the effective tax rate is bounded from above by the value of the price, beyond which production is zero and thus a further increase would be inconsequential, a lower bound exists only to the extent that the negative tax would use up all resources from the society and redirect it to the industrial production. When comparing $\upsilon\,(t=p)$ with $\upsilon\,(t \rightarrow -\infty)$ it is immediately clear that the value of the former corner solution falls short of the latter. [Strictly speaking the latter is not an infinite, but a finite subsidy with a rate implicitly defined by the gross resources of all groups.]

[24] Such a degenerate case could occur only if the environmentalists had little political weight, the pollution damage from production was small, the value of welfare consideration of the government's calculus was low, and if the number of industrialists was small ($\alpha_{\mathrm{E}}$, $\beta$, $a$, $\alpha_{\mathrm{I}}$ were small).

benevolent dictator case the reaction function of a political support maximizing government can take on negative values. Given Assumption 1, the reaction function is downward sloping in the foreign tax rate, as in the Benevolent Dictator case.

Eq. (27) defines the reaction function only for positive production of both firms, not for corner solutions. Analogous to Section 3, we determine the optimal domestic tax rate for zero foreign production, $t_{\mathrm{low}}^{\mathrm{PG}}$, by setting $t^* = p$ in Eq. (27). This yields

$$t_{\mathrm{low}}^{\mathrm{PG}} = \frac{\alpha_{\mathrm{E}}\beta(a+1)p - \alpha_{\mathrm{W}}p}{(a+1)(\alpha_{\mathrm{E}}\beta+1) - 2\alpha_{\mathrm{W}}}. \tag{29}$$

The denominator is positive under Assumption 1, the numerator can have either sign: Contrary to the benevolent dictator case, $t_{\mathrm{low}}^{\mathrm{PG}}$ can take on negative values. A comparison of (27) and (29) shows that $t_{\mathrm{low}}^{\mathrm{PG}} \leq \tilde{t}^{\mathrm{PG}}$.

Next we determine the foreign tax rate $t_{\mathrm{lb}}^{\mathrm{PG}*}$, below which it is optimal for the home country to introduce a prohibitive tax rate $t \geq p$.

$$t_{\mathrm{lb}}^{\mathrm{PG}*} = p\left(\frac{(a+1)\,[\alpha_{\mathrm{E}}\beta - 1] + \alpha_{\mathrm{W}}}{\alpha_{\mathrm{E}}\beta(a+1)}\right) \tag{30}$$

Note that $t_{\mathrm{lb}}^{\mathrm{PG}*}$ may be positive or negative. The reaction function is thus defined by

$$t^{\mathrm{PG}} = \begin{cases} t_{\mathrm{low}}^{\mathrm{PG}}, & \text{for } t^* \geq p \\ \tilde{t}^{\mathrm{PG}}, & \text{for } t_{\mathrm{lb}}^{\mathrm{PG}*} < t^* < p \\ p, & \text{for } t^* \leq t_{\mathrm{lb}}^{\mathrm{PG}*}, \end{cases} \tag{31}$$

where $\tilde{t}^{\mathrm{PG}}$ is defined by Eq. (27). As $t_{\mathrm{low}}^{\mathrm{PG}}$ and $t_{\mathrm{lb}}^{\mathrm{PG}*}$ can have either sign and are not bounded from below, the 'interior' reaction function $\tilde{t}^{\mathrm{PG}}$ may be only in the first quadrant, in the fourth, first, and second, or not in the first quadrant at all. Three possible reaction curves are depicted in Figure 3 below.[25]

[25] The reaction functions are depicted for values $\beta = 6$, $\alpha_{\mathrm{I}} = 0.06$, $p = 5$, $l = 1$, $a = 1$ and differ in the value for $\alpha_{\mathrm{E}}$ which takes on the values 0.03, 0.05, and 0.1 for the reaction curves a, b, and c, respectively.

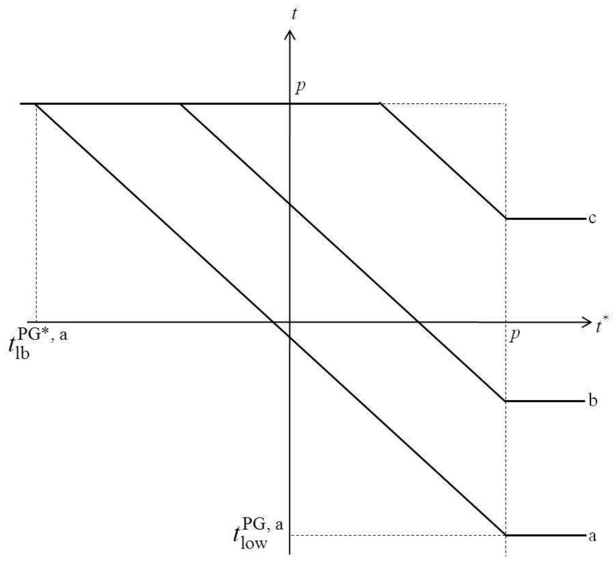



Figure 3: Domestic reaction curves for the political game

Next we analyze the possible equilibria. The slopes of the 'interior' domestic reaction function $\tilde{t}^{\mathrm{PG}}$ and the inverse of the foreign reaction function $\tilde{t}^{\mathrm{PG}*-1}$ are

$$\frac{\partial t}{\partial t^*} = -\frac{\alpha_{\mathrm{E}}\beta(a+1)}{(a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}}$$

and (32)

$$\frac{\partial t^{*-1}}{\partial t} = -\frac{(a^*+1)(\alpha_{\mathrm{E}}{}^*\beta+1)-2\alpha_{\mathrm{W}}{}^*}{\alpha_{\mathrm{E}}{}^*\beta(a^*+1)}.$$

By Assumption 1, both reaction curves are downward sloping. Eq. (32) shows that the inverse of the foreign reaction curve can be flatter or steeper than the domestic reaction curve, which implies the possibility of unstable equilibria.

**Stable equilibria**

We first analyze the case in which $\frac{\partial t}{\partial t^*} > \frac{\partial t^{*-1}}{\partial t}$; i.e. the domestic reaction function is flatter than the inverse of the foreign reaction function. This is depicted in Figure 4. The line in boldface depicts a selected domestic reaction function as shown in Figure 3.[26] The thin lines show three possible inverse of the foreign reaction curve, which results in three different – stable – equilibria. In equilibrium A, the domestic country sets a prohibitive tax rate $t = p$ and the foreign country sets its best response, which is $t_{\mathrm{low}}^*$. Equilibrium C is the mirror case in which the foreign government denies production in its country and the home country sets $t = t_{\mathrm{low}}^{\mathrm{PG}}$. These corner solutions A and C correspond to the panels 3 and 1 in Figure 1 of Section 3. Yet,

[26] For the sake of clarity we selected only one domestic reaction function, but it is clear from Figure 3, that depending on parameter values the domestic reaction function could lie entirely outside the first quadrant or the downward sloping part entirely inside the first quadrant. The same is true for the foreign reaction function so that a resulting interior equilibrium could lie anywhere in the policy space.

while in the benevolent dictator case both tax rates are unambiguously positive, tax rates may (or may not) be negative in the political game. The interior solution B is characterized by both countries producing finite amounts of the polluting good. In Figure 4 this solution is depicted with positive tax rates for both countries; however home and foreign reaction curves could be positioned very differently in the policy space $(t, t^*)$ – as shown in Figure 3 for the domestic reaction function – so that any combination between taxes and subsidies is possible in the equilibrium. In other words, it is possible that both countries subsidize the production of the polluting good or that one country taxes the negative externality while the other country subsidizes it.

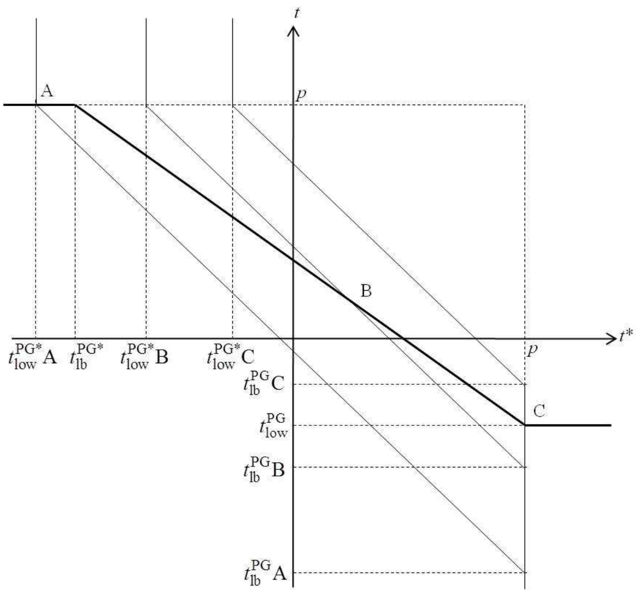



Figure 4: Stable equilibria in the political game

The conditions for corner solutions versus inner solutions can be seen in Figure 4 by comparing the values of $t_{low}^{PG}, t_{lb}^{PG}, t_{low}^{PG*}$ and $t_{lb}^{PG*}$:[27]

$$\begin{aligned} X = 0, X^* > 0 \quad &\Leftrightarrow \quad t_{low}^{PG*} < t_{lb}^{PG*} \\ X, X^* > 0 \quad &\Leftrightarrow \quad t_{low}^{PG*} > t_{lb}^{PG*} \text{ and } t_{low}^{PG} > t_{lb}^{PG} \\ X > 0, X^* = 0 \quad &\Leftrightarrow \quad t_{low}^{PG} < t_{lb}^{PG} \end{aligned} \tag{33}$$

The three conditions in (33) describe the equilibria A, B, C in Figure 4. For instance, for equilibrium A: $t_{low}^{PG*} < t_{lb}^{PG*}$. The first line of (33) corresponds with the tax rates $t = p, t^* = t_{low}^{PG*}$, the second line with $(\tilde{t}^{PG}, \tilde{t}^{PG*})$, and the third line with $t = t_{low}^{PG}, t^* = p$ .

We analyze the political-economic determinants for the equilibria A, B, and C and their position by demonstrating how the domestic reaction function shifts in response to changes in $\alpha_E$, $\alpha_I$, $\beta$ and $a$. All derivations are relegated to Appendix 2. An increase in $\alpha_E$ shifts the domestic reaction function to the Northeast and increases $t_{lb}^{PG*}$ and $t_{low}^{PG}$. An increasing $\beta$ shifts

[27] Note that for $\frac{\partial t}{\partial t^*} > \frac{\partial t^{*-1}}{\partial t}$ the condition $t_{low}^{PG*} < t_{lb}^{PG*}$ implies that $t_{low}^{PG} > t_{lb}^{PG}$.

the domestic reaction curve upwards in the same way; however the foreign reaction curve shifts as well in the described manner so that a new inner equilibrium, if it exists, must lie to the Northeast of the old inner equilibrium. Thus both shifts are similar to the benevolent dictator case depicted in

Figure 2 (with the exception that they are not confined to the first quadrant). Increases in $\alpha_{\mathrm{I}}$ and $a$ reduce $t_{\mathrm{lb}}^{\mathrm{PG}*}$ and make the reaction curve flatter; they increase $t_{\mathrm{low}}^{\mathrm{PG}}$ if $\alpha_{\mathrm{E}}\beta < 1$ and decreases it otherwise.[28] The new reaction curve may either be entirely below the old reaction curve or intersect with it.

To sum up, subsidies in both countries occur if the industry lobbies are relatively strong and if damages are relatively small (small $\beta$). A subsidy in one country may also occur if the political distortion in this country is much larger than in the other country ($a \ll a^*$). Equilibria with positive tax rates are more likely the higher the environmental damages and the more powerful the environmental lobbies in both countries are (i.e., the higher $\alpha_{\mathrm{E}}$ and $\alpha_{\mathrm{E}}{}^{*}$). The government's reaction to an increased size of the industrial lobby group depends on the entire political system. The domestic equilibrium tax rates will decrease if $\alpha_{\mathrm{E}}\beta > 1$ but may increase if $\alpha_{\mathrm{E}}\beta < 1$. If the domestic tax rate decreases, the foreign tax rate will increase.

**Unstable equilibria**

We now turn to the case of the inverse of the foreign reaction curve being flatter than the domestic reaction curve, i.e. $\frac{\partial t}{\partial t^*} < \frac{\partial t^{*-1}}{\partial t}$. This condition implies that both countries react relatively strongly in their tax setting to changes in the other country's tax rate. It occurs if the political distortion is very strong, either in favor of the industry lobby or in favor of the environmental lobby. A case of unstable equilibria is depicted in Figure 5. Again the line in boldface depicts the domestic reaction function; the thin line represents the foreign reaction function.

[28] This follows straightforwardly from differentiating (29) and (30) w.r.t. $\alpha_{\mathrm{I}}$ and $a$.

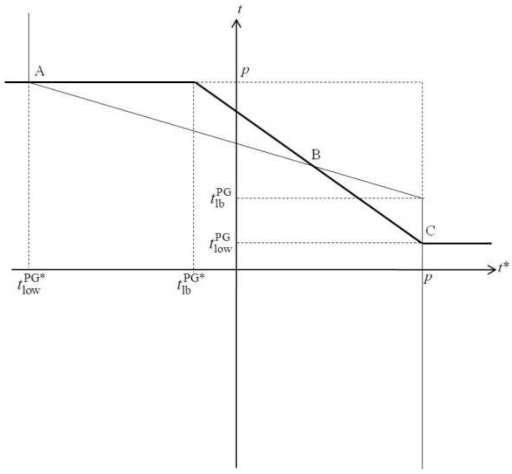



Figure 5: Unstable equilibria in the political game

In this case the reaction curves intersect three times with intersection B representing an unstable equilibrium and points A and C representing stable corner solutions. That is, if countries react relatively strongly to their opponent's tax setting, initial conditions or coincidence decide about which country will produce the polluting good, possibly under heavy subsidization, and which country will not produce the good at all. In that case it is no longer the political-economic characteristics of the two countries alone which decide about the pattern of production and pollution, but any factor that happens to tip the unstable equilibrium to the left or the right of point B with the consequences being most radical. A government may heavily subsidize the production of the polluting good, which it would have banned under only slightly different circumstances. Overall, the possibility of unstable outcomes makes corner solutions more likely.

For this case the equilibria can be characterized as follows:

$$\begin{aligned}
&t_{\mathrm{low}}^{\mathrm{PG*}} < t_{\mathrm{lb}}^{\mathrm{PG*}} \text{ and } t_{\mathrm{low}}^{\mathrm{PG}} > t_{\mathrm{lb}}^{\mathrm{PG}} &\Rightarrow\quad & X = 0,\ X^* > 0\\
&t_{\mathrm{low}}^{\mathrm{PG*}} < t_{\mathrm{lb}}^{\mathrm{PG*}} \text{ and } t_{\mathrm{low}}^{\mathrm{PG}} < t_{\mathrm{lb}}^{\mathrm{PG}} &\Rightarrow\quad & X = 0, X^* > 0 \text{ or } X, X^* > 0 \text{ or } X > 0, X^* = 0\\
&t_{\mathrm{low}}^{\mathrm{PG*}} > t_{\mathrm{lb}}^{\mathrm{PG*}} \text{ and } t_{\mathrm{low}}^{\mathrm{PG}} < t_{\mathrm{lb}}^{\mathrm{PG}} &\Rightarrow\quad & X > 0,\ X^* = 0
\end{aligned} \tag{34}$$

The second line refers to the situation depicted in Figure 5. The comparative static properties are the same as described above. Note that increases in $a$ and $\alpha_{\mathrm{I}}$ will make the reaction curve flatter, which may turn an unstable equilibrium into a stable one.

We now turn to the derivation of the equilibrium. For an interior political-economic equilibrium $(\tilde{T}^{\mathrm{PG}}, \tilde{T}^{\mathrm{PG*}})$ the 'interior' reaction functions $\tilde{t}^{\mathrm{PG}}$ and $\tilde{t}^{\mathrm{PG*}}$ need to intersect. From Eq. (27) and its foreign equivalent follows

$$\tilde{T}^{\mathrm{PG}} = p\{1 + \frac{\alpha_{\mathrm{W}}^*[(a+1)(\alpha_{\mathrm{E}}\beta+1) - 2\alpha_{\mathrm{W}}] + (a^*+1)\left[\alpha_{\mathrm{W}}(2-\alpha_{\mathrm{E}}^*\beta) - (a+1)\left(1+\beta(\alpha_{\mathrm{E}}-\alpha_{\mathrm{E}}^*)\right)\right]}{(a+1)(\alpha_{\mathrm{E}}\beta+1)(a^*+1-2\alpha_{\mathrm{W}}) + (a^*+1)(\alpha_{\mathrm{E}}^*\beta+1)(a+1-2\alpha_{\mathrm{W}}) + 4\alpha_{\mathrm{W}}\alpha_{\mathrm{W}}^* - (a+1)(a^*+1)}\}. \tag{35}$$

We summarize our first main result in the following Proposition.

**Proposition 3:**

(i) The equilibrium tax rates on production for two political support-maximizing governments, $T^{\mathrm{PG}}, T^{\mathrm{PG}*}$, are given by

| | | |
|---|---|---|
| (1) | $\tilde{T}^{\mathrm{PG}}, \tilde{T}^{\mathrm{PG}*}$ | for $t_{\mathrm{low}}^{\mathrm{PG}*} > t_{\mathrm{lb}}^{\mathrm{PG}*}$ and $t_{\mathrm{low}}^{\mathrm{PG}} > t_{\mathrm{lb}}^{\mathrm{PG}}$ |
| (2) | $p, t_{\mathrm{low}}^{\mathrm{PG}*}$ | for $t_{\mathrm{low}}^{\mathrm{PG}*} < t_{\mathrm{lb}}^{\mathrm{PG}*}$ and $t_{\mathrm{low}}^{\mathrm{PG}} > t_{\mathrm{lb}}^{\mathrm{PG}}$ |
| (3) | $t_{\mathrm{low}}^{\mathrm{PG}}, p$ | for $t_{\mathrm{low}}^{\mathrm{PG}*} > t_{\mathrm{lb}}^{\mathrm{PG}*}$ and $t_{\mathrm{low}}^{\mathrm{PG}} < t_{\mathrm{lb}}^{\mathrm{PG}}$ |
| (4) | Multiple equilibria | for $t_{\mathrm{low}}^{\mathrm{PG}*} < t_{\mathrm{lb}}^{\mathrm{PG}*}$ and $t_{\mathrm{low}}^{\mathrm{PG}} < t_{\mathrm{lb}}^{\mathrm{PG}}$ |

where $t_{\mathrm{low}}^{\mathrm{PG}}$ is defined by Eq. (29) and $t_{\mathrm{lb}}^{\mathrm{PG}*}$ by Eq. (30).

(ii) Equilibria (1) to (3) are unique and stable. In situation (4) there exist two stable corner solutions with $(p, t_{\mathrm{low}}^{\mathrm{PG}*})$ and $(t_{\mathrm{low}}^{\mathrm{PG}}, p)$ and an unstable interior equilibrium with $(\tilde{T}^{\mathrm{PG}}, \tilde{T}^{\mathrm{PG}*})$.

(iii) Equilibrium tax rates $\tilde{T}^{\mathrm{PG}}, \tilde{T}^{\mathrm{PG}*}, t_{\mathrm{low}}^{\mathrm{PG}}, t_{\mathrm{low}}^{\mathrm{PG}*}$ may be positive or negative. They will be more likely to be positive if environmental lobbies are strong in both countries or the damages are large.

### 4.3 Comparison of the Political Game with the Benevolent Dictator Solutions

The conditions (1) to (4) in Proposition 3 can be rewritten in terms of the parameter of the model, which allows an easier comparison to the benevolent dictator case as described in Eq. (19). Using Eqs. (29), (30) and the other country's equivalents, we can rewrite the conditions in Proposition 3 as:

| | | |
|---|---|---|
| (1) | $\tilde{T}^{\mathrm{PG}}, \tilde{T}^{\mathrm{PG}*}$ | $-\frac{1}{\beta}\left(1-\frac{\delta\alpha_{\mathrm{W}}}{a+1}\right) < \alpha_{\mathrm{E}}\delta - \alpha_{\mathrm{E}}{}^{*}\delta^{*} < \frac{1}{\beta}\left(1-\frac{\delta^{*}\alpha_{\mathrm{W}}{}^{*}}{a^{*}+1}\right)$ |
| (2) | $p, t_{\mathrm{low}}^{\mathrm{PG}*}$ | $\frac{1}{\beta}\left(1-\frac{\delta^{*}\alpha_{\mathrm{W}}{}^{*}}{a^{*}+1}\right) < \alpha_{\mathrm{E}}\delta - \alpha_{\mathrm{E}}{}^{*}\delta^{*}$ and $\alpha_{\mathrm{E}}\delta - \alpha_{\mathrm{E}}{}^{*}\delta^{*} > -\frac{1}{\beta}\left(1-\frac{\delta\alpha_{\mathrm{W}}}{a+1}\right)$ |
| (3) | $t_{\mathrm{low}}^{\mathrm{PG}}, p$ | $\frac{1}{\beta}\left(1-\frac{\delta^{*}\alpha_{\mathrm{W}}{}^{*}}{a^{*}+1}\right) > \alpha_{\mathrm{E}}\delta - \alpha_{\mathrm{E}}{}^{*}\delta^{*}$ and $\alpha_{\mathrm{E}}\delta - \alpha_{\mathrm{E}}{}^{*}\delta^{*} < -\frac{1}{\beta}\left(1-\frac{\delta\alpha_{\mathrm{W}}}{a+1}\right)$ |
| (4) | Multiple equilibria | $-\frac{1}{\beta}\left(1-\frac{\delta\alpha_{\mathrm{W}}}{a+1}\right) > \alpha_{\mathrm{E}}\delta - \alpha_{\mathrm{E}}{}^{*}\delta^{*} > \frac{1}{\beta}\left(1-\frac{\delta^{*}\alpha_{\mathrm{W}}{}^{*}}{a^{*}+1}\right)$ |

$\delta$ is a measure of the political distortion in the home country. It is defined as: $\delta := \frac{a+1}{a+1-\alpha_{\mathrm{W}}} > 1$. The definition of $\delta^{*}$ is analogous.

For the inner solution in the benevolent dictator case, countries need to be not too dissimilar in the sense that the disutility from pollution must not differ by more than $\frac{1}{\beta}$ in absolute terms, i.e. $|\alpha_{\mathrm{E}} - \alpha_{\mathrm{E}}{}^{*}| < \frac{1}{\beta}$, as shown in Eq. (19). For the political game a comparable condition exists that bounds a weighted difference in the $\alpha_{\mathrm{E}}$s from above and below for a stable interior equilibrium; yet this difference now takes the political distortions into account and thus the lower and upper limits are different. Comparing the two sets of conditions in the above table and in Eq. (19) shows that it is impossible to determine in general whether the area with interior solutions is larger in the benevolent dictator case or in the political game.

Comparing $\tilde{T}^{PG}$ in Eq. (35) with $\tilde{T}^{BD}$ in Eq. (18) demonstrates that in principle the politically optimal tax rate can be smaller or larger than the tax rate that a benevolent dictator would set for the same economy; yet the deviation of the political-economic equilibrium from the benevolent dictators' solution depends systematically on the structure of the political, economic and ecological system. We illustrate this with the following three examples displayed in Table 1. Columns 2-4 give the values for the political economic equilibrium and for the benevolent dictators' equilibrium (in parentheses) for three different sets of parameter values. All equilibria are interior and stable and represent maxima of the governments' conditional objective functions (i.e., Assumption 1 is fulfilled and the domestic reaction curve is flatter than the inverse of the foreign reaction curve).[29]

| | Case (1) | Case (2) | Case (3) |
|---|---|---|---|
| *Parameter values* | $\alpha_E = \alpha_E^* = 0.1$ , $\alpha_I = \alpha_I^* = 0.1$ , p = 1, $l = l^* = 1$ | | |
| | $a = a^* = 1$ <br> $\beta = 1$ | $a = 10, a^* = 1$ <br> $\beta = 1$ | $a = a^* = 1$ <br> $\beta = 10$ |
| *Variables* | | | |
| Home tax rate PG (BD) | -0.5 (0.17) | 0.21 (0.17) | 0.73 (0.67) |
| Foreign tax rate PG (BD) | -0.5 (0.17) | -0.74 (0.17) | 0.73 (0.67) |
| Home production level PG (BD) | 0.75 (0.42) | 0.39 (0.42) | 0.14 (0.17) |
| Foreign production level PG (BD) | 0.75 (0.42) | 0.87 (0.42) | 0.14 (0.17) |
| Total Pollution PG (BD) | 2.25 (0.69) | 1.59 (0.69) | 0.74 (1.11) |
| Home Welfare PG (BD) | 0.96 (1.17) | 1.08 (1.17) | 1.04 (1.03) |
| Foreign Welfare PG (BD) | 0.96 (1.17) | 0.95 (1.17) | 1.04 (1.03) |
| Overall Welfare PG (BD) | 1.92 (2.34) | 2.03 (2.34) | 2.08 (2.06) |

*Note*: PG denotes values for the political game; BD denotes values for the benevolent dictator game. They are given in parentheses.

Table 1: Simulated equilibria in the benevolent dictator and the political game

In case 1, both countries are symmetric; the welfare maximizing governments would levy a 17 percent tax on the value of the output of their polluting firms. Under the same parameter

[29] It is straightforward to construct examples in which the political game results in corner solutions and the benevolent dictator game does not and vice versa. Results are available upon request.

values, the political-support maximizing governments, however, would subsidize production at a rate of 50 percent, with the consequence that production and pollution is significantly higher. In this case, the political-economic calculus leads to a sizeable deterioration in welfare and a strong increase in environmental degradation as the governments cater to the industrialist lobby group.

In case 2, both countries have the same parameter values as in case 1 with the exception that the domestic government places a much larger weight on welfare considerations in its political-economic calculus (i.e. the parameter $a$ is higher). The benevolent dictators' equilibrium is thus the same as in case 1, but the political equilibrium is qualitatively different: The domestic government now levies a tax rate, which is even higher than in the benevolent dictator case while the foreign country subsidizes production, yet more strongly than in case 1. The comparison between the two political economic equilibria in case 1 and 2 shows the interdependence of the political support maximizing governments' behaviors: As the domestic government is taxing the production of the polluting good thereby reducing the negative externality, the foreign government can increase its subsidy further thereby enhancing its political support. It is free-riding on the domestic government. Conversely, the domestic government anticipates such behavior and therefore taxes production more heavily than if the foreign government would tax its production as well. The welfare in the political equilibrium in case 2 is higher than in case 1, but it is lower than in the benevolent dictator case.

Case 3 is again completely symmetrical and has the same parameter values as case 1 except for the damage parameter $\beta$, which is now much higher. As a result, the benevolent dictators now tax production more heavily than in case 1 and the resulting welfare level is lower. This is intuitive. More striking, however, is the comparison between the political game and benevolent dictators' game. The political support maximizing governments tax production of the polluting good *more* heavily than a benevolent dictator would! As a result environmental degradation is lower and the welfare is higher in the political game.[30] The reason for this result is that the distortion created by the political-economic calculus — "too" high tax rates — now counteracts the distortion created by the strategic interaction of two benevolent dictators, who set tax rates on transboundary pollution too low. In our case, such a result occurs if there are strong environmental lobby groups (high $\beta$ and $\alpha_E$).[31]

We summarize these findings in the following corollary.

---

[30] Cf. Bhagwati (1982) for an analysis of countervailing distortions in a different context.

[31] However, the influence of the environmental lobby must not exceed a certain limit since the resulting equilibrium would be unstable leaving it to chance whether welfare would be higher under such conditions.

**Corollary 1:**

(i) The tax rates of the benevolent dictator can be either higher or lower than the tax rates set by political support maximizing governments.

(ii) The political game may result in higher or in lower welfare than the strategic interaction of non-cooperative benevolent dictators.

(iii) The welfare is lower if both tax rates are lower than in the benevolent dictator case. It may be higher if both tax rates exceed the benevolent dictators' tax rates.

(iv) Strong environmental lobby groups may increase welfare in the presence of transnational pollution.

(v) The political game leads to instability if either lobby group becomes too influential.

The intuition behind this result is that there are different forces that shift the equilibrium from the benevolent dictator solution to the political game solution. The direction of the political distortion depends on the relative strengths of the interest groups (and the value of $a$). For instance, if $\alpha_E$ and $\beta$ are high, an increase in the tax rate reduces production and thus profits, but translates into a large reduction in disutility from pollution. Thus, the environmentalists will be lobbying more strongly for an increase in the tax rate than if the damage coefficient and the size of the environmentalists were lower. The resulting political-economic equilibrium will imply a higher tax rate.[32] The political-economic equilibrium is affected in addition by the redistribution of the tax revenue: While in the benevolent dictator case it is a mere redistribution of income between members of the society that does not affect overall welfare, this redistribution affects the political equilibrium as not all groups of the society will reward additional income from tax proceeds as they are not organized (the workers). Thus redistribution effects matter for the political equilibrium, but not for the benevolent dictator. The lower $a$, the stronger the political distortion; for $a \longrightarrow \infty$ the political game solution converges to the benevolent dictator solution.

If the politically optimal tax rates are higher than the benevolent dictators' tax rates, they may reduce a distortion that is created by the strategic interaction of the two welfare-maximizing governments. Non-cooperative governments internalize the externality of transboundary pollution only to the extent that pollution affects domestic welfare. As a result, tax rates are too low compared with joint welfare maximization (Markusen 1975).If tax rates are lower than the in benevolent dictator case, the political distortion reinforces the distortion created by strategic interaction and welfare is even lower.

The fact that a political support maximizing government *may* pursue a welfare superior policy compared to a government that seeks to maximize welfare is thus contingent on a situation of international strategic interaction. It cannot arise in the analysis of a small open economy (e.g. Fredriksson 1997), where the political-economic calculus of the government unambiguously reduces overall welfare. Yet it may be relevant for a number of situations in

[32] Note that both profits and disutility from pollution are convex in the tax rate, however with different signs and magnitudes.

which environmental damage is high and environmental policies affect the environmental quality of neighboring states.

### 4.4 Welfare Implications of Political Institutions

As shown above, the institutional design of the political system (represented by the lobbying parameters $\alpha_E$, $\alpha_I$, and $a$) determine together with the properties of the ecological system (described by $\beta$) whether the political-economic equilibrium is welfare superior to the equilibrium obtained by welfare maximizing governments, and if so, to what extent. To investigate the welfare implication of different lobby group sizes, we calculate the closed form solution for the aggregate welfare $\Omega_E$ (cf. Eq. (13)) for the political economic equilibrium. It is given in Appendix 3. From this equation, we simulate the change in domestic welfare resulting from a change in $\alpha_E$ and $\alpha_I$ for different values of $a$. The welfare implication of a change in the share of environmentalists is depicted in Figure 6:[33]

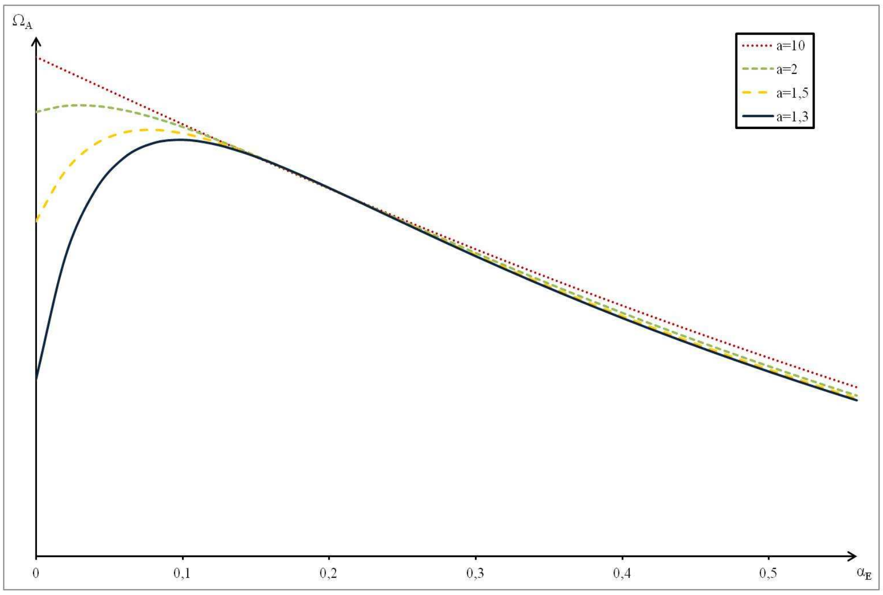



Figure 6: Welfare effects of a change in $\alpha_E$

An increase in $\alpha_E$ gives rise to two effects. First, the negative welfare effect of pollution increases since environmental degradation $S$ affects social welfare negatively only through environmentalists' utility. (cf. Eqs. (10) – (13)). Thus if the group of environmentalists grows, the 'felt' pollution, $\alpha_E S$, increases for any given level of physical pollution $S$ (cf. Eq. (1)) and aggregate welfare declines. Second, an increase in the environmental lobby group increases its lobbying contributions and thus raises $\tilde{t}^{PG}$ (cf. Appendix 2), thereby lowering physical pollution. These are two countervailing effects on aggregate welfare; their relative strength depends on the extent of the political distortion. This is shown in Figure 6: If the political distortion is high, i.e. $a$ is low, welfare first increases with a rising share of the environmental lobby group and declines after a maximum. If the political distortion is largely absent (high $a$)

[33] Parameters take on the values: $l = l^* = 1$; $\beta = 10$; $a^* = 10$; $\alpha_I = \alpha_I^* = 0.1$; $\alpha_E^* = 0.1$

the welfare implications are negative because now more people are negatively affected by the same level of physical pollution.

Quite contrary, the welfare implication of an increased number of industrialists is unambiguous. If $\alpha_I$ increases aggregate welfare increases. This is shown in Figure 7:[34]

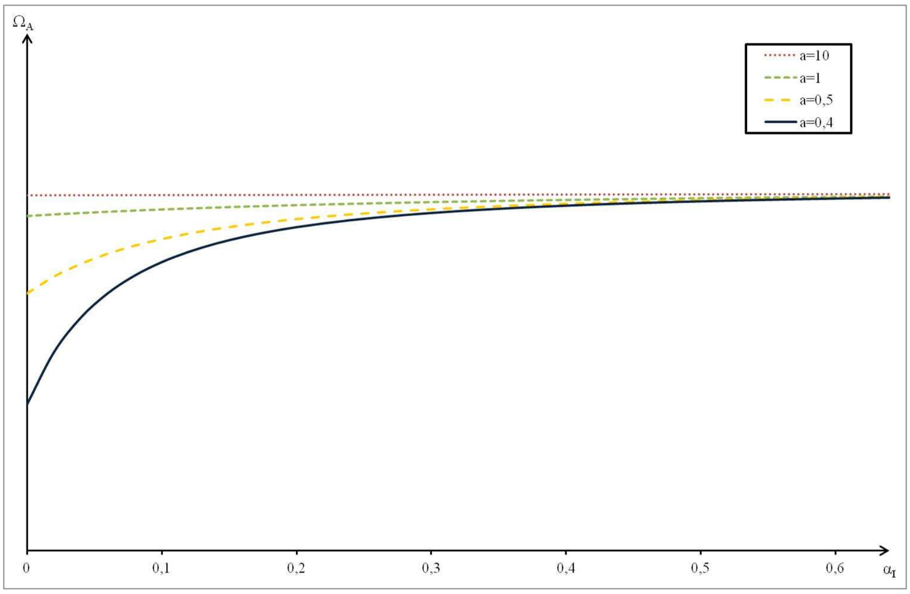



Figure 7: Welfare effects of a change in $\alpha_I$

An increase in $\alpha_I$ reduces the lobbying effort of the industrialists, because the net gain from lower taxation declines: while the increase in profits from lower taxation remains the same the reduction in redistributed tax revenue declines more strongly as the industrialists receive a large share of tax revenue with rising $\alpha_I$. As a consequence, $\tilde{t}^{PG}$ increases with $\alpha_I$ causing a rise in overall welfare. This effect is more pronounced the larger the political distortion as it makes lobbying more effective and is absent in the benevolent dictator case.

Proposition 4 summarizes our second main result.

**Proposition 4:**

(i) An increase in the size of the environmental lobby group may raise aggregate welfare up to a critical point and reduce it thereafter.

(ii) An increase in the size of the industrial lobby group unambiguously increases aggregate welfare.

*Proof*: Follow straightforwardly from differentiating the aggregate welfare equation (A3.1) with respect to $\alpha_E$ and $\alpha_I$. □

It is this asymmetry in welfare effects of increased lobbying that sets our situation apart from other – domestic – political economic situations. While in those other cases increased

[34] Parameters take on the values: $l = l^* = 1$; $\beta = 10$; $a^* = 10$; $\alpha_E = \alpha_E{}^* = 0.1$; $\alpha_I{}^* = 0.1$

lobbying from either side creates a larger political distortion and therefore a larger deviation from the welfare maximum, in our case increased lobbying of environmentalists increase welfare (up to a point)[35] while decreased lobbying of industrialists (due to a higher $\alpha_I$) increases welfare monotonously. Thus from a constitutional economic perspective, it may be reasonable to increase lobbying possibilities of environmental action groups and curb those of industry associations, if there are distortions through transnational pollution.

## 5. Concluding Remarks

In this paper, we have analyzed endogenous environmental policy formation of two countries that are small on the world markets, but are linked through transboundary pollution. Three major results emerge. First, the environmental policy adopted by self-interested governments may be more stringent than by social welfare maximizing, but uncooperative governments. The distortion created by the transboundary pollution may be alleviated (or exacerbated) by the distortion created through the political system. Therefore, under certain circumstances, a political process that does not take all individuals into consideration equally, may work in favor of the society at large. In our model, this is the case when the influence of the environmental lobby group is higher than of the industrial lobby group. We find that a marginal increase in the size of the environmentalist group increases aggregate welfare up to a point and declines thereafter. In contrast aggregate welfare decreases unambiguously with reduced lobbying by industrialists. Thus measured restrictions on industrialists' lobbying possibilities and support for lobbying by environmental action groups on a constitutional economics level may lead to a welfare-superior outcome. [36]

Second, the space of optimal policies in the political-economic game is larger than in the game played by benevolent dictators: While uncooperative benevolent governments will always set positive, but inefficiently low tax rates (from the perspective of joint welfare maximization), the politically optimal tax rates may even be too high to optimally internalize the transboundary externality (for strong environmental lobby groups), but they may also be too low (for strong industrial lobby groups). Political support maximizing governments may indeed subsidize the production of the polluting good rather than taxing it, if the relative strength of the industrial lobby group is large. In equilibrium, it is possible that one government subsidizes the production of the polluting good while the other taxes it. Again this suggests a moderate and asymmetric restriction of political influence.

Third, the political distortion might create instability, if the relative influence of either lobby group is too large: while the resulting equilibria in the case of social welfare maximizing governments are always unique and stable, the possibility of multiple equilibria

[35] Environmental degradation declines unambiguously.

[36] This could be brought about by a number of concrete measures such as a restriction of donations to parties as this would affect industry associations and corporations scope for lobbying disproportionately.

in the political game with one equilibrium being unstable cannot be excluded. This increases the probability of corner solutions with one country ceding production to the other.

We believe that the strategic interaction in environmental policy formation of self-interested governments in the presence of transboundary, but non-global pollution has not been sufficiently examined. Our theoretical analysis is aimed at improving our understanding of this issue, the scope of which can be broadened in many ways. First, it would be interesting to study what the incentives for political support maximizing governments are to cooperate and what the welfare effects would be. It is obvious that international cooperation would eliminate the distortion created by strategic interaction (Markusen 1975), but could also lead to a welfare deterioration if the political distortion has an offsetting effect. Second, interest groups could be assumed to lobby across the border. While industrialist lobby groups in both countries have opposing interests with respect to the national regulations; they favor higher regulation abroad and lower at home, environmentalist groups' interests in both countries are aligned. Third, extending the model to a multi-country setup with incomplete spill-overs would provide many important insights on real world applications of regional pollution.

Our model shows how distortions created by the strategic interaction of national governments interact with distortions created by the political processes in both countries. We show that these two sets of distortions could either reinforce or counteract each other. Which scenario is more realistic, however, remains an empirical question. It could be the subject of a fruitful empirical analysis.

# Appendices

### Appendix 1: Uniqueness and stability of the equilibria in the benevolent dictator case

In this appendix we demonstrate the existence and uniqueness of the welfare maximizing non-cooperative equilibrium of Section 3.

First, we note that both reaction curves $\tilde{t}^{\mathrm{BD}}$ and $\tilde{t}^{\mathrm{BD}*}$ are linear in the opponent‘s tax rate (cf. Eq. (16)). In the $(t, t^*)$ space, the slope of the ‘interior’ domestic reaction curve $\tilde{t}^{\mathrm{BD}}$, is

$$\frac{\partial \tilde{t}^{\mathrm{BD}}}{\partial t^*} = -\frac{\alpha_{\mathrm{E}}\beta}{1+\alpha_{\mathrm{E}}\beta} > -1\,, \tag{36}$$

while the slope of the inverse of the ‘interior’ foreign reaction curve, $\tilde{t}^{\mathrm{BD}*-1}$ is

$$\frac{\partial \tilde{t}^{\mathrm{BD}*-1}}{\partial t^*} = -\frac{1+\alpha_{\mathrm{E}}{}^*\beta}{\alpha_{\mathrm{E}}{}^*\beta} < -1\,. \tag{37}$$

For $t^* > p$ and for $t^* < t^*_{\mathrm{lb}}$, the domestic reaction function is flat, while the inverse of the foreign reaction function is vertical for $t > p$ and for $t < t^{\mathrm{BD}}_{\mathrm{lb}}$. That is, the inverse of the foreign reaction curve is always steeper than the domestic reaction curve and thus the reaction curves $t^{\mathrm{BD}}$ and $t^{\mathrm{BD}*}$ intersect exactly once. The resulting equilibrium is thus unique and stable.

□

### Appendix 2: Properties of the reaction curve in the political game

We first derive that the reaction function in the political game is downward sloping. This is shown by differentiating Eq. (27) w.r.t. the foreign tax rate.

$$\frac{\partial \tilde{t}^{\mathrm{PG}}}{\partial t^*} = -\frac{\alpha_{\mathrm{E}}\beta(a+1)}{(a+1)(1+\alpha_{\mathrm{E}}\beta)-2\alpha_{\mathrm{W}}} < 0\,. \tag{38}$$

An analogous expression can be derived for the inverse of the foreign reaction function.

Next we analyze the comparative static properties of the domestic reaction function with respect to $\alpha_{\mathrm{E}}$, $\alpha_{\mathrm{I}}$, $\beta$ and $a$. We calculate the change of $t^{\mathrm{PG}*}_{\mathrm{lb}}$ and $t^{\mathrm{PG}}_{\mathrm{low}}$ in response to a change in the relevant parameters using Eqs. (29) and (30).

$$\frac{\partial t^{\mathrm{PG}*}_{\mathrm{lb}}}{\partial \alpha_{\mathrm{E}}} = \frac{p(\alpha_{\mathrm{E}}+a)}{\alpha_{\mathrm{E}}{}^2\beta(a+1)} > 0\,, \tag{39}$$

$$\frac{\partial t^{\mathrm{PG}}_{\mathrm{low}}}{\partial \alpha_{\mathrm{E}}} = \frac{p(a+1)(\alpha_{\mathrm{I}}\beta + a\beta + 1)}{\left((a+1)(1+\alpha_{\mathrm{E}}\beta)-2\alpha_{\mathrm{W}}\right)^2} > 0\,, \tag{40}$$

$$\frac{\partial \tilde{t}^{\mathrm{PG}}}{\partial \alpha_{\mathrm{E}}} = \frac{a(p+1)+(a+1)\beta[\alpha_{\mathrm{I}}(3p-2t^*)+p]+\beta(a^2+1)(2p-t^*)}{\left((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\right)^2} > 0\,. \tag{41}$$

An increase in $\alpha_{\mathrm{E}}$ shifts the domestic reaction function to the Northeast; at the same time the range of foreign tax rates increases, for which the domestic government sets prohibitive tax rates as best response. The tax rate for zero foreign production increases as well. The slope of the reaction function can either increase or decrease:

$$\frac{\partial^2 \tilde{t}^{\mathrm{PG}}}{\partial t^* \partial \alpha_{\mathrm{E}}} = -\frac{\beta(a+1)(a+2\alpha_{\mathrm{I}}-1)}{\big((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\big)^2} \lesseqgtr 0\,. \tag{42}$$

If $\alpha_{\mathrm{I}}$ increases, the range of foreign tax rates for which the home country sets prohibitive taxes decreases and the reaction curve becomes flatter. $t_{\mathrm{low}}^{\mathrm{PG}}$ decreases for $\alpha_{\mathrm{E}}\beta > 1$ and increases otherwise.

$$\frac{\partial t_{\mathrm{lb}}^{\mathrm{PG}*}}{\partial \alpha_{\mathrm{I}}} = -\frac{p}{\alpha_{\mathrm{E}}\beta(a+1)} < 0\,, \tag{43}$$

$$\frac{\partial t_{\mathrm{low}}^{\mathrm{PG}}}{\partial \alpha_{\mathrm{I}}} = -\frac{p(a+1)(\alpha_{\mathrm{E}}\beta-1)}{\big((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\big)^2} \lesseqgtr 0\,, \tag{44}$$

$$\frac{\partial^2 \tilde{t}^{\mathrm{PG}}}{\partial t^* \partial \alpha_{\mathrm{I}}} = \frac{2\alpha_{\mathrm{E}}\beta(a+1)}{\big((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\big)^2} > 0\,. \tag{45}$$

The reaction curve shifts similarly in response to an increase in $a$:

$$\frac{\partial t_{\mathrm{lb}}^{\mathrm{PG}*}}{\partial a} = -\frac{p\alpha_{\mathrm{W}}}{\alpha_{\mathrm{E}}\beta(a+1)^2} < 0\,, \tag{46}$$

$$\frac{\partial t_{\mathrm{low}}^{\mathrm{PG}}}{\partial a} = \frac{p\alpha_{\mathrm{W}}(1-\alpha_{\mathrm{E}}\beta)}{\big((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\big)^2} \lesseqgtr 0\,, \tag{47}$$

$$\frac{\partial^2 \tilde{t}^{\mathrm{PG}}}{\partial t^* \partial a} = \frac{2\alpha_{\mathrm{E}}\beta\alpha_{\mathrm{W}}}{\big((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\big)^2} > 0\,, \tag{48}$$

$$\frac{\partial \tilde{t}^{\mathrm{PG}}}{\partial a} = \frac{\alpha_{\mathrm{W}}[p-\alpha_{\mathrm{E}}\beta(3p-2t^*)]}{\big((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\big)^2} \lesseqgtr 0\,. \tag{49}$$

In other words, if $a$ increases the new reaction curve is flatter than the old reaction curve and it may lie completely below the old one or may intersect with it.

An increase in $\beta$ shifts the domestic reaction curve to the Northeast, which is qualitatively the same reaction to an increase in $\alpha_{\mathrm{E}}$. However, if $\beta$ rises, the foreign reaction curve shifts as well making the effect of the equilibrium qualitatively different.

$$\frac{\partial t_{\mathrm{lb}}^{\mathrm{PG}*}}{\partial \beta} = \frac{p(a+1-\alpha_{\mathrm{W}})}{\alpha_{\mathrm{E}}\beta^2(a+1)} > 0\,, \tag{50}$$

$$\frac{\partial t_{\mathrm{low}}^{\mathrm{PG}}}{\partial \beta} = \frac{p\alpha_{\mathrm{E}}(a+1)(a+1-\alpha_{\mathrm{W}})}{\big((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\big)^2} > 0\,, \tag{51}$$

$$\frac{\partial^2 \tilde{t}^{\mathrm{PG}}}{\partial t^* \partial \beta} = -\frac{\alpha_{\mathrm{E}}(a+1)(a+1-2\alpha_{\mathrm{W}})}{\big((a+1)(\alpha_{\mathrm{E}}\beta+1)-2\alpha_{\mathrm{W}}\big)^2} \lesseqgtr 0 \,, \tag{52}$$

$$\frac{\partial t_{\mathrm{lb}}^{\mathrm{PG}}}{\partial \beta} = \frac{p(a^*+1-\alpha_{\mathrm{W}}{}^*)}{\alpha_{\mathrm{E}}{}^*\beta^2(a^*+1)} > 0 \,, \tag{53}$$

$$\frac{\partial t_{\mathrm{low}}^{\mathrm{PG}*}}{\partial \beta} = \frac{p\alpha_{\mathrm{E}}{}^*(a^*+1)(a^*+1-\alpha_{\mathrm{W}}{}^*)}{\big((a^*+1)(\alpha_{\mathrm{E}}{}^*\beta+1)-2\alpha_{\mathrm{W}}{}^*\big)^2} > 0 \,. \tag{54}$$

□

## Appendix 3: Welfare in the Political Game

Below we give the closed form solution for the aggregate welfare for the political economic equilibrium. It is derived by plugging the politically optimal tax rates into eq. (14). In this notation, capital letters refer to foreign values, i.e. $A \equiv a^{*}$, $A_{\mathrm{E}} \equiv \alpha_{\mathrm{E}}^{*}$ and $A_{\mathrm{I}} \equiv \alpha_{\mathrm{I}}^{*}$. This equation is used to simulate Figures 6 and 7.

$$
\begin{aligned}
\omega_A :=& -\frac{1}{4}\Big(\mathrm{Preis}^2\big(1 - a\alpha_E\beta + \alpha_I A_E\beta + 3\alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + 2\alpha_E\beta A + \alpha_E A A_E\beta + \alpha_I A A_E\beta - A_E\beta \\
& + \alpha_I A + 2\alpha_I A_E - 2A_E + \alpha_E A - 2A_I + 2\alpha_E A_I - \alpha_I + 2\alpha_I A_I + 3a\alpha_E\beta A_I - \alpha_E - \alpha_E\beta - A + 2a\alpha_E\beta A \\
& + 3a\alpha_E\beta A_E + 2\alpha_E A_E\big)^2\Big) \Big/ \big(1 - a - a\alpha_E\beta + 2\alpha_I A_E\beta + 2\alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + \alpha_E\beta A \\
& + 2\alpha_E A A_E\beta + 2\alpha_I A A_E\beta - A_E\beta + 2\alpha_I A + A A_E\beta a + 4\alpha_I A_E - 2A_E + 2\alpha_E A - 2A_I + 4\alpha_E A_I - 2\alpha_I \\
& + 4\alpha_I A_I + 2a\alpha_E\beta A_I + 2A_E a - 2\alpha_E - \alpha_E\beta + A a - A + 2A_I a + a\alpha_E\beta A + a\beta A_E + 2a\alpha_E\beta A_E + 4\alpha_E A_E\big)^2 \\
& + 1 + \frac{1}{4}\mathrm{Preis}^2 - \alpha_E\beta\,\mathrm{Preis}^2 + \Big(\alpha_E\beta\,\mathrm{Preis}^2\big(1 - a\alpha_E\beta + \alpha_I A_E\beta + 3\alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta \\
& + 2\alpha_E\beta A + \alpha_E A A_E\beta + \alpha_I A A_E\beta - A_E\beta + \alpha_I A + 2\alpha_I A_E - 2A_E + \alpha_E A - 2A_I + 2\alpha_E A_I - \alpha_I + 2\alpha_I A_I \\
& + 3a\alpha_E\beta A_I - \alpha_E - \alpha_E\beta - A + 2a\alpha_E\beta A + 3a\alpha_E\beta A_E + 2\alpha_E A_E\big)\Big) \Big/ \big(1 - a - a\alpha_E\beta + 2\alpha_I A_E\beta + 2\alpha_E\beta A_I \\
& - A A_E\beta + 4\alpha_E A_E\beta + \alpha_E\beta A + 2\alpha_E A A_E\beta + 2\alpha_I A A_E\beta - A_E\beta + 2\alpha_I A + A A_E\beta a + 4\alpha_I A_E - 2A_E \\
& + 2\alpha_E A - 2A_I + 4\alpha_E A_I - 2\alpha_I + 4\alpha_I A_I + 2a\alpha_E\beta A_I + 2A_E a - 2\alpha_E - \alpha_E\beta + A a - A + 2A_I a + a\alpha_E\beta A \\
& + a\beta A_E + 2a\alpha_E\beta A_E + 4\alpha_E A_E\big) + \Big(\alpha_E\beta\,\mathrm{Preis}^2\big(1 - a - a\alpha_E\beta + 3\alpha_I A_E\beta + \alpha_E\beta A_I - A A_E\beta \\
& + 4\alpha_E A_E\beta + 3\alpha_E A A_E\beta + 3\alpha_I A A_E\beta - A_E\beta + 2A A_E\beta a + 2\alpha_I A_E - A_E - A_I + 2\alpha_E A_I - 2\alpha_I + 2\alpha_I A_I \\
& + a\alpha_E\beta A_I + A_E a - 2\alpha_E - \alpha_E\beta + A_I a + 2a\beta A_E + a\alpha_E\beta A_E + 2\alpha_E A_E\big)\Big) \Big/ \big(1 - a - a\alpha_E\beta + 2\alpha_I A_E\beta \\
& + 2\alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + \alpha_E\beta A + 2\alpha_E A A_E\beta + 2\alpha_I A A_E\beta - A_E\beta + 2\alpha_I A + A A_E\beta a + 4\alpha_I A_E \\
& - 2A_E + 2\alpha_E A - 2A_I + 4\alpha_E A_I - 2\alpha_I + 4\alpha_I A_I + 2a\alpha_E\beta A_I + 2A_E a - 2\alpha_E - \alpha_E\beta + A a - A + 2A_I a \\
& + a\alpha_E\beta A + a\beta A_E + 2a\alpha_E\beta A_E + 4\alpha_E A_E\big) - \frac{1}{4}\Big(\alpha_E\beta\,\mathrm{Preis}^2\big(1 - a\alpha_E\beta + \alpha_I A_E\beta + 3\alpha_E\beta A_I \\
& - A A_E\beta + 4\alpha_E A_E\beta + 2\alpha_E\beta A + \alpha_E A A_E\beta + \alpha_I A A_E\beta - A_E\beta + \alpha_I A + 2\alpha_I A_E - 2A_E + \alpha_E A - 2A_I \\
& + 2\alpha_E A_I - \alpha_I + 2\alpha_I A_I + 3a\alpha_E\beta A_I - \alpha_E - \alpha_E\beta - A + 2a\alpha_E\beta A + 3a\alpha_E\beta A_E + 2\alpha_E A_E\big)^2\Big) \Big/ \big(1 - a \\
& - a\alpha_E\beta + 2\alpha_I A_E\beta + 2\alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + \alpha_E\beta A + 2\alpha_E A A_E\beta + 2\alpha_I A A_E\beta - A_E\beta + 2\alpha_I A \\
& + A A_E\beta a + 4\alpha_I A_E - 2A_E + 2\alpha_E A - 2A_I + 4\alpha_E A_I - 2\alpha_I + 4\alpha_I A_I + 2a\alpha_E\beta A_I + 2A_E a - 2\alpha_E - \alpha_E\beta \\
& + A a - A + 2A_I a + a\alpha_E\beta A + a\beta A_E + 2a\alpha_E\beta A_E + 4\alpha_E A_E\big)^2 - \frac{1}{2}\Big(\alpha_E\beta\,\mathrm{Preis}^2\big(1 - a\alpha_E\beta + \alpha_I A_E\beta \\
& + 3\alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + 2\alpha_E\beta A + \alpha_E A A_E\beta + \alpha_I A A_E\beta - A_E\beta + \alpha_I A + 2\alpha_I A_E - 2A_E + \alpha_E A \\
& - 2A_I + 2\alpha_E A_I - \alpha_I + 2\alpha_I A_I + 3a\alpha_E\beta A_I - \alpha_E - \alpha_E\beta - A + 2a\alpha_E\beta A + 3a\alpha_E\beta A_E + 2\alpha_E A_E\big)\big(1 - a \\
& - a\alpha_E\beta + 3\alpha_I A_E\beta + \alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + 3\alpha_E A A_E\beta + 3\alpha_I A A_E\beta - A_E\beta + 2A A_E\beta a + 2\alpha_I A_E \\
& - A_E - A_I + 2\alpha_E A_I - 2\alpha_I + 2\alpha_I A_I + a\alpha_E\beta A_I + A_E a - 2\alpha_E - \alpha_E\beta + A_I a + 2a\beta A_E + a\alpha_E\beta A_E + 2\alpha_E A_E\big) \\
& \Big) \Big/ \big(1 - a - a\alpha_E\beta + 2\alpha_I A_E\beta + 2\alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + \alpha_E\beta A + 2\alpha_E A A_E\beta + 2\alpha_I A A_E\beta \\
& - A_E\beta + 2\alpha_I A + A A_E\beta a + 4\alpha_I A_E - 2A_E + 2\alpha_E A - 2A_I + 4\alpha_E A_I - 2\alpha_I + 4\alpha_I A_I + 2a\alpha_E\beta A_I + 2A_E a \\
& - 2\alpha_E - \alpha_E\beta + A a - A + 2A_I a + a\alpha_E\beta A + a\beta A_E + 2a\alpha_E\beta A_E + 4\alpha_E A_E\big)^2 - \frac{1}{4}\Big(\alpha_E\beta\,\mathrm{Preis}^2\big(1 \\
& - a - a\alpha_E\beta + 3\alpha_I A_E\beta + \alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + 3\alpha_E A A_E\beta + 3\alpha_I A A_E\beta - A_E\beta + 2A A_E\beta a \\
& + 2\alpha_I A_E - A_E - A_I + 2\alpha_E A_I - 2\alpha_I + 2\alpha_I A_I + a\alpha_E\beta A_I + A_E a - 2\alpha_E - \alpha_E\beta + A_I a + 2a\beta A_E + a\alpha_E\beta A_E \\
& + 2\alpha_E A_E\big)^2\Big) \Big/ \big(1 - a - a\alpha_E\beta + 2\alpha_I A_E\beta + 2\alpha_E\beta A_I - A A_E\beta + 4\alpha_E A_E\beta + \alpha_E\beta A + 2\alpha_E A A_E\beta \\
& + 2\alpha_I A A_E\beta - A_E\beta + 2\alpha_I A + A A_E\beta a + 4\alpha_I A_E - 2A_E + 2\alpha_E A - 2A_I + 4\alpha_E A_I - 2\alpha_I + 4\alpha_I A_I \\
& + 2a\alpha_E\beta A_I + 2A_E a - 2\alpha_E - \alpha_E\beta + A a - A + 2A_I a + a\alpha_E\beta A + a\beta A_E + 2a\alpha_E\beta A_E + 4\alpha_E A_E\big)^2
\end{aligned}
$$

(A3.1)

## References


Aidt, T. (1998). Political internalization of economic externalities and environmental policy. *Journal of Public Economics* 69: 1–16.

Aidt, T. (2005). The rise of environmentalism, pollution taxes and intra-industry trade. *Economics of Governance* 6: 1–12.

Aidt, T. (2010). Green taxes: refunding rules and lobbying. *Journal of Environmental Economics and Management* 60(1): 31–43.

Akimoto, H. (2003). Global air quality and pollution. *Science* (302): 1716–1719.

Barrett, S. (1994). Strategic environmental policy and international trade. *Journal of Public Economics* 54: 325–38.

Barrett, S. (2003). *Environment and statecraft: the strategy of environmental treaty-making*. Oxford University Press, Oxford.

Bernheim, B. and Whinston, M. (1986). Menu auctions, resource allocation, and economic influence. *Quarterly Journal of Economics* 101(1): 1–31.

Bhagwati, J. (1982). Directly unproductive, profit-seeking (DUP) activities. *Journal of Political Economy* 90(5): 988–1002.

Binder, S. and Neumayer, E. (2005). Environmental pressure group strength and air pollution: an empirical analysis. *Ecological Economics* 55(4): 527–538.

Bommer, R. and Schulze, G. (1999). Environmental improvement with trade liberalization. *European Journal of Political Economy* 15(4): 639–661.

Cansier, D. and Krumm, R. (1997). Air pollutant taxation: an empirical survey. *Ecological Economics* 23: 59–70.

Conconi, P. (2003). Green lobbies and transboundary pollution in large open economies. *Journal of International Economics* 59: 399–422.

Fredriksson, P. (1997). The political economy of pollution taxes in a small open economy. *Journal of Environmental Economics and Management* 33: 44–58.

Fredriksson, P., Neumayer, E., Damania, R., and Gates, S. (2005). Environmentalism, democracy, and pollution control. *Journal of Environmental Economics and Management* 49(2): 343–365.

Fredriksson, P. and Svensson, J. (2003). Political instability, corruption and policy formation: the case of environmental policy. *Journal of Public Economics* 87(7–8): 1383–1405.

Gren, I.-M. (2001). International versus national actions against nitrogen pollution of the Baltic Sea. *Environmental and Resource Economics* 20: 41–59.

Grossman, G. and Helpman, E. (1994). Protection for sale. *American Economic Review* 84(4): 833–850.

Grossman, G. and Helpman, E. (1995). Trade wars and trade talks. *Journal of Political Economy* 103(4): 675–708.

HELCOM (2010). *Ecosystem Health of the Baltic Sea 2003–2007: HELCOM Initial Holistic Assessment*. Baltic Sea Environment Proceedings No. 122. Helsinki Commission, Helsinki: Baltic Marine Environment Protection Commission.

Hillman, A. and Ursprung, H. (1988). Domestic politics, foreign interests, and international trade policy. *American Economic Review* 78(4): 719–45.

Ichikawa, Y. and Fujita, S. (1995). An analysis of wet deposition of sulfate using a trajectory model for East Asia. *Water, Air, and Soil Pollution* 85: 1927–1932.

List, J. and Sturm, D. (2006). How elections matter: theory and evidence from environmental policy. *Quarterly Journal of Economics* 121(4): 1249–1281.

Liu, J. and Diamond, J. (2005). China's Environment in a Globalizing World. *Nature* 435: 1179–1186.

Lu, Z. et al. (2010). Sulfur dioxide emissions in China and sulfur trends in East Asia since 2000. *Atmospheric Chemistry and Physics* 10: 6311–6331.

Markusen, J. (1975). Cooperative control of international pollution and common property resources. *Quarterly Journal of Economics* 89(4): 618–32.

Olson, M. (1965). *The Logic of Collective Action*. Harvard University Press, Cambridge, MA.

Rauscher, M. (1997). *International trade, factor movements, and the environment*. Oxford: Oxford University Press.

Rauscher, M. (2005). International trade, foreign investment, and the environment. In: K.-G. Mäler, J. Vincent, eds., *Handbook of Environmental and Resource Economics* Vol. 3. Amsterdam: Elsevier (North-Holland Handbooks in Economics). 1403–1456.

Schleich, J. (1999). Environmental quality with endogenous domestic and trade policies. *European Journal of Political Economy* 15(1) 53–71.

Schleich, J. and Orden, D. (2000). Environmental quality and industry protection with noncooperative versus cooperative domestic and trade policies. *Review of International Economics* 8(4) 681–697.

Ramanathan, V. et al. (2008). *Atmospheric Brown Clouds: Regional Assessment Report with Focus on Asia*. United Nations Environment Programme. Nairobi, Kenya.