ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Eichner, Thomas; Pethig, Rüdiger

Working Paper

Global Environmental Agreements and International Trade: Asymmetry of Countries Matters

CESifo Working Paper, No. 5051

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Eichner, Thomas; Pethig, Rüdiger (2014) : Global Environmental Agreements and International Trade: Asymmetry of Countries Matters, CESifo Working Paper, No. 5051, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/105098

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Global Environmental Agreements and International Trade: Asymmetry of Countries Matters

Thomas Eichner
Rüdiger Pethig

CESIFO WORKING PAPER NO. 5051
CATEGORY 10: ENERGY AND CLIMATE ECONOMICS
NOVEMBER 2014

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

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5051*

Global Environmental Agreements and International Trade: Asymmetry of Countries Matters

Abstract

We investigate the formation of global climate agreements (= stable grand climate coalitions) in a model, in which climate policy takes the form of carbon emission taxation and fossil fuel and consumption goods are traded on world markets. We expand the model of Eichner and Pethig (2014) by considering countries that are identical within each of two groups but differ across groups with respect to climate damage or fossil fuel demand. Our numerical analysis suggests that climate damage asymmetry tends to discourage cooperation in the grand coalition. The effects of fuel-demand asymmetry depend on fossil fuel abundance. If fuel is very abundant, the grand coalition fails to be stable independent of the degree of fuel demand asymmetry. If fuel is sufficiently scarce, low degrees of fuel demand asymmetry discourage cooperation whereas higher degrees of asymmetry stabilize the grand coalition.

JEL-Code: C720, F020, Q500, Q580.

Keywords: fuel demand, climate damage, international trade, asymmetry, stability, grand coalition.

Thomas Eichner
Department of Economics
University of Hagen
Universitätsstr. 41
Germany – 58097 Hagen
thomas.eichner@fernuni-hagen.de

Rüdiger Pethig
Department of Economics
University of Siegen
Hölderlinstr. 3
Germany – 57068 Siegen
pethig@vwl.wiwi.uni-siegen.de

1 The problem

The reduction of global carbon emissions necessary to stabilize the world climate at safe levels requires a broad and deep international environmental agreement (IEA), i.e. an IEA that secures cooperation of all (major) countries and strives for maximum world welfare. The first legally binding IEA on climate change, the Kyoto Protocol, accomplished only little more than global non-cooperation. It expired in 2012, and the prospects are uncertain for reaching a more effective follow-up agreement. That calls for further efforts to improve our understanding of how to reach a broad and deep climate agreement.

In line with the prediction of the standard theory of voluntary contributions to a public good, the countries' efforts to reduce carbon emissions are sub-optimally low in the absence of cooperation. The impact of cooperation - or coalition formation - has been studied in what Rubio and Ulph (2006) call the workhorse model of the coalition formation literature[1] that specifies the welfare of identical autarkic countries as a function, which is increasing and concave in domestic, and declining and concave in aggregate carbon emissions. In that model, large coalitions are unstable.[2] Eichner and Pethig (2013, 2014) expand the workhorse model by introducing international trade and thus add the trade interdependence to the countries' interdependence through climate damage.[3,4] In their model with identical countries, international trade and tax policy, Eichner and Pethig (2014) identify a subset of parameters under which the grand coalition is stable and implements the social optimum regardless of whether Nash or Stackelberg coalitions are assumed. That remarkable result motivates us to investigate in the present paper whether the conditions for stable grand coalitions are even more favorable if we deviate from Eichner and Pethig (2014) by dropping the assumption of symmetric countries.[5]

[1]Here we refer to the workhorse model employed by Carraro and Siniscalco (1991), Hoel (1992), Barrett (1994), surveyed by Finus (2003) and rigorously characterized by Diamantoudi and Sartzetakis (2006) and by Rubio and Ulph (2006). We consider corner solutions (e.g. zero emissions) an implausible artefact of the parametric approach and therefore follow Diamantoudi and Sartzetakis (2006) who restrict their focus on interior solutions. Under these conditions, stable coalitions turn out to be small and shallow in the workhorse model.

[2]The concept of stability applied in the literature of self-enforcing IEAs was originally introduced by D'Aspremont et al. (1983) in the context of cartel theory. A climate coalition is stable – or equivalently, an IEA is self-enforcing – if non-member wants to join and no member wants to leave the coalition.

[3]Eichner and Pethig (2013, 2014) place a game model on top of a simple general competitive equilibrium of the world economy, which increases analytical complexity considerably.

[4]There is a large literature on trade and trans-frontier pollution, e.g. Copeland and Taylor (2005) and Chen and Woodland (2013), which does not address the formation of coalitions, however.

[5]Since we restrict our attention to the grand coalition, it suffices for the check of stability to envisage the coalition with all countries but one and an individual country outside that coalition that determines its net benefit of joining the coalition.

Obviously, country heterogeneity is both a general characteristic of the real world and an important determinant of coalition formation. Barrett (1997) investigates self-enforcing IEAs for two types of countries in the workhorse model by means of numerical simulations. Fuentes-Albero and Rubio (2010) and de Zeeuw and Pavlova (2013) analytically characterize stable coalition in a simpler model with linear damage. These papers conclude that the size of stable coalitions can be larger but large stable coalitions go hand in hand with small gains of cooperation. Based on that pessimistic view, the impact of transfers on the performance of self-enforcing IEAs with asymmetric countries has been studied with various simplifying assumptions and degrees of complexity, e.g. Carraro and Botteon (1997), Carraro and Siniscalco (1998), Barrett (2001), Bosello et al. (2003), Weikard et al. (2006) or McGinty (2007). The results are mixed concerning the possibility to improve the performance of self-enforcing IEAs via transfers. In the present paper we disregard transfers but focus instead on the question, how strong the capacity of international trade is to generate gains from cooperation.

Asymmetry boosts complexity and makes it difficult to derive general results. Unless models with asymmetric countries are extremely simple, they can be solved numerically only, even if one assumes, as we will do in line with most of the theoretical literature, that countries are identical within each of two groups but different across groups. The price to be paid is that apart from a few general results we need to resort to numerical analysis. Specifically, we elaborate and discuss five samples. In the first two samples we focus on climate damage asymmetry by varying continuously the size and asymmetry of a climate damage parameter that determines the countries' scale of damage for given levels of aggregate carbon emissions. We find that if climate damage parameters are low and hence the grand coalition is stable in the absence of asymmetries, as shown by Eichner and Pethig (2014), growing damage asymmetry eventually turns stability into instability. If climate damage is high, the grand coalition is unstable at all degrees of asymmetry - including symmetry. These findings suggest that climate damage asymmetry generally discourages cooperation in the grand coalition.

The other three samples deal with fuel demand asymmetry. We perform continuous variations of fuel preference parameters and thus shift up and down the size and asymmetry of the countries' fuel demand. We find that if extraction costs are low – and hence fuel is abundant – stability is not attained, neither with nor without fuel demand asymmetry. Low degrees of fuel demand asymmetry discourage cooperation whereas higher asymmetry stabilizes the grand coalition (again), if extraction costs are sufficiently high. The size of extraction costs plays a similar role as the size of fuel demand, because an increase in either tends to boost fuel consumption and hence aggregate emissions.

The paper is organized as follows. Section 2 outlines the analytical framework. Section 3 derives the welfare for an individual country of either group when it is a member of the grand coalition and when it free rides. We need to determine and compare these welfares, because the country's decision to join or not to join depends on the sign of the welfare difference. Section 5 presents the numerical analysis and the key results based on five samples each of which comprises continuous variations of one of the asymmetries. Section 6 summarizes and concludes.

2 Analytical framework

Consider a world economy, in which each country $i = 1, \ldots, m+n$ produces two consumer goods, a standard composite consumption good (quantity x_i^s) and fossil energy (quantity e_i^s), called fuel, from domestic fossil reserves according to[6,7]

$$x_i^s = X\left(e_i^s\right) := \bar{x} - \frac{\xi}{2}\left(e_i^s\right)^2 \quad \text{for } i = 1, \ldots, m+n, \tag{1}$$

where ξ and $\bar{x}$ are positive parameters. The transformation function (1) is the same for all countries and implies that both commodities are produced by means of domestic productive factors with given endowments. Consumers derive utility from the consumption good (quantity x_i^d) and fuel (quantity e_i^d). Fuel consumption generates the greenhouse gas carbon dioxide whose emissions are proportional to fuel consumption. By suitable choice of units, we denote by e_i^d both fuel consumption and emissions. Global emissions cause climate damage that hits all countries in the form of a loss of utility of the countries' representative consumer. However, the disutility of climate damage and/or the consumers' preferences for fuel may differ. Specifically, we distinguish two groups of countries, group $M := \{1, \ldots, m\}$ and group $N := \{1, \ldots, n\}$, and specify the utility function of the representative consumer of country i as[8]

$$\begin{array}{ll} V^m\left(e_i^d\right) + x_i^d - D^m\left(\sum_{j=1}^{m+n} e_j^d\right) & \text{if } i \in M, \\ V^n\left(e_i^d\right) + x_i^d - D^n\left(\sum_{j=1}^{m+n} e_j^d\right) & \text{if } i \in N, \end{array} \tag{2}$$

where $V^m\left(e_i^d\right) = a_m e_i^d - \frac{b}{2}\left(e_i^d\right)^2$, $V^n\left(e_i^d\right) = a_n e_i^d - \frac{b}{2}\left(e_i^d\right)^2$, $D^m\left(\sum_{j=1}^{m+n} e_j^d\right) := \frac{\delta_m}{2}\left(\sum_{j=1}^{m+n} e_j^d\right)^2$, $D^n\left(\sum_{j=1}^{m+n} e_j^d\right) := \frac{\delta_n}{2}\left(\sum_{j=1}^{m+n} e_j^d\right)^2$, and where the parameters in (2) satisfy $a_m, a_n, b, \delta_m > 0$ and $\delta_n \in \mathbb{R}$.

[6] The superscripts s and d indicate quantities supplied and demanded, respectively. Upper-case letters denote functions. Subscripts attached to them indicate partial derivatives.

[7] We use general functional forms such as X in equation (1) only for convenience of notation. For reasons of tractability, we will make use of linear-quadratic functional forms throughout the paper.

[8] We also refer to the utility of country i's representative consumer as country i's welfare.

According to (2), consumers in both groups of countries may differ with respect to the parameters a_m and a_n implying differences in their demand for fuel, and/or with respect to the parameters δ_m and δ_n implying differences in the disutility of climate damage. For analytical convenience, we define $a_m := a$, $a_n := a + c$, $c \in \mathbb{R}$ and $\delta_m = \delta$, $\delta_n = \delta + \rho$, $\rho \in \mathbb{R}$. Exogenous variations of the parameters c and ρ are variations of the degree of asymmetry and exogenous variations of the parameters a and δ are variations in the size of fuel demand and climate damage, respectively.

The aggregate resource constraints

$$\sum_{j=1}^{m+n} x_j^d = \sum_{j=1}^{m+n} x_j^s \quad \text{and} \quad \sum_{j=1}^{m+n} e_j^d = \sum_{j=1}^{m+n} e_j^s \tag{3}$$

complete the description of the basic model.

The next step is to establish the existence of the general competitive equilibrium with world markets for the consumption good (price $p_x \equiv 1$) and for fuel (producer price p), when each country regulates its domestic emissions by means of an arbitrarily chosen emission tax t_i. Consider first country i's fuel supply

$$e_i^s = \frac{p}{\xi} \quad \text{for } i \in M \cup N \tag{4}$$

that is straightforward from firm i's maximization of profits (including factor income) $x_i^s + pe_i^s$ subject to (1). Consumer i takes climate damage and her income[9] $y_i := x_i^s + pe_i^s + t_i e_i^d$ as given and maximizes consumption utility $V^h\left(e_i^d\right) + x_i^d$ subject to $x_i^d + (p + t_i)e_i^d = y_i$. Due to the one-to-one relation between emissions and fuel consumption, the emission tax is equivalent to a tax on fuel consumption. The resultant demands are

$$e_i^d = \begin{cases} \frac{a-p-t_i}{b} & \text{for } i \in M, \\ \frac{(a+c)-p-t_i}{b} & \text{for } i \in N, \end{cases} \tag{5}$$

$$x_i^d = X\left(e_i^s\right) + p\left(e_i^s - e_i^d\right). \tag{6}$$

The fuel demand functions (5) differ across types of countries. For given c the parameter a determines the size of fuel demand and for given a the parameter c determines the degree of asymmetry of fuel demand.

The combination of (3), (4) and (5) yields the market-clearing[10] producer price of fuel,

$$p = P(\mathbf{t}, c) := \frac{\xi\left[(m+n)a + nc - \sum_{j=1}^{m+n} t_j\right]}{(m+n)(b+\xi)}, \tag{7}$$

[9] Income is profits plus factor income plus recycled tax revenue.

[10] By Walras Law, the market for the consumption good is in equilibrium, if the fuel market is in equilibrium.

where $\mathbf{t} := (t_1, \ldots, t_{m+n})$ is short for the (given) profile of tax rates. The countries' equilibrium demands and supplies for fuel as functions of tax rates follow from inserting (7) in (4) and (5):

$$e_i^d = \frac{[a_m - P(\mathbf{t}, c) - t_i]}{b} \text{ for } i \in M, \; e_i^d = \frac{[a_n - P(\mathbf{t}, c) - t_i]}{b} \text{ for } i \in N \text{ and } e_i^s = \frac{P(\mathbf{t}, c)}{\xi}. \quad (8)$$

Combined with (7), these functions determine the equilibrium demands and supplies of the consumption good as functions of tax rates via (1) and (6). Thus, each tax profile $\mathbf{t}$ maps into a unique competitive equilibrium of the world economy. The corresponding equilibrium welfare of country i reads

$$\begin{aligned} W^m(\mathbf{t}, c, \rho) &:= V^m\left(e_i^d\right) + X\left(e_i^s\right) + P(\mathbf{t}, c)\left(e_i^s - e_i^d\right) - D^m\left(\textstyle\sum_{j=1}^{m+n} e_j^d\right) \quad \text{if } i \in M, \\ W^n(\mathbf{t}, c, \rho) &:= V^n\left(e_i^d\right) + X\left(e_i^s\right) + P(\mathbf{t}, c)\left(e_i^s - e_i^d\right) - D^n\left(\textstyle\sum_{j=1}^{m+n} e_j^d\right) \quad \text{if } i \in N, \end{aligned} \quad (9)$$

where e_i^s and e_i^d satisfy (8).[11] Note that due to the additivity of the utility function (2) the equilibrium allocation $\left(e_i^d, e_i^s, x_i^d, x_i^s\right)_{\in M \cup N}$ depends on the asymmetry parameter c, but not on the damage difference parameter ρ. The parameter ρ affects the equilibrium welfare (9), however.

Central to our subsequent analysis is the standard concept of self-enforcing IEAs, or stable environmental coalitions, introduced by D'Aspremont et al. (1983),[12] which requires that no signatory has an incentive to defect and no non-signatory has an incentive to sign. As pointed out in the introduction, we restrict our attention to the stability of the grand coalition. It therefore suffices to compare the situation, in which a grand coalition exists, with situations, in which a coalition of all countries but one, i.e. a coalition of $m + n - 1$ countries, coexists with a single country of group M or N outside the coalition. The outsider then has to decide whether to join the coalition or free ride. Obviously, the outcome is a stable grand coalition, if and only if it is in the outsider's self-interest to join.

We denote as h-country ($h = m, n$) a country of group $H = M, N$ *inside* the coalition and as f-country ($f = m, n$) a country of group $F = M, N$ *outside* the coalition. If a country is in the grand coalition, it enjoys the first-best equilibrium welfare which we denote as $w_h^* = \mathcal{W}^{*h}(c, \rho)$. If a country stays outside the coalition of size $m + n - 1$, it calculates its equilibrium welfare $\bar{w}_f = \bar{\mathcal{W}}^f(c, \rho)$ based on the expectation that the coalition of size $m + n - 1$ will continue to exist.[13] The functions $\mathcal{W}^{*h}$ and $\bar{\mathcal{W}}^f$ (with $f = h = m, n$) describe the dependence of welfares on the degree of asymmetry that is measured by the asymmetry parameters[14] c and ρ. Using that notation, the grand coalition is said to be

[11] We refrain from directly inserting the terms e_i^d and e_i^s from (8) into (9) to ease the exposition.

[12] We use the terms "stable (environmental) coalition" and "self-enforcing IEA" synonymously.

[13] This assumption is implicit in the concept of coalition stability referred to in the introduction.

[14] These welfare functions depend on all model parameters, of course. However, we suppress the parameters other than c and ρ, because our focus will be on these asymmetry parameters.

stable for some given asymmetry parameters c and ρ,[15] if the welfare difference $\bar{\mathcal{D}}^f(c,\rho) := \mathcal{W}^{*h}(c,\rho) - \bar{\mathcal{W}}^f(c,\rho)$ satisfies $\bar{\mathcal{D}}^f(c,\rho) \geq 0$ for $f = h = m$ and for $f = h = n$. The stability of the grand coalition obviously requires that signing the agreement makes m- **and** n-countries at least as well off as free riding. That is, the coalition is unstable, if m-countries **or** n-countries, **or** both, prefer free riding to signing.

3 Welfare functions

We proceed by specifying the (first-best) welfare $w_h^* = \mathcal{W}^{*h}(c,\rho)$, of an individual country of group $H = M, N$ in the grand coalition. Our point of departure is the welfare $W^h(\mathbf{t}, c, \rho)$ in (9) that represents the welfare a country of group $H = M, N$ attains in the competitive equilibrium with tax profile $\mathbf{t} = (t_1, \ldots, t_{m+n})$. An obvious necessary condition for maximizing the aggregate welfare of the grand coalition - as of any coalition - is a uniform tax rates of its members.[16] Therefore, we assign to all countries the same tax rate, denoted t_z, and write the welfare of countries of group $H = M, N$ in the grand coalition as $W^h(\mathbf{t}, c, \rho) = W^h(t_z, c, \rho)$, when the tax profile is $\mathbf{t} = \underbrace{(t_z, \ldots, t_z)}_{(m+n)-\text{times}}$. With this notation, we rewrite the fuel price function (7) as

$$P(\mathbf{t}, c) = P(t_z, c) := \frac{\xi\left[(m+n)(a - t_z) + nc\right]}{(m+n)(b+\xi)} \tag{10}$$

and account for (10) to turn the fuel demand and supply functions (5) into

$$e_h^d = \frac{a_h - P(t_z, c) - t_z}{b} \quad \text{for } h = m, n \quad \text{and} \quad e_m^s = e_n^s = \frac{P(t_z, c)}{\xi}. \tag{11}$$

The equilibrium welfare of a country of group $H = M, N$ is

$$W^h(t_z, c, \rho) := V^h\left(e_h^d\right) + X\left(e_h^s\right) + P(t_z, c)\left(e_h^s - e_h^d\right) - D^h\left[(m+n)e_h^s\right] \quad \text{for } h = m, n, \tag{12}$$

where e_m^d, e_n^d, e_m^s and e_n^s satisfy (11). The maximization of the grand coalition's aggregate welfare with respect to t_z yields the socially optimal tax rate

$$\begin{aligned} t_z^* = T(c,\rho) \quad &:= \quad \arg\max_{t_z}\left[mW^m(t_z, c, \rho) + nW^n(t_z, c, \rho)\right] \\ &= \quad \frac{\left[(m+n)a + nc\right]\left[m\delta + n(\delta+\rho)\right]}{b + \xi + (m+n)\left[m\delta + n(\delta+\rho)\right]}. \end{aligned} \tag{13}$$

[15] And for some given parameters $a, b, m, n, \xi, \bar{x}$ in the set of feasible parameters.

[16] Maximization of aggregate welfare requires equalizing marginal willingness-to-pay for fuel across countries. The FOC of maximizing $V^h\left(e_i^d\right) + x_i^d$ subject to $x_i^d + (p + t_i)e_i^d = y$ is $V^h_{e_i^d} = p + t_i$. Hence, $V^m_{e_i^d} = V^n_{e_j^d}$ for all $i \in M$ and $j \in N$, if and only if $t_i = t_j$ for all $i, j \in M \cup N$.

We insert t_z^* from (13) in (12) and finally get[17,18]

$$w_h^* := W^h\left(t_z^*, c, \rho\right) = W^h\left[T(c,\rho), c, \rho\right] := \mathcal{W}^{*h}(c,\rho) \quad \text{for } h = m, n. \tag{14}$$

Next, we derive the welfare $\bar{w}_f = \bar{\mathcal{W}}^f(c,\rho)$ of an f-country, $f = m, n$. As above, we assign to all coalition countries the same tax rate t_z and denote by t_f the f-country's tax rate. Accordingly, the welfare (9) now reads $W^f(\mathbf{t}, c, \rho) = W^f(t_f, t_z, c, \rho)$ for the f-country and $W^h(\mathbf{t}, c, \rho) = W^{hf}(t_f, t_z, c, \rho)$ for a coalition country of group $H = M, N$, if the coalition size is $m + n - 1$ and if the f-country belongs to group $F = M, N$. Since the f-country participates in international trade, all countries face the same world price of fuel, p. We readily derive $e_f^s = \frac{p}{\xi}$ and $e_f^d = \frac{a_f - p - t_f}{b}$ for the f-country and $e_{hf}^s = \frac{p}{\xi}$ and $e_{hf}^d = \frac{a_h - p - t_z}{b}$ for the coalition countries.[19]

Combining these equations with (3) results in the equilibrium fuel price

$$p = \bar{P}(t_f, t_z, c) := \frac{\xi\left[(m+n)a + nc - t_f - (m+n-1)t_z\right]}{(m+n)(b+\xi)}.$$

Accounting for $\bar{P}(t_f, t_z, c)$ yields the fuel supply functions,

$$e_f^s = e_{mf}^s = e_{nf}^s = \frac{\bar{P}(t_f, t_z, c)}{\xi} =: e^s, \tag{15}$$

the fuel demand of a coalition country of group H,

$$e_{hf}^d = \frac{a_h - \bar{P}(t_f, t_z, c) - t_z}{b} \quad \text{for } h = m, n, \tag{16}$$

and the fuel demand of the outside f-country,

$$e_f^d = \frac{a_f - \bar{P}(t_f, t_z, c) - t_f}{b} \quad \text{for } h = m, n. \tag{17}$$

This information allows specifying the f-country's welfare as

$$\bar{W}^f(t_f, t_z, c, \rho) := V^f\left(e_f^d\right) + X(e^s) + \bar{P}(t_f, t_z, c)\left(e^s - e_f^d\right) - D^f\left[(m+n-1)e^s\right]. \tag{18}$$

[17] It is also worth noting that in the socially optimal market equilibrium of the grand coalition the exports and imports of fuel are $e_m^{d*} - e^{s*} = nc/(m+n)b$ and $e_n^{d*} - e^{s*} = nc/(m+n)b$. Hence the h-countries with parameter $a_h = \min[a_m, a_n]$ export fuel (and import the consumption good X), if $c = a_n - a_m \neq 0$, and there is no trade, if $c = 0$.

[18] $\mathcal{W}^{*h}(c)$ is a formidably complex expression of parameters. First, we substitute t_z with t_z^* in (10), then we insert $P\left(t_z^*, c\right)$ into (12), and finally we plug the resultant quite big terms for e^s and e_h^d from (11) into (12) at several places.

[19] The notation e_{hf}^s $\left[e_{hf}^d\right]$ stands for the fuel supply [demand] of a country of group $H = M, N$ in the coalition of size $m + n - 1$, if the f-country belongs to group $F = M, N$.

To prepare for the calculation of the Nash equilibrium, we also need to calculate the welfare of h-countries, $h = m, n$, in the coalition of size $m + n - 1$ when the f-country belongs to group $F = M, N$. These countries' welfare is

$$\bar{W}^{hf}(t_f, t_z, c, \rho) := V^h\left(e^d_{hf}\right) + X(e^s) + \bar{P}(t_f, t_z, c)\left(e^s - e^d_{hf}\right) - D^h\left[(m+n-1)e^s\right], \quad (19)$$

where e^s and e^d_{hf} satisfy (15), (16) and (17), respectively. The aggregate welfare of the coalition is

$$\bar{W}^{zf}(t_f, t_z, c, \rho) := \begin{cases} (m-1)\bar{W}^{mf}(t_f, t_z, c, \rho) + n\bar{W}^{nf}(t_f, t_z, c, \rho) & \text{if } f = m, \\ m\bar{W}^{mf}(t_f, t_z, c, \rho) + (n-1)\bar{W}^{nf}(t_f, t_z, c, \rho) & \text{if } f = n. \end{cases} \quad (20)$$

The coalition and the f-country are players in a non-cooperative game with payoffs $\bar{W}^{zf}$ (t_f, t_z, c, ρ) and $\bar{W}^f(t_f, t_z, c, \rho)$ whose strategies are t_z and t_f, respectively. Our solution concept is the Nash equilibrium that consists of the (unique) pair of tax rates $\left[\bar{t}_f = \bar{T}^f(c, \rho), \bar{t}_{zf} = \bar{T}^{zf}(c, \rho)\right]$ satisfying $\bar{W}^f(\bar{t}_f, \bar{t}_{zf}, c, \rho) \geq \bar{W}^f(t_f, \bar{t}_{zf}, c, \rho)$ for all t_f and $\bar{W}^{zf}(\bar{t}_f, \bar{t}_{zf}, c, \rho) \geq \bar{W}^{zf}(\bar{t}_f, t_z, c, \rho)$ for all t_z. In Appendix A we calculate the Nash equilibrium tax rates and determine the Nash equilibrium welfare of the f-country as

$$\bar{w}_f := \bar{W}^f\left(\bar{t}_f, \bar{t}_{zf}, c, \rho\right) = \bar{W}^f\left[\bar{T}^f(c, \rho), \bar{T}^{fz}(c, \rho), c, \rho\right] =: \bar{\mathcal{W}}^f(c, \rho). \quad (21)$$

4 Stability of the grand coalition: Size and asymmetry of climate damage ($\delta \uparrow\downarrow, \rho \uparrow\downarrow, c = 0$)

Throughout this section, we assume that consumer preferences for fuel and for the consumption good are the same in all countries ($c = 0$), whereas the climate damage may differ across countries ($\rho \neq 0$). Under these conditions, the grand coalition is stable [unstable], if the welfare difference $\bar{\mathcal{D}}^f(c = 0, \rho) := \mathcal{W}^{*h}(c = 0, \rho) - \bar{\mathcal{W}}^f(c = 0, \rho)$ is non-negative for $f = h = m, n$. To avoid clutter, we will suppress the qualifier $c = 0$ in the present section and write $\bar{\mathcal{D}}^f(\rho)$ short for $\bar{\mathcal{D}}^f(c = 0, \rho)$. Before we investigate the case of climate damage differences ($\rho \neq 0$), it is useful to recall the benchmark result of Eichner and Pethig (2014) on the stability of the grand coalition in the special case of uniform climate damage.

Proposition 1. *(Eichner and Pethig 2014)*
In the world economy with identical countries ($c = 0$ and $\rho = 0$), the grand coalition is stable, if and only if

$$G(b, \delta, m+n, \xi) := [4(m+n-1) + k]\left[\xi(\xi + b) - (m+n)^2 b\delta\right] - k \cdot (\xi + b)^2 \geq 0, \quad (22)$$

where $k := (m+n)^2(m+n-3)$.

Proposition 1 is a sharp result because the inequality (22) is a necessary and sufficient condition for the stability of the grand coalition. If identical countries sign a global IEA, each country obviously enjoys the same socially optimal welfare $\mathcal{W}^{*m}(0,0) \equiv \mathcal{W}^{*n}(0,0)$, and there is no trade. If the coalition consists of all countries but one, trade takes place between the f-country and the coalition of size $n+m-1$, but not within the coalition. Since the function G in (22) is decreasing in δ and increasing in ξ, the stability of the grand coalition is the more likely, the smaller the climate damage parameter δ and the larger the extraction cost parameter[20] ξ. The link between δ and climate damage is obvious. Turning to marginal extraction costs, we find that the more expensive fuel is due to high marginal extraction costs, the lower is fuel consumption and the lower are total emissions and climate damage. Since the role of the parameters b and $m+n$ for the inequality (22) is less obvious, we find it most promising to focus on the impact of exogenous variations of the parameters δ and ξ. In what follows, we combine variations of the damage difference $\rho = \delta_n - \delta$ with variations of the level of damage, δ, and in Section 5 below we combine variations of the preference difference $c = a_n - a$ with variations of the parameters a and ξ.

The question we address now is whether the conditions for the stability of the grand coalition improve, when the climate damage differs across groups ($\rho \neq 0$). Since it is not possible to derive informative general results due to the complexity of the analytical approach, we resort to numerical analysis and proceed in two steps in the remaining part of Section 4. First, we determine the (in)stability of the grand coalition in the Samples 1 and 2 of Table 1 that differ only with respect to the size of the parameter δ and include the continuous variation of the damage asymmetry parameter ρ in the interval[21] $[\rho_{\min}, \rho_{\max}]$, the set of all values of ρ, for which the equilibrium demands and supplies of fuel are positive.[22] After that, we explore the conditions for the stability of the grand coalition on the set of feasible parameters (ρ, δ) - that includes the Samples 1 and 2 - keeping all other parameters

[20] To interpret the production parameter ξ introduced in (1), observe that the extraction e_i^s of fossil fuel causes the extraction costs,

$$\bar{x} - X\left(e_i^s\right) = \bar{x} - \left[\bar{x} - \frac{\xi}{2}\left(e_i^s\right)^2\right] = \frac{\xi}{2}\left(e_i^s\right)^2,$$

expressed in units of the consumption good, because $X\left(e_i^s\right) = \bar{x}$ for $e_i^s = 0$ and $x_i^s = X\left(e_i^s\right) < \bar{x}$ for $e_i^s > 0$. Therefore, ξ is a measure of marginal fuel extraction costs, ξe_i^s.

[21] For details see Appendix B. We keep our analysis within that interval, because negative quantities are an artefact of the linear-quadratic functional forms (1) and (2), and zero consumption of fuel is highly unrealistic. Positive values of the demand and supply of the consumption good are secured by choosing a sufficiently high value of the parameter $\bar{x}$. To avoid clutter, we write $[\rho_{\min}, \rho_{\max}]$ instead of $[\rho_{\min}(\delta), \rho_{\max}(\delta)]$. Also, we set $\rho_{\min}(\delta) = -\delta$ and thus avoid dealing with climate change benefits (= negative climate damage).

[22] We use the term sample rather than example because each sample characterizes the outcome along the asymmetry dimension ρ (and later along the asymmetry dimension c).

as specified in Table 1.

	$\delta_m = \delta$	$a_m = a$	$m = n$	b	ξ	$\rho = \delta_n - \delta$	$c = a_n - a$
Sample 1	$1,600$	100	5	$2,000$	$100,000$	$\in [\rho_{\min}, \rho_{\max}]$	0
Sample 2	3	100	5	$2,000$	$100,000$	$\in [\rho_{\min}, \rho_{\max}]$	0

Table 1: Degree of asymmetry and level of climate damage: Samples 1 and 2

There is a rationale of choosing a very large value of δ in Sample 1 and a very small value in Sample 2. Inspection of (22) shows that $G_\delta < 0$ such that for given parameters $b, m+n$ and ξ there exists $\tilde{\delta} > 0$ such that $G(b, \delta, m+n, \xi) \gtreqless 0 \iff \delta \lesseqgtr \tilde{\delta}$. For the parameters $m+n = 10, b = 2,000$ and $\xi = 100,000$ which are common to the Samples 1 and 2 we calculate $\tilde{\delta} = 1,524$. Hence, in view of Proposition 1 the relevant difference between the Samples 1 and 2 is that for $\rho = 0$ the grand coalition is unstable in Sample 1, but stable in Sample 2.

We now proceed to determine for both samples the sign of the welfare difference $\bar{\mathcal{D}}^f(\rho) = \mathcal{W}^{*h}(\rho) - \bar{\mathcal{W}}^f(\rho)$ for $f = h = m, n$ and for all $\rho \in [\rho_{\min}, \rho_{\max}]$. That requires focusing on the switching points, say $\bar{\rho}_f$, defined by $\bar{\mathcal{D}}^f(\bar{\rho}_f) = 0$ or, equivalently, by $\mathcal{W}^{*h}(\bar{\rho}_f) = \bar{\mathcal{W}}^f(\bar{\rho}_f)$ for $f = h = m, n$, at which the welfare difference $\bar{\mathcal{D}}^f(\rho)$ turns from negative to positive or vice versa.

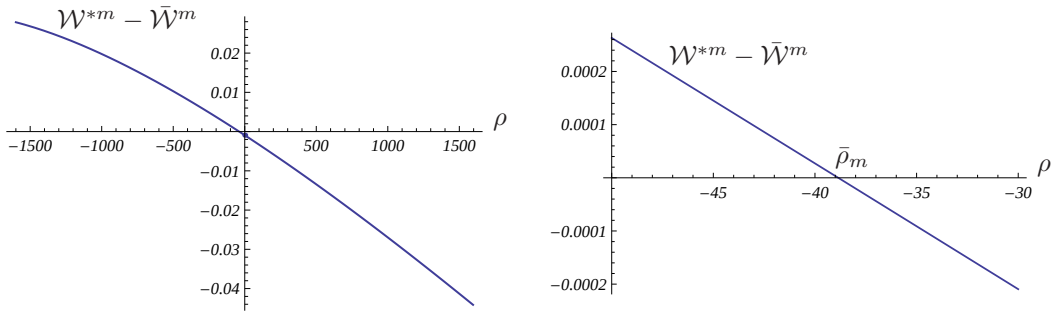

Figure 1: Welfare difference function $\bar{\mathcal{D}}^m = \mathcal{W}^{*m} - \bar{\mathcal{W}}^m$ in Sample 1 ($\delta = 1,600$)

The Figures 1 and 2 depict the graphs of the welfare difference functions $\bar{\mathcal{D}}^f$ for $f = m$ and for $f = n$, respectively, in Sample 1. Their right-hand side panels present enlarged sections of these curves in the close neighborhood of $\rho = 0$. In Figure 1, there exists $\bar{\rho}_m$ satisfying $\bar{\mathcal{D}}^m(\bar{\rho}_m) = 0$ and $\bar{\mathcal{D}}^m_\rho(\bar{\rho}_m) < 0$ and in Figure 2 there is $\bar{\rho}_n$ satisfying $\bar{\mathcal{D}}^n(\bar{\rho}_n) = 0$

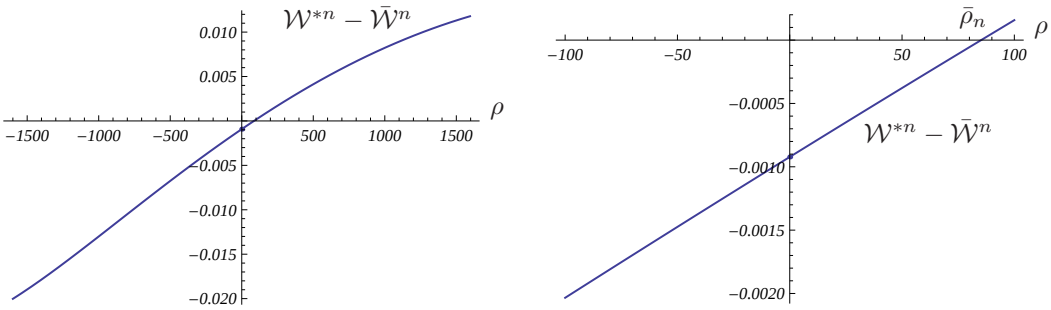

Figure 2: Welfare difference function $\bar{\mathcal{D}}^n = \mathcal{W}^{*n} - \bar{\mathcal{W}}^n$ in Sample 1 ($\delta = 1,600$)

and $\bar{\mathcal{D}}^n_\rho(\bar{\rho}_n) > 0$. Our numerical calculations yield the precise numbers: $\bar{\rho}_m = -38.85$ and $\bar{\rho}_n = 85.22$. Since $\rho_{\min} = -1600$ and $\rho_{\max} = 1596$, the ranking is $\rho_{\min} < \bar{\rho}_m < 0 < \bar{\rho}_n < \rho_{\max}$.

m-country free rides:	no	yes			
n-country free rides:	yes				no
Grand coalition is:	unstable				
Parameter ρ:	$\rho_{\min}$	$\bar{\rho}_m$	$\rho = 0$	$\bar{\rho}_n$	$\rho_{\max}$

Table 2: Stability of the grand coalition in Sample 1 ($\delta = 1,600$)

Table 2 accounts for the correct ranking $\rho_{\min} < \bar{\rho}_m < 0 < \bar{\rho}_n < \rho_{\max}$, but not for the exact numbers $\rho_{\min}, \bar{\rho}_n, \bar{\rho}_m$ and $\rho_{\max}$, for convenience of exposition.[23] It shows that an f-country of either group finds it advantageous to join the coalition in some range of ρ but, unfortunately, these ranges do not overlap. Hence the grand coalition is unstable for all feasible values of ρ. Recall that in Sample 1 the grand coalition is unstable in case of symmetry ($\rho = 0$) according to Eichner and Pethig (2014). Hence, the message of Sample 1 is that, given the parameters underlying that sample, asymmetric climate damage fails to turn the unstable grand coalition into a stable one - no matter how strong the asymmetry is.

Next consider the Figures 3 and 4 that are analogues of the Figures 1 and 2 for Sample 2. In Figure 3, there exists $\bar{\rho}_m = 2.15$ satisfying $\bar{\mathcal{D}}^m(\bar{\rho}_m) = 0$ and $\bar{\mathcal{D}}^m_\rho(\bar{\rho}_m) < 0$ and in

[23] This qualification also applies for the Tables 3, 5, 6 and 7 below.

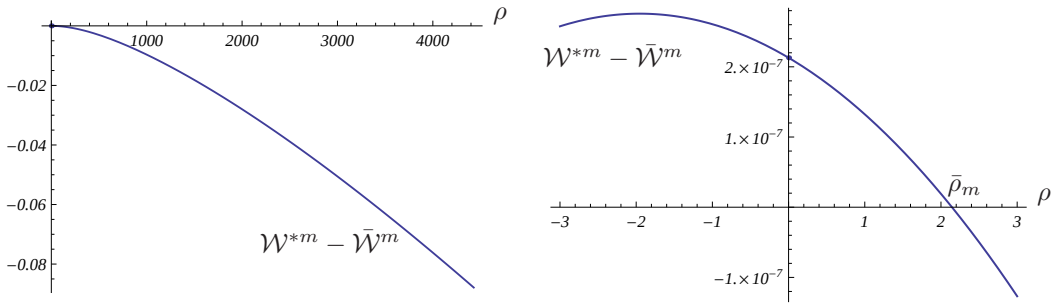

Figure 3: Welfare difference function $\bar{\mathcal{D}}^m$ in Sample 2 ($\delta = 3$)

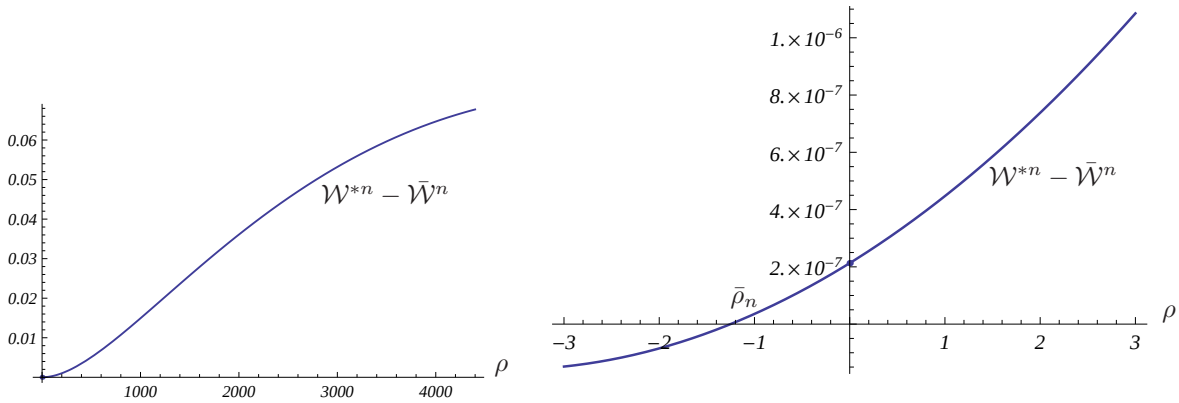

Figure 4: Welfare difference function $\bar{\mathcal{D}}^n$ in Sample 2 ($\delta = 3$)

Figure 4 there is $\bar{\rho}_n = -1.25$ satisfying $\bar{\mathcal{D}}^n(\bar{\rho}_n) = 0$ and $\bar{\mathcal{D}}^n_\rho(\bar{\rho}_n) > 0$. Since $\rho_{\min} = -3$ and $\rho_{\max} = 4,429$, the ranking is $\rho_{\min} < \bar{\rho}_n < 0 < \bar{\rho}_m < \rho_{\max}$. We conclude that the grand coalition is stable in the interval $[\bar{\rho}_n, \bar{\rho}_m]$ and unstable in the intervals $[\rho_{\min}, \bar{\rho}_n[$ and $]\bar{\rho}_m, \rho_{\max}]$, as conveniently illustrated in Table 3.

m-country free rides:	no			yes
n-country free rides:	yes	no		
Grand coalition is:	unstable	stable		unstable
Parameter ρ:	$\rho_{\min}$	$\bar{\rho}_n$	$\rho = 0$	$\bar{\rho}_m$	$\rho_{\max}$

Table 3: Stability of the grand coalition in Sample 2 ($\delta = 3$)

Note that $\rho = 0$ is in the interior of the stability interval $[\bar{\rho}_n, \bar{\rho}_m]$. Hence, the thrust of Sample 2 is that the grand coalition remains stable as long as the climate damage parameter of n-countries does not deviate too much from the climate damage parameter of m-countries. If the climate damage asymmetry $|\rho|$ between countries of the groups M and N is sufficiently large, the grand coalition becomes unstable. That means, growing climate damage asymmetry eventual destabilizes the grand coalition, that was stable in cases of symmetry and low degrees of asymmetry.

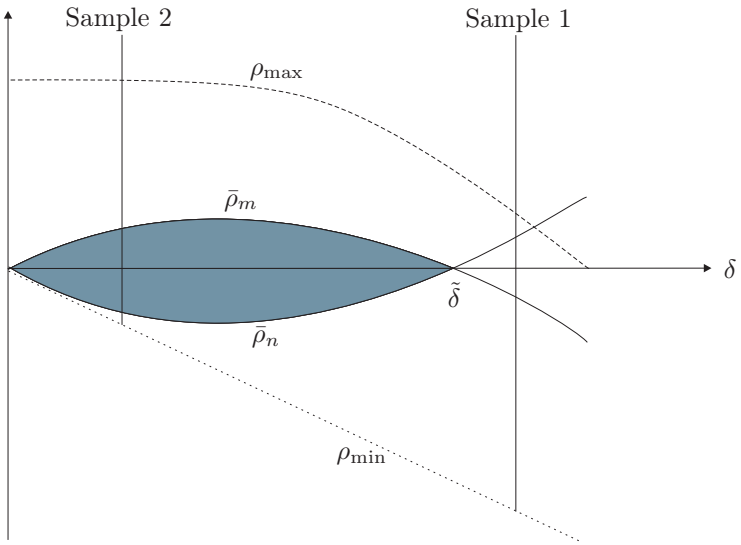

Figure 5: Stability of the grand coalition in the (δ, ρ) parameter space

The Samples 1 and 2 analyze how the conditions for the stability of the grand coalition depend on two dimensions: the size (δ) and asymmetry (ρ) of climate damage. We have varied the parameter ρ continuously, but assigned only two specific values to the parameter δ. Figure 5 is based on the parameters $a = 100$, $m = n = 5$, $b = 2,000$, $\xi = 100,000$ and $c = 0$ which are common to the Samples 1 and 2 and illustrates the outcome of continuous variations of both parameters. It shows how the boundary points $\rho_{\min}$ and $\rho_{\max}$ and the switching points $\bar{\rho}_m$ and $\bar{\rho}_n$ depend on δ and provides three pieces of interesting information. First, it illustrates Proposition 1. For $c = \rho = 0$ and given parameters b, $m + n$ and ξ, there is $\tilde{\delta} > 0$ such that the grand coalition is stable [unstable], if $\delta \leq \tilde{\delta}$ [$\delta > \tilde{\delta}$]. Second, it illustrates the results of the Samples 1 and 2, because the threshold $\tilde{\delta}$ in Figure 5 takes on the value $\tilde{\delta} = 1,524$. Since Sample 1 [Sample 2] satisfies $1,600 = \delta > \tilde{\delta}$ [$3 = \delta < \tilde{\delta}$], the information provided by Sample 1 in Table 2 [by Sample 2 in Table 3] corresponds to the information of a vertical line in Figure 5 that intersects the δ-axis to the right [left] of $\tilde{\delta}$. Finally, it generalizes Proposition 1 as well as the Samples 1 and 2, because it demonstrates that in qualitative terms the result of Sample 1 [Sample 2] holds for all $\delta > \tilde{\delta}$ [$\delta < \tilde{\delta}$]. We summarize these findings in

Result 1. *(Stability depending on size and asymmetry of climate damage)*
Consider the parameters $a = 100$, $b = 2,000$, $c = 0$, $m = n = 5$, $\xi = 100,000$ *and* $\tilde{\delta} = 1,524$ *that are common to the Samples 1 and 2.*

(i) If the climate damage is low ($\delta \leq \tilde{\delta}$)*, the grand coalition is stable in the interval* $[\bar{\rho}_n, \bar{\rho}_m]$ *of low damage asymmetry (containing* $\rho = 0$*) and becomes unstable at sufficiently high degrees of damage asymmetry* ($\rho < \bar{\rho}_n$ *and* $\rho > \bar{\rho}_m$*).*

(ii) If the climate damage is high ($\delta > \tilde{\delta}$)*, the grand coalition is unstable in the entire interval* $[\rho_{\min}, \rho_{\max}]$ *of feasible degrees of climate damage asymmetry.*

Result 1 suggests that the incentives to form a stable grand coalition are declining in both the climate damage asymmetry ($|\rho| \uparrow$) and the severity of climate damage ($\delta \uparrow$).

5 Stability of the grand coalition: Size and asymmetry of fuel demand and size of extraction costs ($a \uparrow\downarrow, c \uparrow\downarrow, \xi \uparrow\downarrow, \rho = 0$)

Now we assume that the climate damage hits all countries in the same way ($\rho = 0$), whereas the preferences for fuel may differ across countries ($c \neq 0$). Under these conditions, the grand coalition is stable [unstable], if the welfare differences[24] $\bar{\mathcal{D}}^f(c) := \mathcal{W}^{*h}(c) - \bar{\mathcal{W}}^f(c)$ for $f = h = m, n$ satisfy $\bar{\mathcal{D}}^f(c) \geq 0$ for $f = m$ and $f = n$. In order to see, how the stability depends on other parameters, we also vary the size of fuel demand (parameter a) and the production parameter ξ. It is clear that Proposition 1 serves as a benchmark again for the special case of uniform preferences (and $\rho = 0$ and ξ constant) - as it did in the previous section for the special case of uniform climate damage and $c = 0$.

	$\delta_m = \delta$	$a_m = a$	$m = n$	b	ξ	$\rho = \delta_n - \delta$	$c = a_n - a$
Sample 3	3	100*)	5	2,000	100,000	0	$\in [c_{\min}, c_{\max}]$
Sample 4	3	100	5	2,000	5,000	0	$\in [c_{\min}, c_{\max}]$
Sample 5	3	100	5	2,000	500	0	$\in [c_{\min}, c_{\max}]$

*) For the case of sample 3, Section 5.1 also studies variations of parameter α

Table 4: Fuel preference asymmetry and extraction costs: Sample 3, 4 and 5

[24] Analogous to our procedure in the previous section, we suppress the qualifier $\rho = 0$ and write $\bar{\mathcal{D}}^f(c)$ short for $\bar{\mathcal{D}}^f(c, \rho = 0)$ etc.

As in the last section, the complexity of the asymmetry analysis leads us to resort to numerical analysis.[25] Specifically, we will explore the impact on the stability of the grand coalition of continuous variations of the fuel demand difference parameter c by means of the Samples 3, 4 and 5 specified in Table 3. These samples differ with respect to the parameter ξ only and in each of them, the parameter c will be varied in the interval[26] $[c_{\min}, c_{\max}]$, the set of all values of c, for which the equilibrium demands and supplies of fuel are positive. Sample 5 equals Sample 1 except that we turn the asymmetry $\rho \neq 0$ and $c = 0$ from Sample 1 into the asymmetry $c \neq 0$ and $\rho = 0$ in Sample 5. Recall from footnote 20 that the production parameter ξ introduced in (1) is a measure of marginal fuel extraction costs. More generally high or low values of ξ indicate high or low extraction costs or more or less fuel abundance. For any given fuel price, an increase of extraction costs or fuel scarcity reduces total fuel extraction and with it total emissions and climate damage. The values of ξ chosen in the Samples 3 - 5 will turn out to cover the three possible qualitatively different outcomes. Analogous to the analysis of the Samples 1 and 2 in Section 4, we will determine for the Samples 3, 4 and 5 the sign of the welfare difference $\bar{\mathcal{D}}^f(c) \geq 0, f = m, n$, on the interval $[c_{\min}, c_{\max}]$. In these samples, the fuel demand parameter a will be kept constant at $a = 100$. In addition, we perform continuous variations of (a, c) to understand the interdependence of the size and asymmetry of fuel demand.

5.1 Size and asymmetry of fuel demand ($c \uparrow\downarrow, a \uparrow\downarrow, \rho = 0$)

We begin with determining the sign of the welfare differences $\bar{\mathcal{D}}^m(c)$ and $\bar{\mathcal{D}}^n(c)$ on $[c_{\min}, c_{\max}]$ for Sample 3 from Table 4 in Figure 6.

In the left panel of Figure 6 we find $\bar{c}_{m1} = -2.57$ and $\bar{c}_{m2} = 47.86$ satisfying $\bar{\mathcal{D}}^m(\bar{c}_{m1}) = \bar{\mathcal{D}}^m(\bar{c}_{m2}) = 0$, $\bar{\mathcal{D}}^m_c(\bar{c}_{m1}) < 0$ and $\bar{\mathcal{D}}^m_c(\bar{c}_{m2}) > 0$. In the right panel of Figure 6 there exist $\bar{c}_{n1} = -32.36$ and $\bar{c}_{n2} = 2.64$ satisfying $\bar{\mathcal{D}}^n(\bar{c}_{n1}) = \bar{\mathcal{D}}(\bar{c}_{n2}) = 0$, $\bar{\mathcal{D}}^n_c(\bar{c}_{n1}) < 0$ and $\bar{\mathcal{D}}^n_c(\bar{c}_{n2}) > 0$. Since $c_{\min} = -41$ and $c_{\max} = 70$, the ranking is[27] $c_{\min} < \bar{c}_{n1} < \bar{c}_{m1} < 0 < \bar{c}_{n2} < \bar{c}_{m2} < c_{\max}$. Table 5 illustrates the conclusion regarding the stability of the grand coalition.

[25] The Appendix C presents a general result that relates to Result 2 below but provides an incomplete characterization only.

[26] The feasibility interval $[c_{\min}, c_{\max}]$ is the analogue to the interval $[\rho_{\min}, \rho_{\max}]$ in Section 4. To avoid clutter, we write $[c_{\min}, c_{\max}]$ instead of $[c_{\min}(\xi), c_{\max}(\xi)]$.

[27] The results for m- and n-countries are not mirror-symmetric, because the preferences for fuel of a free-riding m-country are invariant in c, while the preference parameter $a_n = 100 + c$ of a free-riding n-country obviously depends on c.

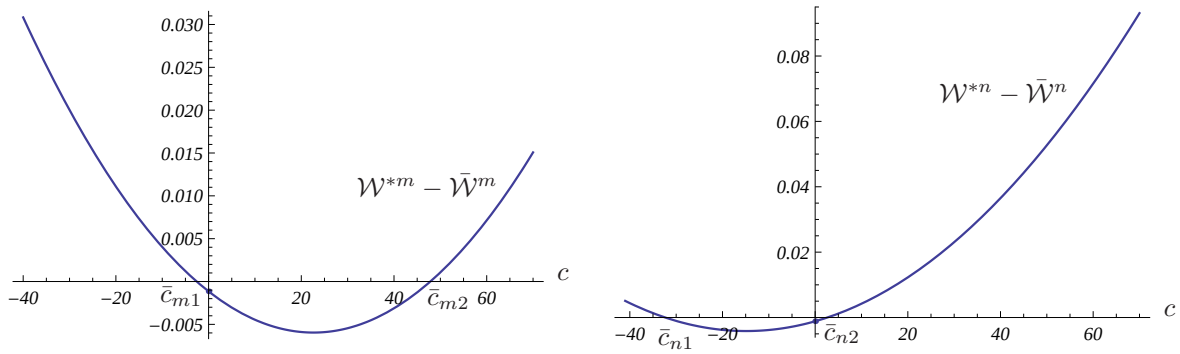

Figure 6: Welfare difference functions $\bar{\mathcal{D}}^m$ and $\bar{\mathcal{D}}^n$ in Sample 3 ($\xi = 5,000$)

m-country free rides:	no		yes				no
n-country free rides:	no	yes				no	
Grand coalition is:	stable	unstable					stable
Parameter c:	$c_{\min}$	$\bar{c}_{n1}$	$\bar{c}_{m1}$	$c = 0$	$\bar{c}_{n2}$	$\bar{c}_{m2}$	$c_{\max}$

Table 5: Stability of the grand coalition in Sample 3 ($\xi = 5,000$)

The striking result is that in Sample 3 the grand coalition is unstable in the interval $]\bar{c}_{n1}, \bar{c}_{m2}[$ of zero, low and medium fuel demand asymmetry $|c|$ and eventually turns from unstable to stable when $|c|$ becomes sufficiently large. Hence asymmetric fuel demand improves the incentives to cooperate. This is so, because increasing asymmetry raises the gains from trade, which in turn reduce the free-rider advantage. To be more specific, we split a country's total welfare into consumption welfare, $V^i(\cdot) + x_i^d$, and climate welfare, $-D^i(\cdot)$.[28] As demonstrated in Eichner and Pethig (2014) for the case $c = 0$, an outside country of either group increases its climate welfare and reduces its consumption welfare by joining the coalition, and the increase of the former falls short of the reduction of the latter under the conditions of Sample 3. If the fuel-demand asymmetry $|c|$ becomes positive, the f-country's consumption welfare is larger than in case of $c = 0$, because of greater gains from trade. Since within the grand coalition trade and trade gains increase with increasing asymmetry $|c|$, the consumption welfare of coalition countries rises. Thus, the advantage of the f-country's trade gains over those of countries in the grand coalition diminishes and

[28]The graphs showing how the consumption and climate welfare of f-countries and coalition countries depend on fuel-demand asymmetry can be obtained from the authors upon request.

eventually vanishes when the fuel-demand asymmetry $|c|$ becomes sufficiently high. In fact, for very large $|c|$ the f-country raises its consumption welfare *and* climate welfare by joining the coalition such that the grand coalition is stable as shown in Table 5.

In Sample 3 the parameter a is fixed at the level $a = 100$. An important question is how the result (of Table 5) changes by considering continuous variations of the parameters c and a while keeping unchanged all other parameters of Sample 3. To answer that question we show in Appendix B that all values $c_{\min}, \bar{c}_{m1}, \bar{c}_{m2}, \bar{c}_{n1}, \bar{c}_{n2}$ and $c_{\max}$ are linear functions of the parameter a. $c_{\min}, \bar{c}_{m1}$ and $\bar{c}_{n1}$ are decreasing in a and $\bar{c}_{m2}, \bar{c}_{n2}$ and $c_{\max}$ are increasing in a. The implication for the stability of the grand coalition in the (a, c) space is readily illustrated in Figure 7. The vertical line in Figure 7 illustrates the intervals of stability and instability of Sample 3 in a way that is different but equivalent to Table 5. Other parameters being unchanged, exogenous variations of the parameter a expand or shrink the feasibility interval $[c_{\min}, c_{\max}]$ in Figure 7 and with it the sub-intervals of stability and instability. This is also true for the Samples 4 and 5 we study below which is why we will not go through the exercise analogous to Figure 7 again in the subsequent analysis Section 5.2.

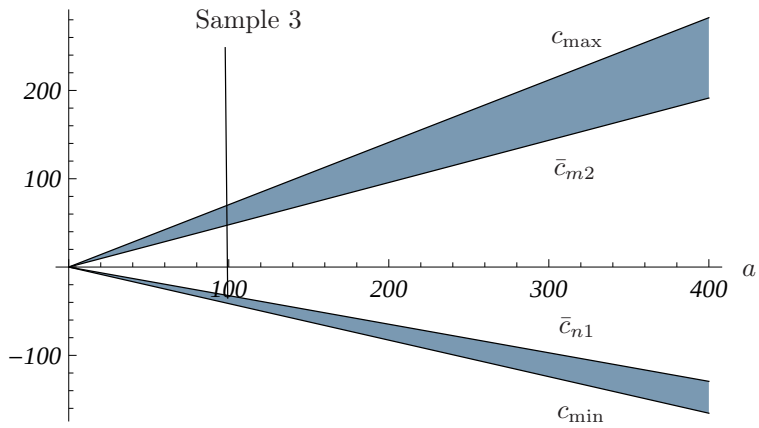

Figure 7: Stability of the grand coalition in the (a, c) parameter space, when all other parameters are as in Sample 3

To summarize, for any given value of a, the pattern of intervals with a stable or unstable grand coalition is shown as in Table 5 with the qualification that all intervals expand [shrink], if the parameter a becomes larger [smaller]. As an implication, the asymmetry $|c|$ necessary to turn instability into stability is increasing in the parameter a. Hence, if we start with a tuple (a, c) for which the grand coalition is stable and increase a while keeping c constant

we find some value $\tilde{a} > a$ such that the grand coalition is unstable for all (a', c) satisfying $a' > \tilde{a}$.

5.2 Asymmetric fuel demand and the abundance of fossil fuel

Knowing the result of Sample 3 (Table 5 and Figure 7) with its medium-size extraction costs, we are now interested in the outcome of cases of lower (Sample 4) and higher (Sample 5) extraction costs.

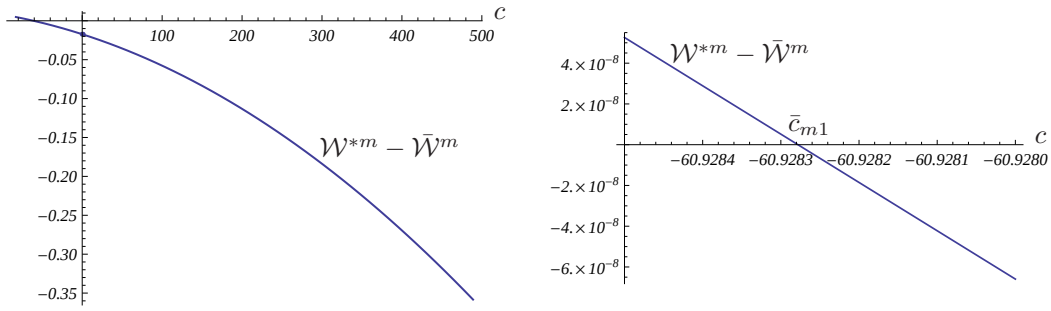

Figure 8: Welfare difference function $\bar{\mathcal{D}}^m$ in Sample 4 ($\xi = 500$)

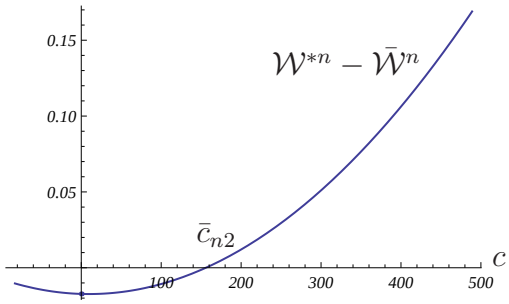

Figure 9: Welfare difference function $\bar{\mathcal{D}}^n$ in Sample 4 ($\xi = 500$)

Consider the Figures 8 and 9 that correspond to Sample 4. In Figure 8 there exists $\bar{c}_{m1} = -60.928$ satisfying $\bar{\mathcal{D}}^m(\bar{c}_{m1}) = 0$ and $\bar{\mathcal{D}}^m_c(\bar{c}_{m1}) < 0$. In Figure 9 there exists[29] $\bar{c}_{n2} = 155.94$ satisfying $\bar{\mathcal{D}}^n(\bar{c}_{n2}) = 0$ and $\bar{\mathcal{D}}^n_c(\bar{c}_{n2}) > 0$. Since $c_{\min} = -83$ and $c_{\max} = 489$,

[29] There also exists $\bar{c}_{n1} < 0$ satisfying $\bar{\mathcal{D}}^n(\bar{c}_{n1}) = 0$, but $\bar{c}_{n1} < c_{\min}$.

the ranking is[30] $c_{\min} < \bar{c}_{m1} < 0 < \bar{c}_{n2} < c_{\max}$. We conclude and illustrate in Table 6 that there is no $c \in [c_{\min}, c_{\max}]$ such that $\bar{\mathcal{D}}^m(c) \geq 0$ and $\bar{\mathcal{D}}^n(c) \geq 0$.

m-country free rides:	no	yes			
n-country free rides:	yes				no
Grand coalition is:	unstable				
Parameter c:	$c_{\min}$	$\bar{c}_{m1}$	$c = 0$	$\bar{c}_{n2}$	$c_{\max}$

Table 6: Stability of the grand coalition in Sample 4 ($\xi = 500$)

Recall that due to Proposition 1 the grand coalition is unstable in both Sample 3 and Sample 4 for the special case $c = 0$. However, lower (marginal) extraction costs in Sample 4 tend to increase total fuel consumption and hence affect climate damage in a similar way as an increase of the damage parameter δ. Growing fuel demand asymmetry $|c|$ may reduce the free-rider advantage in Sample 4, as in Sample 3, but not enough to turn the grand coalition from unstable to stable.

Sample 5 exhibits extraction costs that are higher than in the Samples 3 and 4 and so high, in fact, that in contrast to the Samples 3 and 4 the grand coalition is stable in Sample 5 for $c = 0$ (Eichner and Pethig 2014). The welfare difference curves of Sample 5 are plotted in Figures 10 and 11.

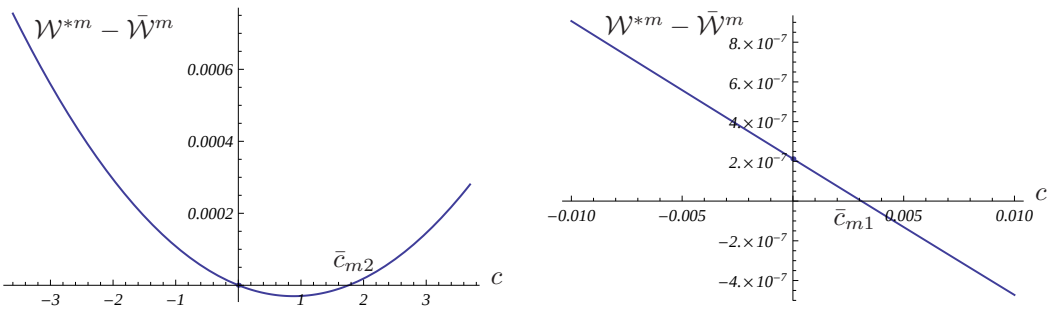

Figure 10: Welfare difference function $\bar{\mathcal{D}}^m$ in Sample 5 ($\xi = 100,000$)

[30]The results for m- and n-countries are not mirror-symmetric, because the preferences for fuel of a free-riding m-country are invariant in c, while the preference parameter $a_n = 100 + c$ of a free-riding n-country obviously depends on c.

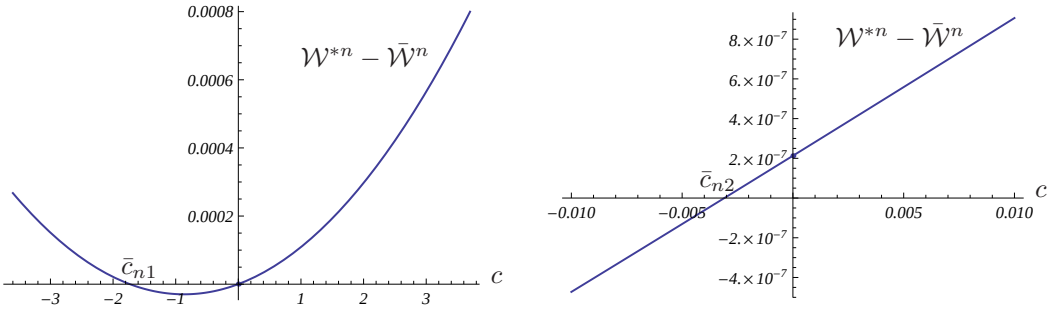

Figure 11: Welfare difference function $\bar{\mathcal{D}}^n$ in Sample 5 ($\xi = 100,000$)

In Figure 10, there exist $\bar{c}_{m1} = 0.0031$ and $\bar{c}_{m2} = 1.761$ satisfying $\bar{\mathcal{D}}^m(\bar{c}_{m1}) = \bar{\mathcal{D}}^m(\bar{c}_{m2}) = 0, \bar{\mathcal{D}}^m_c(\bar{c}_{m1}) < 0$ and $\bar{\mathcal{D}}^m_c(\bar{c}_{m2}) > 0$. In Figure 11 there exist $\bar{c}_{n1} = -1.73$ and $\bar{c}_{n2} = -0.031$ satisfying $\bar{\mathcal{D}}^n(\bar{c}_{n1}) = \bar{\mathcal{D}}^n(\bar{c}_{n2}) = 0$, $\bar{\mathcal{D}}^n_c(\bar{c}_{n1}) < 0$ and $\bar{\mathcal{D}}^n_c(\bar{c}_{n2}) > 0$. Since $c_{\min} = -3.6$ and $c_{\max} = 3.7$, the ranking is $c_{\min} < \bar{c}_{n1} < \bar{c}_{n2} < 0 < \bar{c}_{m1} < \bar{c}_{m2} < c_{\max}$. Table 7 illustrates the conclusion with regard to the stability of the grand coalition.

m-country free rides:	no			yes	no
n-country free rides:	no	yes	no		
Grand coalition is:	stable	unstable	stable	unstable	stable

Parameter c: $c_{\min}$ $\bar{c}_{n1}$ $\bar{c}_{n2}$ $c = 0$ $\bar{c}_{m1}$ $\bar{c}_{m2}$ $c_{\max}$

Table 7: Stability of the grand coalition in Sample 5 ($\xi = 100,000$)

The last row of Table 7 shows that the stability of the grand coalition either requires sufficiently small or sufficiently large asymmetries $|c|$. The common feature of the Samples 2 and 5 is that the grand coalition is stable in case of full symmetry and that the grand coalition remains stable if the asymmetries ($|\rho| \neq 0$ in Sample 2 and $|c| \neq 0$ in Sample 5) are small. The Samples 2 and 5 differ, however, because large asymmetries restore stability in Sample 5 but not in Sample 2.[31] In Sample 5 the effect of asymmetry on stability is non-monotonic. The grand coalition is stable in case of full symmetry, it gets unstable at intermediate degrees of fuel-demand asymmetry and stabilizes again if the asymmetry gets sufficiently large. The explanation for the switch from instability to stability, when the

[31] Our conjecture is that such differences are a somewhat artificial limitation of the linear-quadratic model. For example, if in Figure 12 one would have $c_{\max} > \bar{c}_{m2}$ and $c_{\min} < \bar{c}_{n1}$ for all $\xi \in [0, \bar{\xi}]$ the results of the Samples 3 and 4 would be the same in qualitative terms.

asymmetry increases from intermediate to large values, is as in the context of Sample 3, where we investigated how the consumption welfare $\left[V^i(\cdot) + x_i^d\right]$ and the climate welfare $[-D^i(\cdot)]$ change with variations in fuel-demand asymmetry. In Sample 5 there exists a (small) positive level of c at which the gains of trade and the consumption welfare within the grand coalition attain their minimum. Consumption welfare within the grand coalition increases with convexity around this minimum. In contrast, the f-country's gains of trade and their consumption welfare have their minimum at $c = 0$. Starting from a stable coalition at $c = 0$ and increasing the asymmetry parameter c increases the outsider's gains of trade. In contrast, within the coalition the gains of trade initially decrease and then increase in c. In the range at which the gains of trade within the coalition diminish the grand coalition gets unstable. Further increasing c the trade gains within the coalition increase faster than the trade gains of the f-country which reduces the free-rider benefit of the f-country and stabilizes the grand coalition (analoguous to Sample 3).

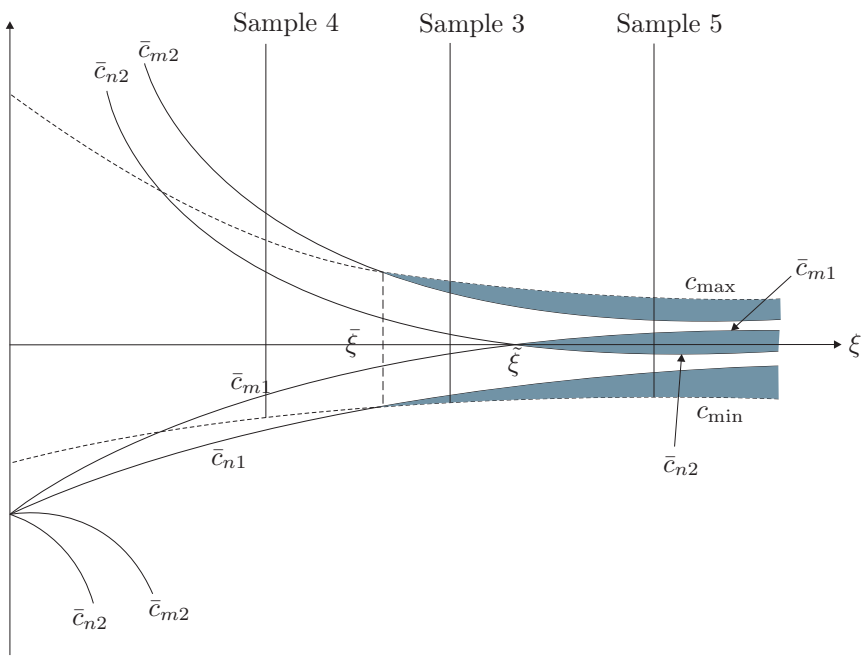

Figure 12: Stability of the grand coalition in the (c, ξ) parameter space

The strong message of the Samples 3 - 5 is that the extraction costs (parameter ξ) is crucial for the stability of the grand coalition. To strengthen that important finding we will generalize the Samples 3 - 5 by characterizing the parameter space (c, ξ) for the parameters $\delta = 3$, $a = 100$, $m = n = 5$, $b = 2,000$ and $\rho = 0$ which are common to the Samples 3 - 5. Figure 12 illustrates the result of that exercise. The shaded areas in Figure 12 mark those subsets of parameters (c, ξ) in which the grand coalition is stable. Note that $\bar{\xi} = 2,504$ and $\tilde{\xi} = 39,186$ such that $\bar{\xi}$ and $\tilde{\xi}$ satisfy $\xi_4 < \bar{\xi} < \xi_3 < \tilde{\xi} < \xi_5$, where ξ_i is the value of ξ in Sample $i = 3, 4, 5$ defined in Table 4. In view of Figure 12 this observation implies that in

qualitative terms the outcome regarding the stability of the grand coalition is the same as in Sample 5 for all $\xi > \tilde{\xi}$, the same as in Sample 3 for all $\xi \in]\bar{\xi}, \tilde{\xi}[$ and the same as in Sample 4 for all $\xi < \tilde{\xi}$. We summarize the findings of Section 5 in

Result 2. *(Stability, asymmetric fuel demand and extraction costs)*
Consider the parameters $a_m = 100$, $b = 2,000$, $m = n = 5$, $\delta = 3$ and $\rho = 0$ which are common to the Samples 3 - 5.

(i) Suppose, the parameter ξ is fixed – as in the Section 3, 4 and 5 – and consider variations of the parameter a. Increasing [reducing] the parameter a expands [shrinks] the intervals on $[\rho_{\min}, \rho_{\max}]$ with a stable or unstable grand coalition without changing the patterns of those intervals.

(ii) There exist parameters $\bar{\xi}$ and $\hat{\xi}$ satisfying $0 < \bar{\xi} < \hat{\xi}$ such that

- *the grand coalition is unstable in the interval $[c_{\min}, c_{\max}]$, if $\xi \in [0, \tilde{\xi}[$;*
- *the grand coalition is stable in the intervals $[\bar{c}_{\min}, \bar{c}_{n1}]$ and $[\bar{c}_{m2}, \bar{c}_{\max}]$, if $\xi \in]\bar{\xi}, \tilde{\xi}[$;*
- *the grand coalition is stable in the intervals $[c_{\min}, \bar{c}_{n1}[$, $[\bar{c}_{m1}, \bar{c}_{n2}]$ and $]\bar{c}_{m2}, c_{\max}]$, if $\xi > \tilde{\xi}$.*

6 Summary and concluding remarks

This paper investigates the formation of stable grand climate coalitions, when climate policy takes the form of carbon emission taxation and when fossil fuel and consumption goods are traded on world markets. We expand the model of Eichner and Pethig (2014) by considering two groups of asymmetric countries. Within groups, countries are alike but across groups, they differ with respect to climate damage or fossil fuel demand. Despite our very simple "linear-quadratic" modeling of the world economy, the analysis gets too complicated to yield informative general results. We therefore proceed with the numerical analysis of five samples, each of which contains continuous changes in one of the asymmetries. We find that if climate damage is low (δ small) and hence the grand coalition is stable in the absence of asymmetries, growing damage asymmetry eventually turns stability into instability (Sample 2). Otherwise (δ large), the grand coalition is unstable at all degrees of asymmetry - including symmetry (Sample 1). These findings suggest that climate damage asymmetry generally discourages cooperation in the grand coalition.

The Samples 3, 4 and 5 investigate fuel-demand asymmetry. These samples yield more complex results that crucially depend on the size of the extraction cost parameter ξ. The

impact of extraction costs, or fossil fuel abundance, on stability is similar to that of the climate damage parameter δ, since increasing extraction costs ($\xi \uparrow$) as well as decreasing climate damage ($\delta \downarrow$) tend to make the grand coalition stable. If fuel is very abundant (Sample 4), stability is not attained, neither with nor without fuel demand asymmetry. However, an integrated view on the Samples 3, 4 and 5 suggests that low degrees of fuel demand asymmetry discourage cooperation whereas higher asymmetry stabilizes the grand coalition (again), if fuel is sufficiently scarce (Sample 3 and 5).

We need to add that the insights provided by these results are limited and therefore must be assessed with care. First, the model is extremely simple, in particular with respect to the additive and quasi-linear utility function (2). Unfortunately, the analytical relief given by the utility function (2) is indispensable for informative results. Even with this relief, sharp results cannot be attained unless one relies on numerical calculations. Second, our characterization of the determinants of (un)stable grand coalitions within our simple model is far from complete. Although we provide significantly more insight than would be gained by a limited number of examples, the samples we selected fail to cover all relevant relations between the stability of the grand coalition and possible country asymmetries. We only consider two groups of countries of equal size and we neither vary all parameters that are uniform across countries nor do we investigate asymmetries among countries with respect to all parameters.

References

Barrett, S. (1994): Self-enforcing international environmental agreements. *Oxford Economic Papers* 46, 878-894.

Barrett, S. (1997). Heterogeneous international environmental agreements. In: Carraro, C. (Ed.), *International Environmental Negotiations*. Edward Elgar, Cheltenham, UK.

Barrett, S. (2001): International cooperation for sale. *European Economic Review* 45, 1835-1850.

Bosello, F., Buchner, B. and C. Carraro (2003): Equity, development and climate change control. *Journal of the European Economic Association* 1, 601-611.

Botteon, M. and C. Carraro (2001): Environmental coalitions with heterogenous countries: burden sharing and carbon leakage, in A. Ulph (ed.): Environmental Policy, *International Agreements and International Trade*, Oxford University Press, New York.

Carraro, C. and D. Siniscalco (1998): International Environmental agreements: incentives

and political economy. *European Economic Review* 42, 561-572.

Chen, X. and A. Woodland (2013): International trade and climate change. *International Tax and Public Finance* 20, 381-413.

Copeland, B.R. and M.S. Taylor (1995): Trade and transboundary pollution, *American Economic Review* 85, 716-737.

D'Aspremont, L., Jacquemin, A., Gabszewicz, J.J and J.A. Weymark (1983): On the stability of collusive price leadership. *The Canadian Journal of Economics* 16, 17-25.

Diamantoudi, E. and E. Sartzetakis (2006): Stable international environmental agreements: An analytical approach. *Journal of Public Economic Theory* 8, 247-263.

Eichner, T. and R. Pethig (2013): Self-enforcing environmental agreements and international trade. *Journal of Public Economics* 102, 37-50.

Eichner, T. und R. Pethig (2014): Self-enforcing environmental agreements and trade: taxes versus caps, CESifo Working paper No. 4954.

Finus, M. (2003): Stability and design of international environmental agreements: The case of transboundary pollution, in: H. Folmer and T. Tietenberg (eds.), *The International Yearbook of Environmental and Reource Economics* 2003/2004, Edward Elgar, Cheltenham, 82 - 158.

Fuentes-Abero, C. and S. Rubio (2010): Can international environmental cooperation be bought? *European Journal of Operations Research* 202, 255-264.

McGinty, M (2007): International environmental agreements among asymmetric nations. *Oxford Economic Papers* 59, 45-62.

Pavlova, Y. and A. de Zeeuw (2013): Asymmetries in international environmental agreements. *Environment and Development Economics* 18, 51-68.

Rubio, S.J. and A. Ulph (2006): Self-enforcing agreements and international trade in greenhouse emission rights. *Oxford Economic Papers* 58, 233-263.

Weikard, H.-P., Finno, M. and J.C. Altamirano-Cabrera (2006): The impact of surplus sharing on the stability of international climate agreements, *Oxford Economic Papers* 58, 209-232.

Appendix

Throughout the Appendix it is assumed that the groups are of equal size ($m = n$).

A: Nash equilibrium tax rates

The f-country maximizes $\bar{W}^f(t_f, t_z, c, \rho)$ from (18) with respect to t_f. The associated first-order condition reads

$$\bar{W}^f_{t_f} = \begin{cases} \frac{\partial e^d_f}{\partial t_f} t_f + (e^s - e^d_f)\frac{\partial \bar{P}}{\partial t_f} - \delta[(m+n)e^s]\frac{\partial e^s}{\partial t_f} = 0 & \text{if } f = m, \\ \frac{\partial e^d_f}{\partial t_f} t_f + (e^s - e^d_f)\frac{\partial \bar{P}}{\partial t_f} - (\delta+\rho)[(m+n)e^s]\frac{\partial e^s}{\partial t_f} = 0 & \text{if } f = n. \end{cases} \tag{A1}$$

The coalition maximizes $\bar{W}^{zf}(t_f, t_z, c, \rho)$ from (20) with respect to t_z. The corresponding first-order condition is

$$\begin{aligned} \bar{W}^{zf}_{t_z} \quad = \quad & \begin{cases} (m-1)\frac{\partial e^d_{mf}}{\partial t_z} t_z + n\frac{\partial e^d_{nf}}{\partial t_z} t_z + [(m+n-1)e^s - (m-1)e^d_{mf} - ne^d_{nf}]\frac{\partial \bar{P}}{\partial t_z} \\ m\frac{\partial e^d_{mf}}{\partial t_z} t_z + (n-1)\frac{\partial e^d_{nf}}{\partial t_z} t_z + [(m+n-1)e^s - me^d_{mf} - (n-1)e^d_{nf}]\frac{\partial \bar{P}}{\partial t_z} \end{cases} \\ & -[(m-1)\delta + n(\delta+\rho)][(m+n)e^s]\frac{\partial e^s}{\partial t_z} = 0 \quad \text{if } f = m, \\ & -[m\delta + (n-1)(\delta+\rho)][(m+n)e^s]\frac{\partial e^s}{\partial t_z} = 0 \quad \text{if } f = n. \end{aligned} \tag{A2}$$

Solving (A1)-(A2) with respect to t_f and t_z yields the Nash equilibrium tax rates:

$$\bar{T}^m(c,0) = \frac{-c\xi^3 + 2b^2(-cn\xi + 8an^3\delta + 4cn^3\delta) + b\xi[16an^3\delta - c(\xi + 2n\xi - 4n^2\delta)]}{4(b+\xi)[2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 4n^2(2n^2-2n-1)\delta]]}, \tag{A3}$$

$$\bar{T}^{mz}(c,0) = \frac{c(2n-1)\xi^3 + 2b^2[8an^3(2n-1)\delta + cn(\xi + 4n^2(2n-1)\delta)] + b\xi[16an^3(2n-1)\delta + c[(4n-1)\xi + 4n^2(4n^2-2n+1)]]}{4(b+\xi)[2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 4n^2(2n^2-2n-1)\delta]]}, \tag{A4}$$

$$\bar{T}^n(c,0) = \frac{c\xi^3 + 2b^2[8an^3\delta + cn(\alpha + 4n^2\delta)] + b\xi[16an^3\delta + c[\xi + 2n\xi + 4n^2(4n-1)\delta]]}{4(b+\xi)[2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 4n^2(2n^2-2n-1)\delta]]}, \tag{A5}$$

$$\bar{T}^{nz}(c,0) = \frac{c(2n-1)\xi^3 + 2b^2[8an^3(2n-1)\delta + c(8n^4\delta - 4n^3\delta - n\xi)] + b\xi[16an^3(2n-1)\delta + c[\xi - 4n\xi + 4n^2(4n^2-2n-1)\delta]]}{4(b+\xi)[2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 4n^2(2n^2-2n-1)\delta]]}, \tag{A6}$$

$$\bar{T}^m(0,\rho) = \frac{abn^2\,[4bn\delta + \xi[2n(2\delta+\rho) - \rho]]}{(\xi+b)\,[2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 2n^2[(4n^2-4n+2)\delta + n(2n-1)\rho]]]}, \tag{A7}$$

$$\bar{T}^{mz}(0,\rho) = \frac{abn^2\,[4bn[(2n-1)\delta + n\rho] + (2n-1)\xi[2n(2\delta+\rho) + \rho]]}{(\xi+b)\,[2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 2n^2[(4n^2-4n+2)\delta + n(2n-1)\rho]]]}, \tag{A8}$$

$$\bar{T}^n(0,\rho) = \frac{abn^2\,[4bn(\delta+\rho) + \xi[2n(2\delta+\rho) + \rho]]}{(\xi+b)\,[2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 2n^2[(4n^2-4n+2)\delta + (2n^2-3n+2)\rho]]]}, \tag{A9}$$

$$\bar{T}^{nz}(0,\rho) = \frac{abn^2\,[4bn[(2n-1)\delta + (n-1)\rho] + (2n-1)\xi[2n(2\delta+\rho) - \rho]]}{(\xi+b)\,[2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 2n^2[(4n^2-4n+2)\delta + (2n^2-3n+2)\rho]]]}. \tag{A10}$$

B: The domain of the functions $\mathcal{W}^{*h}(c,0) : \mathcal{S}^c \to \mathbb{R}$ for $h = m,n$, $\bar{\mathcal{W}}^f(c,0) : \mathcal{S}^c \to \mathbb{R}$ for $f = m,n$, $\mathcal{W}^{*h}(0,\rho) : \mathcal{S}^\rho \to \mathbb{R}$ for $h = m,n$, $\bar{\mathcal{W}}^f(0,\rho) : \mathcal{S}^\rho \to \mathbb{R}$ for $f = m,n$

In the sequel, we restrict the set of parameters such that all fossil fuel demands and supplies and all consumption good supplies and demands are strictly positive.

Suppose first that $c \neq 0$ and $\rho = 0$. Inserting (13) into (11) for $\rho = 0$ we obtain in the first-best equilibrium

$$\begin{aligned} e_m^d > 0 \quad &\iff \quad c < \frac{2ab}{\xi + 4n^2\delta} =: c_1^*, \\ e_n^d > 0 \quad &\iff \quad c > -\frac{2ab}{2b + \xi + 4n^2\delta} =: c_2^*, \\ e_m^s = e_n^s > 0 \quad &\iff \quad c > -2a =: c_3^*. \end{aligned}$$

Observe that $c_1^* > 0$, $c_2^* < 0$ and $c_3^* < 0$. In addition, from $a_n = a + c > 0$ follows $c > -a$.

Inserting the Nash equilibrium tax rates (A3)-(A6) in (15)-(17) for $\rho = 0$ we get in the free-trade equilibrium if country m free rides:

$$
\begin{aligned}
e^s > 0 \quad &\Longleftrightarrow \quad c > -\frac{2a[2bn^2 + \alpha(2n-1)]}{n(2bn+\xi)} =: \bar{c}_1, \\
e^d_{mm} > 0 \quad &\Longleftrightarrow \quad c < \frac{4ab(\xi+b)[2bn^2 + (2n-1)\xi - 4(n-1)n^2\delta]}{(4n-1)\xi^3 + 2b^2n[\xi(2n+1) + 4n^2(2n-1)\delta] + b\xi[(4n^2+6n-1)\xi + 4n^2(4n^2-2n+1)\delta]} =: \bar{c}_2, \\
e^d_{nm} > 0 \quad &\Longleftrightarrow \quad c > -\frac{4ab(\xi+b)[2bn^2 + (2n-1)\xi - 4(n-1)n^2\delta]}{8b^3n^2 + (4n-3)\xi^3 + 2b^2[(6n^2+3n-2)\xi + 4n^2(2n^2-3n+2)\delta] + b\xi[(4n^2+10n-7)\xi + 4n^2(4n^2-6n-3)\delta]} =: \bar{c}_3, \\
e^d_{m} > 0 \quad &\Longleftrightarrow \quad c < \frac{4ab(\xi+b)[2bn^2 + (2n-1)(\xi + 4(n-1)n^2\delta)]}{(2bn+\xi)[(2n-1)\xi^2 + b((2n-1)\xi + 4n^2\delta)]} =: \bar{c}_4.
\end{aligned}
$$

If country n free rides, the free-trade equilibrium is characterized by

$$
\begin{aligned}
e^s > 0 \quad &\Longleftrightarrow \quad c > -\frac{2a[2bn^2 + \xi(2n-1)]}{2bn^2 + (3n-2)\xi} =: \bar{c}_5, \\
e^d_{mn} > 0 \quad &\Longleftrightarrow \quad c < \frac{4ab(\xi+b)[2bn^2 + (2n-1)\xi - 4(n-1)n^2\delta]}{(4n-3)\xi^3 + 2b^2n(2n-1)(\xi + 4n^2\delta) + b\xi[(4n^2+2n-3)\xi + 4n^2(4n^2-2n-1)\delta]} =: \bar{c}_6, \\
e^d_{nn} > 0 \quad &\Longleftrightarrow \quad c > -\frac{4ab(\xi+b)[2bn^2 + (2n-1)\xi - 4(n-1)n^2\delta]}{8b^3n^2 + (4n-1)\xi^3 + 2b^2[(6n^2+5n-2)\xi + 4n^2(2n^2-3n+2\delta)] + b\xi[(4n^2+14n-5)\xi + 4n^2(4n^2-6n+5)]} =: \bar{c}_7, \\
e^d_{n} > 0 \quad &\Longleftrightarrow \quad c > -\frac{4ab(\xi+b)[2bn^2 + 2(n-1)(\xi + 4(n-1)n^2\delta)]}{8b^3n^2 + (2n-1)\xi^3 + 2b^2[(6n^2+3n-2)\xi + 4n^2(4n^2-5n+2)\delta] + b\xi[(4n^2+8n-5)\xi + 4n^2(8n^2-12n+5)\delta]} =: \bar{c}_8.
\end{aligned}
$$

Observe that $\bar{c}_1 < 0$, $\bar{c}_4 > 0$, $\bar{c}_5 < 0$, $\bar{c}_8 < 0$. In the sequel we assume $b + \frac{2n-1}{2n^2}\xi > 2(n-1)\delta$ which is a sufficient condition for $\bar{c}_2 > 0$, $\bar{c}_3 < 0$, $\bar{c}_4 > 0$, $\bar{c}_6 > 0$, $\bar{c}_7 < 0$. Moreover, sufficiently large values of $\bar{x}$ ensure that all final good supplies and demands are strictly positive. It follows that for sufficiently large parameters $\bar{x}$ the domain of the functions $\mathcal{W}^{*h}(c,0)$ and $\bar{\mathcal{W}}^f(c,0)$ is

$$
\mathcal{S}^c := \left\{ (a,b,\xi,n,\delta,c,\rho) | (a,b,\xi,n,\delta,c) \in \mathbb{R}^6_{++}, \rho = 0, n \geq 2, b + \frac{2n-1}{2n^2}\xi > 2(n-1)\delta \wedge c \in [c_{\min}, c_{\max}] \right\}, \tag{B1}
$$

where $c_{\min} := \max\{c_2^*, c_3^*, -a, \bar{c}_1, \bar{c}_3, \bar{c}_5, \bar{c}_7, \bar{c}_8\}$ and $c_{\max} := \min\{c_1^*, \bar{c}_2, \bar{c}_4, \bar{c}_6\}$.

Next, suppose that $c = 0$ and $\rho \neq 0$. In the first-best equilibrium we obtain

$$e_m^d = e_n^d = e_m^s = e_n^s = \frac{a}{\xi + b + 2n^2(2\delta + \rho)} > 0.$$

If country m free rides, the free-trade equilibrium is characterized by

$$e^s = \frac{a[2bn^2 + (2n-1)\xi]}{2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 2n^2[(4n^2-4n+2)\delta + (2n-1)n\rho]]} > 0,$$

$$\begin{aligned} e_{mm}^d = e_{nm}^d > 0 \quad &\Longleftrightarrow \quad \rho < \frac{(\xi+b)[2bn^2 + (2n-1)\xi + 4(n-1)n^2\delta]}{n^2[2bn + (2n-1)\xi]} =: \bar{\rho}_1, \\ e_m^d > 0 \quad &\Longleftrightarrow \quad \rho > -\frac{(\xi+b)[2bn^2 + (2n-1)[\xi + 4(n-1)n^2\delta]]}{n^2(2n-1)[2bn + (2n-1)\xi]} =: \bar{\rho}_2. \end{aligned}$$

If country n free rides, in the free-trade equilibrium we get

$$e^s = \frac{a[2bn^2 + (2n-1)\xi]}{2b^2n^2 + (2n-1)\xi^2 + b[(2n^2+2n-1)\xi + 2n^2[(4n^2-4n+2)\delta + (2n^2-3n+2)\rho]]} > 0$$

$$\begin{aligned} e_{mn}^d = e_{nn}^d > 0 \quad &\Longleftrightarrow \quad \rho < \frac{(\xi+b)[2bn^2 + (2n-1)\xi - 4(n-1)n^2\delta]}{n^2[(n-2)2b + (2n-3)\xi]} =: \bar{\rho}_3, \\ e_n^d > 0 \quad &\Longleftrightarrow \quad \rho > -\frac{(\xi+b)[2bn^2 + (2n-1)[\xi + 4(n-1)n^2\delta]]}{n^2(2n-1)[2b(n-2) + (2n-3)\xi]} =: \bar{\rho}_4. \end{aligned}$$

In addition, from $\delta_n = \delta + \rho > 0$ follows $\rho > -\delta$. Observe that $\bar{\rho}_1 > 0$, $\bar{\rho}_2 < 0$ and $\bar{\rho}_4 < 0$. Moreover, $b + \frac{2n-1}{2n^2}\xi > 2(n-1)\delta$ is sufficient for $\bar{\rho}_3 > 0$. For sufficiently large parameters $\bar{x}$ the domain of the functions $\mathcal{W}^{*h}(0,\rho)$ and $\bar{\mathcal{W}}^f(0,\rho)$ is

$$\mathcal{S}^\rho := \left\{ (a,b,\xi,n,\delta,c,\rho) | (a,b,\xi,n,\delta,\rho) \in \mathbb{R}^6_{++}, c = 0, n \geq 2, b + \frac{2n-1}{2n^2}\xi > 2(n-1)\delta \wedge \rho \in [\rho_{\min}, \rho_{\max}] \right\}, \tag{B2}$$

where $\rho_{\min} := \max\{-\delta, \bar{\rho}_2, \bar{\rho}_4\}$ and $\rho_{\max} := \min\{\bar{\rho}_1, \bar{\rho}_3\}$.

C: Some analytical results

The welfare functions $\mathcal{W}^{*h}(c,0)$

The welfare functions $\mathcal{W}^{*h}(c,0)$ for $h=m,n$ are specified as[32]

$$\mathcal{W}^{*h}(c,0)=\omega_{ho}^{*}+\omega_{h1}^{*}c+\omega_{h2}^{*}c^{2}, \tag{C1}$$

where

$$\begin{aligned}
&\omega_{mo}^{*}=\omega_{no}^{*}:=\frac{4a^{2}bn^{2}}{z^{*}}>0,\quad \omega_{m1}^{*}:=-\frac{16an^{4}\delta}{z^{*}}<0,\quad \omega_{m2}^{*}:=\frac{n^{2}(\xi-4n^{2}\delta)\delta}{z^{*}},\\
&\omega_{n1}^{*}:=\frac{8an^{2}(b+2n^{2}\delta)}{z^{*}}>0,\quad z^{*}:=8bn^{2}(\xi+b+4n^{2}\delta)>0,\\
&\omega_{n2}^{*}:=\frac{4bn^{2}+n[7n^{3}\delta+n(\xi+5n^{2}\delta)]}{z^{*}}>0.
\end{aligned} \tag{C2}$$

Lemma 1.

(i) The $\mathcal{W}^{*n}(c,0)$-curve is u-shaped and attains its minimum at $\psi_{n}^{*}<0$.

(ii) If $\xi>4n^{2}\delta$, then the $\mathcal{W}^{*m}(c,0)$-curve is u-shaped and attains its minimum at $\psi_{m}^{*}>0$.

(iii) If $\xi<4n^{2}\delta$, then the $\mathcal{W}^{*m}(c,0)$-curve is inverse u-shaped and attains its maximum at $\psi_{m}^{*}<0$.

Proof: Differentiation of $\mathcal{W}^{*h}(c,0)$ with respect to c yields $\mathcal{W}_{c}^{*h}(c,0)=\omega_{h1}^{*}+2\omega_{h2}^{*}c$ and $\mathcal{W}_{cc}^{*h}(c,0)=2\omega_{h2}^{*}$ for $h=m,n$. Solving $\mathcal{W}_{c}^{*h}(\psi_{h}^{*},0)=0$ we obtain $\psi_{h}^{*}=-\frac{\omega_{h1}^{*}}{\omega_{h2}^{*}}$ for $h=m,n$.
Ad (i): Follows from $\omega_{n1}^{*}>0$, $\omega_{n2}^{*}>0$ which in turn implies $\mathcal{W}_{cc}^{*n}>0$ and $\psi_{n}^{*}<0$.
Ad (ii): Follows from $\omega_{m1}^{*}<0$, $\omega_{m2}^{*}>0$ which in turn implies $\mathcal{W}_{cc}^{*m}>0$ and $\psi_{m}^{*}>0$.
Ad (iii): Follows from $\omega_{m1}^{*}<0$, $\omega_{m2}^{*}<0$ which in turn implies $\mathcal{W}_{cc}^{*m}<0$ and $\psi_{m}^{*}<0$.

[32] The procedure for deriving (C1) is described in footnote 18.

The welfare functions $\bar{\mathcal{W}}^f(c,0)$

Next, inserting $\bar{t}_z$ and $\bar{t}_f$ from (A3) and (A6) in $p=\bar{P}(t_z,t_f,c)$, and then inserting $\bar{P}(t_z,t_f,c)$ in (15)-(16) and finally plugging the resultant terms in (16) yields the f-country's welfare function

$$\bar{\mathcal{W}}^f(c,0)=\bar{w}_{fo}+\bar{w}_{f1}+\bar{w}_{f2}c^2, \tag{C3}$$

where

$$\begin{aligned}
\bar{\omega}_{mo} &= \bar{\omega}_{no}=\frac{a^2[4b^3n^4+(\xi-2n\xi)^2(\xi-4n^2\delta)+4b^2n^2[(n^2+2n-1)\xi+4(n-1)^2n^2\delta]]}{\bar{z}} \\
&\quad +\frac{a^2[b(2n-1)[(4n^2+2n-1)\xi^2-8(2n-1)n^2\xi\delta+16n^4(2n^3-3n^2+2n-1)\delta^2]]}{\bar{z}}, && \text{(C4)}\\
\bar{\omega}_{m1} &= -\frac{an^3(2bn+\xi)\delta[4b^2n^2+(4n^2-1)\xi^2+b[(8n^2-1)\xi+4n^2\delta]]}{(b+\xi)\bar{z}}>0, && \text{(C5)}\\
\bar{\omega}_{m2} &= -\frac{(2bn+\xi)^2[-(4n^2-1)\xi^4-2b\xi^2[(6n^2-1)\xi+4n^2\delta-8n^4\delta]+4b^3(4n^4\delta-n^2\xi)+b^2[-(12n^2-1)\xi^2+8n^2(4n^2-1)\xi\delta+16n^4\delta^2]]}{32b(b+\xi)\bar{z}}, && \text{(C6)}\\
\bar{\omega}_{n1} &= \frac{a[4b^4n^4+(2n-1)\xi^3[(2n-1)\xi+n^2(2n^2-7n+4)\delta]+4b^3n^2[(2n^2+2n-1)\xi+2n^2(3n^3-4n+2)\delta]]}{(b+\xi)\bar{z}} \\
&\quad +\frac{ab^2[(4n^4+16n^3-4n^2-4n+1)\xi^2+2n^2(16n^4-14n^3-9n^2+16n-4)\xi\delta+8n^4(8n^4-16n^3+15n^2-8n+2)\delta]}{(b+\xi)\bar{z}} \\
&\quad +\frac{ab\xi[(8n^3+4n^2-8n+2)\xi^2+n^2(8n^4+8n^3-50n^2+47n-12)\xi\delta+4n^4(16n^4-32n^3+28n^2-15n+4)\delta^2]}{(b+\xi)\bar{z}}, && \text{(C7)}\\
\bar{\omega}_{n2} &= \frac{64b^6n^4+(4n^2-1)\xi^6+16b^5n^2[(13n^2+8n-4)\xi+4n^2(7n^2-8n+4)\delta]+}{32b(b+\xi)^2\bar{z}} \\
&\quad +\frac{2b\xi^4[(8n^3+38n^2-34n+7)\xi+4n^2(16n^3-34n^2+28n-7)\delta]}{32b(b+\xi)^2\bar{z}} \\
&\quad +\frac{4b^4[(60n^4+100n^3-33n^2-16n+4)\xi^2+8n^2(36n^4-30n^3-n^2+16n-4)\xi\delta+16n^4(16n^4-32n^3+31n^2-16n+4)\delta^2]}{32b(b+\xi)^2\bar{z}} \\
&\quad +\frac{4b^3\xi[(28n^4+108n^3-n^2-49n+12)\xi^2+4n^2(60n^4-16n^3-71n^2-80n-20)\xi\delta+16n^4(32n^4-64n^3+60n^2-31n+8)\delta^2]}{32b(b+\xi)^2\bar{z}} \\
&\quad +\frac{b^2\xi^2[(16n^4+176n^3+136n^2-200n+47)\xi^2+8n^2(32n^4+40n^3-140n^2+124n-31)\xi\delta]}{32b(b+\xi)^2\bar{z}},
\end{aligned}$$

$$+\frac{b^2\xi^2[16n^4(64n^4-128n^3+112n^2-56n+15)\delta^2]}{32b(b+\xi)^2\bar{z}}, \tag{C8}$$

$$\bar{z} = [2b^2n^2+(2n-1)\xi^2]+b[(2n^2+2n-1)\xi+4n^2(2n^2-2n+1)\delta]^2>0. \tag{C9}$$

Lemma 2.

(i) If ξ is sufficiently large, then

(a) the $\bar{\mathcal{W}}^m(c,0)$-curve is u-shaped and attains its minimum at $\bar{\psi}_m>0$,

(b) the $\bar{\mathcal{W}}^n(c,0)$-curve is u-shaped and attains its minimum at $\bar{\psi}_n<0$.

(ii) If ξ is sufficiently small, then

(a) the $\bar{\mathcal{W}}^m(c,0)$-curve is inverse u-shaped and attains its minimum at $\bar{\psi}_m<0$,

(b) the $\bar{\mathcal{W}}^n(c,0)$-curve is u-shaped and attains its minimum at $\bar{\psi}_n<0$.

Proof: Differentiation of $\bar{\mathcal{W}}^f(c,0)$ with respect to c yields $\bar{\mathcal{W}}^f_c(c,0)=\bar{\omega}_{f1}+2\bar{\omega}_{f2}c$ and $\bar{\mathcal{W}}^f_{cc}(c,0)=2\bar{\omega}_{f2}$ for $f=m,n$. Solving $\bar{\mathcal{W}}^f_c\left(\bar{\psi}_f,0\right)=0$ we obtain $\bar{\psi}_f=-\frac{\bar{\omega}_{f1}}{\bar{\omega}_{f2}}$ for $f=m,n$.
Ad (i): Follows from $\lim_{\xi\to\infty}\bar{\omega}_{m2}=\frac{2n+1}{32b(n-1)}>0$,
$\lim_{\xi\to\infty}\bar{\omega}_{n2}=\frac{2n+1}{32b(2n-1)}>0$, $\lim_{\xi\to\infty}\bar{\omega}_{m1}=0^-$ and[33] $\lim_{\xi\to\infty}\bar{\omega}_{m1}=0^+$.
Ad (ii): Follows from $\lim_{\xi\to 0}\bar{\omega}_{m2}=-\frac{n^2\delta(b+\delta)}{2b[b+2(2n^2-2n+1)\delta]^2}<0$,
$\lim_{\xi\to 0}\bar{\omega}_{n2}=\frac{b^2+b(7n^2-8n+4)\delta+(16n^4-32n^3+31n^2-16n+4)\delta^2}{2b[b+2(2n^2-2n+1)\delta]^2}>0$, $\lim_{\xi\to 0}\bar{\omega}_{m1}=-\frac{2an^2\delta(b+\delta)}{b[b+2(n^2-2n+1)\delta]^2}<0$ and $\lim_{\xi\to 0}\bar{\omega}_{n1}=\frac{a[b^2+2b(3n^2-4n+2)\delta+2(8n^4-16n^3+15n^2-8n+2)\delta^2]}{b[b+2(2n^2-2n+1)\delta]^2}>0$.

Proposition 2. *Suppose the groups M and N are of equal size ($m=n$) and suppose that $\rho=0$.*

(i) Suppose ξ is sufficiently large.

(ia) Then there exists an interval $c\in[\bar{c}_{n2},\bar{c}_{m1}]$ satisfying $\bar{c}_{n2}<0<\bar{c}_{m1}$ and $[\bar{c}_{n2},\bar{c}_{m1}]\subset\,]c_{\min},c_{\max}[$ such that the grand coalition is stable, if $c\in[\bar{c}_{n2},\bar{c}_{m1}]$ and $b<64n^3\delta/(6n-5)$.

(ib) Then there exists an interval $c\in[\bar{c}_{n1},\bar{c}_{m2}]$ satisfying $\bar{c}_{n1}<\bar{c}_{n2}<0<\bar{c}_{m1}<\bar{c}_{m2}$ and $[\bar{c}_{n1},\bar{c}_{m2}]\subset]c_{\min},c_{\max}[$ such that the grand coalition is stable, if $c\in[\bar{c}_{n2},\bar{c}_{m1}]$, $c\in[c_{\min},\bar{c}_{n1}]$ or $c\in[\bar{c}_{m2},c_{\max}]$ and $b>64n^3\delta/(6n-5)$.

(ii) If ξ is sufficiently small, the grand coalition is unstable for every $c\in[c_{\min},c_{\max}]$.

Proof: The properties of the functions $\mathcal{W}^{*h}(c,0)$ and $\bar{\mathcal{W}}^f(c,0)$ are given in the Lemmas 1 and 2. Next, we determine the intersection points of these functions which follow from

$$\bar{\mathcal{W}}^h(c,0)=\mathcal{W}^{*h}(c,0) \quad\Longleftrightarrow\quad (\omega^*_{h2}-\bar{\omega}_{h2})c^2+(\omega^*_{h1}-\bar{\omega}_{h1})c+(\omega^*_{ho}-\bar{\omega}_{ho})=0. \tag{C10}$$

[33] We write $\lim_{\xi\to\infty}\bar{\omega}_m=0^-[0^+]$ if $\bar{\omega}_m$ approaches 0 from negative [positive] values.

Presupposing $(\omega_{h2}^* - \bar{\omega}_{h2}) \neq 0$, the solution to (C10) is

$$\bar{c}_{h1,2} = -\frac{\omega_{h1}^* - \bar{\omega}_{h1}}{2(\omega_{h2}^* - \bar{\omega}_{h2})} \mp \sqrt{\frac{\left(\omega_{h1}^* - \bar{\omega}_{h1}\right)^2}{4(\omega_{h2}^* - \bar{\omega}_{h2})^2} - \frac{\omega_{ho}^* - \bar{\omega}_{ho}}{2(\omega_{h2}^* - \bar{\omega}_{h2})}} \tag{C11}$$

for $h = m, n$. Since $\bar{c}_{m1}$, $\bar{c}_{m2}$, $\bar{c}_{n1}$, $\bar{c}_{n2}$ from (C11) and $c_{\min}$, $c_{\max}$ are functions of the parameter ξ, we write $\bar{c}_{h\ell}(\xi)$ for $h = m, n$, $\ell = 1, 2$ and $c_{\min}(\xi)$ and $c_{\max}(\xi)$.

Ad (i): If ξ is sufficiently large, we get

$$\lim_{\xi\to\infty} \bar{c}_{m\ell}(\xi) = 0^+ \quad \text{and} \quad \lim_{\xi\to\infty} \bar{c}_{n\ell}(\xi) = 0^- \qquad \text{for } \ell = 1, 2. \tag{C12}$$

In addition, the inequalities

$$\lim_{\xi\to\infty} c_{\min}(\xi) < \lim_{\xi\to\infty} \bar{c}_{n2}(\xi) \qquad \text{and} \qquad \lim_{\xi\to\infty} \bar{c}_{n1}(\xi) < \lim_{\xi\to\infty} \bar{c}_{n2}(\xi), \tag{C13}$$

$$\lim_{\xi\to\infty} c_{\max}(\xi) > \lim_{\xi\to\infty} \bar{c}_{m1}(\xi) \qquad \text{and} \qquad \lim_{\xi\to\infty} \bar{c}_{m2}(\xi) > \lim_{\xi\to\infty} \bar{c}_{m1}(\xi), \tag{C14}$$

and the equivalence

$$\left[\lim_{\xi\to\infty} c_{\max}(\xi) \gtreqless \lim_{\xi\to\infty} \bar{c}_{m2}(\xi) \quad \wedge \quad \lim_{\xi\to\infty} c_{\min} \lesseqgtr \lim_{\xi\to\infty} \bar{c}_{n1}\right] \quad \Longleftrightarrow \quad b \gtreqless \frac{64n^3\delta}{6n-5} \tag{C15}$$

hold. (C13) - (C15) prove that the intersection points $\bar{c}_{m1}$, $\bar{c}_{m2}$, $\bar{c}_{n1}$, $\bar{c}_{n2}$ belong to the set of feasible parameters if $b > \frac{64n^3\delta}{6n-5}$. Making use of Lemma 1(i), (ii), 2(i) and (C12)-(C14), we obtain Figure 13 for sufficiently large values of ξ and $b > \frac{64n^3\delta}{6n-5}$. In view of Figure 13, the grand coalition is stable if $c \in [\bar{c}_{n2}, \bar{c}_{m1}]$, $c \in [c_{\min}, \bar{c}_{n1}]$ or if $c \in [\bar{c}_{m2}, c_{\max}]$.

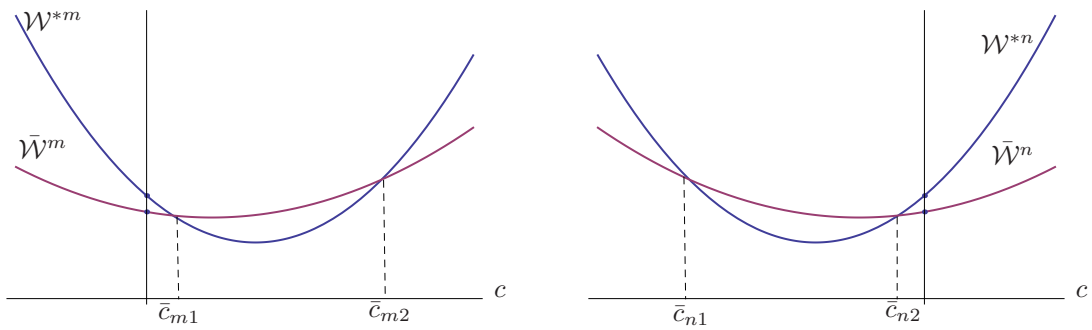

Figure 13: The decision of the f-country, $f = m, n$, to free ride or join the coalition, if ξ is sufficiently large

If $b < \frac{64n^3\delta}{6n-5}$, $\bar{c}_{n1}$ and $\bar{c}_{m2}$ do not belong to the set of feasible parameters. In that case we infer from a figure which equals Figure 13 in qualitative terms that the grand coalition is stable if $c \in [\bar{c}_{n2}, \bar{c}_{m1}]$.

Ad (ii): If ξ is sufficiently small, we find that

$$\bar{c}_{h\ell}(0) = -a\left(1 + \frac{m}{n}\right) < -a \leq c_{\min}(0) \qquad \text{for } h = m, n,\ \ell = 1, 2, \tag{C16}$$

and the curves are continuously at $\xi = 0$. This proves that the intersection points $\bar{c}_{m1}$, $\bar{c}_{m2}$, $\bar{c}_{n1}$, $\bar{c}_{n2}$ do not belong to the set of feasible parameters if ξ is sufficiently low. The information of Lemma 1(i), (ii), 2 (ii) is plotted in Figure 14. From Figure 14 we infer $\bar{\mathcal{W}}^h > \mathcal{W}^{*h}$ for $h = m, n$ for all $c \in [c_{\min}, c_{\max}]$ which establishes that the grand coalition is unstable.

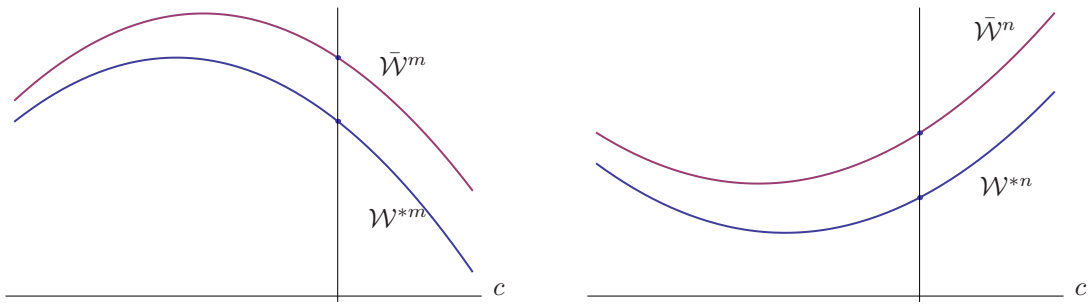

Figure 14: The decision of the f-country, $f = m, n$, to free ride or join the coalition, if ξ is sufficiently small