

Mittelstaedt, Christian; Baumgärtner, Stefan

**Conference Paper**

## Attribution of Collective Causal Responsibility to Individual Actors in a Stochastic System

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Mittelstaedt, Christian; Baumgärtner, Stefan (2022) : Attribution of Collective Causal Responsibility to Individual Actors in a Stochastic System, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/264051>

**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.*

# Attribution of collective causal responsibility to individual actors in a stochastic system

Christian Mittelstaedt\* and Stefan Baumgärtner†

February 24, 2022

## Abstract

We propose an attribution of collective causal responsibility in a stochastic nonlinear system to individual actors. To this end, we take an existing measure of collective causal responsibility from Vallentyne (2008) and Baumgärtner (2020), and adopt Shapley's (1953) fundamental concept of how to divide a collective effect into individual marginal contributions. Our generic setting is a system with a potential regime shift where actions affect the regime-shift probability – e.g. financial markets, managed ecosystems, or the Earth's climate system. This paper closes a gap in the literature on responsibility attribution: our concept has cardinal properties and allows for multiple actors with simultaneous actions. This is relevant for implementing efficient incentive schemes and liability for managers of stochastic systems.

*Keywords:* attribution, causal responsibility, probability, multiple actors, regime shift, Shapley value

*JEL codes:* K13, K32, Q50

---

\*Chair of Environmental Economics and Resource Management, University of Freiburg, Tennenbacher Str. 4, D-79106 Freiburg, Germany, phone: +49 761 203-3784, email: [christian.mittelstaedt@ere.uni-freiburg.de](mailto:christian.mittelstaedt@ere.uni-freiburg.de).

†Chair of Environmental Economics and Resource Management, University of Freiburg.

<sup>0</sup>Acknowledgements: We are grateful to Martin Quaas, Marie-Catherine Riekhof, Michael Stecher for comments, and to the German Federal Ministry of Education and Research (BMBF) for financial support under grant no. 01LA1814A.

# 1 Introduction

Attribution of causal responsibility for an outcome to individual actors poses a challenge in stochastic nonlinear systems with multiple actors who act simultaneously: any given outcome could be due to (individual or collective) actions, or brute luck, or a combination of all. Hence, there is the question: “To what extent is an outcome caused by an individual actor’s action, rather than by some other actor’s action or by brute luck?” This is important for implementing efficient incentive schemes, and liability, for managers of stochastic systems.

We propose an attribution of collective causal responsibility in a stochastic nonlinear system to individual actors. To this end, we take an existing measure of collective causal responsibility from Vallentyne (2008) and Baumgärtner (2020), and adopt Shapley’s (1953) fundamental concept of how to divide a collective effect into individual marginal contributions. Our generic setting is a stochastic system with a potential regime shift, where actions affect the regime-shift probability. Examples include managed ecosystems with alternative stable states such as shallow lakes, rangelands or fisheries (Folke et al., 2004), the tipping of the Earth’s climate system due to greenhouse gas emissions (Lenton et al., 2008; Steffen et al., 2018), or regime-switching in financial markets due to regulation or investor behavior (Ang and Timmermann, 2012).

Our paper closes a gap in the literature on responsibility attribution. Braham and Van Hees (2009) consider multiple actors in a simultaneous situation and study degrees of causation, yet not with a cardinal measure. Dehez and Ferey (2013) and Ferey and Dehez (2016) propose a method for sharing the (absolute) value of damage caused jointly by a group of tortfeasors; yet they assume a sequential structure of the damage causation. Our concept has cardinal properties and allows for multiple actors with simultaneous actions.

The paper is organized as follows. In Section 2, we introduce the model. In Section 3, we propose the attribution measure and discuss its properties. Section 5 discusses and concludes.

## 2 Model

Consider a system which can be in either one of two states. A potential regime shift from the initial state to the other state occurs with probability  $p \in [0, 1]$ . Before any actions take effect,

the regime-shift probability is  $p_0$ . This means, a shift could naturally occur due to stochastic influences.

Each actor  $i \in \mathcal{N} = \{1, \dots, n\}$  chooses simultaneously with, and independently of, the other actors her individual action  $a_i \in \mathbb{R}_+$ . The collective action  $a = (a_1, \dots, a_i, \dots, a_n)$  is an element of the action set  $A \subseteq \mathbb{R}_+^n$ .  $\mathcal{P}(\mathcal{N})$  denotes the power set of  $\mathcal{N}$ , i.e. the set of all subsets of all actors, where every element  $C \in \mathcal{P}(\mathcal{N})$  defines a group (*coalition*) of acting actors.  $|C|$  denotes the number of members in coalition  $C$ .

For any given actual action  $a \in A$  we define hypothetical actions as follows. If only members of a particular coalition  $C$  were considered, and all other potential actors were not, the collective action would be the  $n$ -tuple  $a^C$  with components

$$a_j^C = \begin{cases} a_j & \text{for all } j \in C \\ 0 & \text{else} \end{cases} .$$

If the members of the coalition  $C$  and an additional actor  $i$  were considered, and all other potential actors were not, the collective action would be the  $n$ -tuple  $a^{C \cup \{i\}}$  with components

$$a_j^{C \cup \{i\}} = \begin{cases} a_j & \text{for all } j \in C \cup \{i\} \\ 0 & \text{else} \end{cases} .$$

Note that  $a^{\mathcal{N}} = a$ . The collective action influences the regime shift probability, as described by a continuously differentiable function  $p : A \rightarrow [0, 1]$ . In general, the effect of the individual actions  $a_i$  on  $p$  is not linear, symmetric, or additive.<sup>1</sup> We assume the following properties.

**Assumption 1.** For all  $a \in A$ :

$$p(0, \dots, 0, \dots, 0) = p_0 , \tag{1}$$

$$\frac{\partial p}{\partial a_i} \geq 0 \text{ for all } i . \tag{2}$$

(1) states that a collective null-action is probabilistically neutral: it does not change the initial regime-shift probability  $p_0$ . (2) states that the regime-shift probability is increasing in every individual action  $a_i$ . That is, for simplicity we assume that there are no probability-lowering actions. Together, these two assumptions imply that there is no individual or collective

---

<sup>1</sup>There might be amplifying or extenuating effects in the combination of different individual actions.

action which reduces the regime-shift probability  $p$  below  $p_0$ . Furthermore, (2) implies that

$$C_1 \subseteq C_2 \Rightarrow p(a^{C_1}) \leq p(a^{C_2}) \text{ for all } C_1, C_2 \in \mathcal{P}(\mathcal{N}). \quad (3)$$

That is, the regime-shift probability effected by the collective action of a subgroup  $C_1$  of a coalition  $C_2$  cannot be larger than the regime-shift probability effected by the collective action of the coalition  $C_2$ .

So far, we have described one particular system  $s$  which is defined by the set of actors  $\mathcal{N}$ , the action set  $A$  and the probability function  $p^s : A \rightarrow [0, 1]$  which describes the probability of a regime shift.<sup>2</sup> For the same set of actors  $\mathcal{N}$  and their action set  $A$  one may now consider another system  $s'$  which is defined by some other, independent, regime-shift event and the probability  $p^{s'} : A \rightarrow [0, 1]$  by which it is effected by the action  $a$ . The set of all systems is denoted by  $S$ . For simplicity, we assume that every  $s \in S$  is independent such that a change in one system's regime shift probability does not alter another system's probability. For example one may think of a set of fishing companies ( $\mathcal{N}$ ), each of which decides on investments in boats and gear ( $a = (a_1, \dots, a_n)$ ), which affects the probabilities of two potential regime shifts, collapse of the fish stock due to overfishing ( $s$  and  $p^s$ ) and tipping of the climate system due to CO<sub>2</sub>-emissions from operation of the boats ( $s'$  and  $p^{s'}$ ).

### 3 Responsibility attribution

A quantitative measure of causal responsibility answers the question: “If a regime shift actually occurs, to what extent is it caused by the actor's action, rather than by some other actor's action or by brute luck?” This is a real number between 0 and 1, which can be defined both at the collective and the individual level. The coalition can be considered as a single actor, having a collective responsibility of  $R(a)$  (“If a regime shift actually occurs, to what extent is it caused by the coalition's collective action, rather than by brute luck?”). Also, each individual actor  $i$  has an individual responsibility  $R_i(a)$  (“If a regime shift actually occurs, to what extent is it caused by the individual actor's action, rather than by some other actor's action or by brute luck?”).

---

<sup>2</sup>In order to distinguish between different systems  $s$  we added a superscript  $s$  to the probability function. When only one system is considered and no confusion can arise, we drop the superscript  $s$ .

For measuring the coalition’s collective causal responsibility for a regime-shift we take an existing measure from Baumgärtner (2020) who studies a setting where one single actor takes action  $a$  and formalizes an idea of Vallentyne (2008):

$$R(a) = \frac{p(a) - p_0}{p(a)}. \quad (4)$$

This is the direct change of probability  $p(a) - p_0$  due to the action, which is weighted by  $1/p(a)$  due to the partial attribution of outcome realization luck to the action.<sup>3</sup> In the setting studied here, (4) corresponds to the coalition’s collective responsibility for a regime shift.

We now attribute the collective responsibility  $R(a)$  to individual actors  $i$  and denote this individual responsibility  $R_i(a)$ . We propose that the measure should fulfill the following axioms:

**Axiom 1.** *Symmetry*

$$R_{\pi_i}(\pi a) = R_i(a) \text{ for every permutation } \pi \text{ of individuals } (1, \dots, i, \dots, n).$$

Axiom 1 states that the order in which the actors are numbered is irrelevant for the responsibilities they get attributed.

**Axiom 2.** *Full attribution*

$$0 \leq R_i(a) \leq R(a) \text{ and } \sum_{i=1}^n R_i(a) = R(a).$$

Axiom 2 states that no individual responsibility  $R_i(a)$  can be negative or larger than the collective responsibility, and that the coalition’s collective responsibility is fully attributed to the individual actors.<sup>4</sup>

For the next axiom, we consider several systems  $s \in S$ , such that the collective action  $a$  simultaneously and independently influences probabilities  $p^s$  of all potential regime shifts. In this joint-systems setting, we can independently determine individual responsibilities  $R_i^s(a)$  for the regime shift in system  $s$ , a collective responsibility  $R^s(a)$  of all individuals for the regime shift in system  $s$ , an individual responsibility  $R_i(a)$  for the joint event of regime shifts in all systems  $s \in S$ , and the collective responsibility  $R(a)$  for the joint event of regime shifts in all systems  $s \in S$ .

---

<sup>3</sup>In climate change attribution science, (4) is known as the *fraction of attributable risk* (Otto, 2017; Pfrommer et al., 2019).

<sup>4</sup>Note that Axioms 1 and 2 correspond to Shapley’s (1953) Axioms 1 and 2.

**Axiom 3.** *Consistent joint responsibility*

$$R(a) = \sum_{s=1}^{|S|} w_s R^s(a) \Leftrightarrow R_i(a) = \sum_{s=1}^{|S|} w_s R_i^s(a) \text{ for all } w_s \in [-1, 1].$$

This axiom guarantees consistency over several systems for which all or some actors of the collective action have responsibility.<sup>5</sup> It states that aggregation of individual responsibilities over systems is possible by some weights  $w_s$  if and only if aggregation of collective responsibility is done by the very same weights. The weights  $w_s$  can but do not need to take on negative values, and they can but do not need to add up to 1. Hence, these weights are not relative weights like probabilities, but they are absolute weights as they are known e.g. from linear algebra when decomposing a vector into a weighted sum of its elementary vectors.

Axioms 1-3 uniquely characterize a cardinal measure of individual responsibility.

**Proposition 1.** *For all  $i \in \mathcal{N}$  and all  $a \in A$ ,  $R_i(a)$  satisfies Axioms 1-3 if and only if*

$$R_i(a) = \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} \frac{(n-1-|C|)! \cdot |C|!}{|\mathcal{N}|!} \frac{p(a^{C^i}) - p(a^C)}{p(a)}. \quad (5)$$

**Proof:** See Appendix A1.

The logic behind (5), which extends Shapley's (1953) fundamental concept of how to divide a collective effect into individual marginal contributions to the realm of risky outcomes and to partial causal responsibility, is as follows: The second fraction specifies the marginal relative change in regime-shift probability caused by actor  $i$  when joining a coalition  $C$ . This marginal change differs depending on the coalition  $C$  that  $i$  joins. As all actors act simultaneously, we consider all possible coalitions  $C$  of which  $i$  is not a member, and average over them. To ensure overall consistency, each of these possibilities is weighted by the relative frequency with which it can occur when building up each coalition by sequentially adding one actor at a time. This weighting is achieved by the binomial factor in (5). With this construction,  $i$ 's individual responsibility  $R_i(a)$  does not only depend on her own action  $a_i$  but, in general, on *all* actors' actions  $a$ .<sup>6</sup>

The individual responsibility measure  $R_i$  has the following additional properties.

---

<sup>5</sup>Our Axiom 3 corresponds to Shapley's (1953) Axiom 3. It takes into account that we are working with relative measures instead of absolute payoff or utility amounts.

<sup>6</sup>Technically, this is due to  $p(a)$  not being additive.  $p(a)$  is *additive* if and only if  $p(a^{C_1}) + p(a^{C_2}) = p(a^{C_1 \cup C_2})$  for all  $a \in A$  and all  $C_1, C_2 \in \mathcal{P}(\mathcal{N})$ .

**Proposition 2.** For all  $i, j \in \mathcal{N}$  and all  $a \in A$ :

1.  $R_i(a) \leq R_i(\tilde{a})$  if and only if

$$\frac{p(\tilde{a})}{p(a)} \leq \frac{\sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} (n-1-|C|)! \cdot |C|! \cdot (p(\tilde{a}^{C_i}) - p(\tilde{a}^C))}{\sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} (n-1-|C|)! \cdot |C|! \cdot (p(a^{C_i}) - p(a^C))} \quad (6)$$

with  $\tilde{a} = (a_1, \dots, a_{i-1}, \tilde{a}_i, a_{i+1}, \dots, a_n)$  for any  $\tilde{a}_i \leq a_i$ .

2.  $R_i(a) \geq R_j(a)$  if and only if  $\sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i, j\})} (n-2-|C|)! \cdot |C|! \cdot p(a^{C_i}) \geq \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i, j\})} (n-2-|C|)! \cdot |C|! \cdot p(a^{C_j})$ . In particular, for all  $C$  that contain neither  $i$  nor  $j$ , if  $p(a^{C_i}) = p(a^{C_j})$  then  $R_i = R_j$ .

3. For any actor  $j$  with  $a_j$  such that  $p(a^{C_j}) - p(a^C) = 0$  for all  $C \in \mathcal{P}(\mathcal{N})$ , it holds that  $R_i(a) = R_i(a_1, \dots, a_{j-1}, 0, a_{j+1}, \dots, a_n)$  for all  $i$ ; in particular  $R_j(a) = 0$ .

4.  $R_i(a) \leq R_i(\tilde{a})$  for a particular actor  $i$  and two collective actions  $a, \tilde{a}$  with  $a = (a_1, \dots, a_i, \dots, a_n)$ ,  $\tilde{a} = (0, \dots, \tilde{a}_i, \dots, 0)$  and  $p(a) = p(\tilde{a})$ .

**Proof:** See Appendix A2.

Property 1 states how individual responsibility changes with the actor's action. Actor  $i$ 's responsibility increases if and only if the change from action  $a_i$  to  $\tilde{a}_i$  is such that, ceteris paribus, the ratio of the weighted average over all coalitions of marginal increase in regime-shift probability due to  $\tilde{a}_i$  (RHS nominator of (6)) to that of  $a_i$  (RHS denominator of (6)) is greater than the ratio of the regime-shift probability with collective action  $\tilde{a}$  to the probability with collective action  $a$ . Roughly speaking, this means that the change in actor  $i$ 's responsibility does not only depend on the change in regime-shift probability due to the change in actor  $i$ 's action, but also on the change in the weighted average over all coalitions of marginal increase in regime-shift probability. Note that  $\partial p(a)/\partial a_i \geq 0$  is neither necessary nor sufficient for  $\partial R_i(a)/\partial a_i \geq 0$ , because the other actors' actions matter, too, for attributing individual responsibility according to (5).<sup>7</sup>

Property 2 compares two actors and their individual responsibilities. Actor  $i$  has a larger responsibility than  $j$  if and only if  $i$ 's action has a larger influence than  $j$ 's in the sense that  $i$ 's

<sup>7</sup>Of course,  $\partial p(a)/\partial a_i \geq 0$  implies  $\partial R(a)/\partial a_i \geq 0$ .



marginal contribution to regime-shift probability is larger than that of  $j$  on (weighted) average over all coalitions that do neither contain  $i$  nor  $j$ .<sup>8</sup> In particular, the attributed individual responsibility of two actors is equal if their marginal contribution to regime-shift probability when joining any given coalition is equal.

Property 3 states that dummy actions are irrelevant for attributing responsibility to any actor, where a dummy action is one which does not make a marginal contribution to regime-shift probability in any coalition. As a consequence, for attributing responsibility to those actors whose actions have a probability impact, it does not matter how many dummy actors are present. Property 4 states that an actor  $i$  has lower responsibility when acting in a group compared to committing a probability-equivalent action all by herself.

## 4 Illustrations and applications

### 4.1 A general numerical illustration

Consider a system with three actors where different coalitions have the following impact on the regime-shift probability:  $p(a^\emptyset) = p_0 = 0.1$ ,  $p(a^{\{1\}}) = 0.7$ ,  $p(a^{\{2\}}) = 0.5$ ,  $p(a^{\{3\}}) = 0.2$ ,  $p(a^{\{1,2\}}) = 1$ ,  $p(a^{\{1,3\}}) = 0.9$ ,  $p(a^{\{2,3\}}) = 0.85$ ,  $p(a^{\{1,2,3\}}) = 1$ . There is a small probability that the system shifts on its own due to stochastic processes. The combined actions of 1 and 2 is already sufficient to cause a regime shift with certainty. Furthermore, the combined actions of 2 and 3 have a more-than-proportionally effect on the regime-shift probability.

Suppose all three actors act and a regime shift actually occurs. Table 1 shows for each actor  $i = 1, 2, 3$  the corresponding marginal probability changes for the different sequences of building all possible coalitions, and the individual causal responsibilities  $R_i(a)$  according to (5).

---

<sup>8</sup>This does not imply that simply  $p(a^i) > p(a^j)$ . The difference in influence can result from the higher probability of a coalition that  $i$  is a member of while  $j$  is not.

sequence	$i = 1$	$i = 2$	$i = 3$
1-2-3	$p(a^{\{1\}}) - p(a^\emptyset) = .6$	$p(a^{\{1,2\}}) - p(a^{\{1\}}) = .3$	$p(a^{\{1,2,3\}}) - p(a^{\{1,2\}}) = 0$
1-3-2	$p(a^{\{1\}}) - p(a^\emptyset) = .6$	$p(a^{\{1,2,3\}}) - p(a^{\{1,3\}}) = .1$	$p(a^{\{1,3\}}) - p(a^{\{1\}}) = .2$
2-1-3	$p(a^{\{1,2\}}) - p(a^{\{2\}}) = .5$	$p(a^{\{2\}}) - p(a^\emptyset) = .4$	$p(a^{\{1,2,3\}}) - p(a^{\{1,2\}}) = 0$
2-3-1	$p(a^{\{1,2,3\}}) - p(a^{\{2,3\}}) = .15$	$p(a^{\{2\}}) - p(a^\emptyset) = .4$	$p(a^{\{2,3\}}) - p(a^{\{2\}}) = .35$
3-1-2	$p(a^{\{1,3\}}) - p(a^{\{3\}}) = .7$	$p(a^{\{1,2,3\}}) - p(a^{\{1,3\}}) = .1$	$p(a^{\{3\}}) - p(a^\emptyset) = .1$
3-2-1	$p(a^{\{1,2,3\}}) - p(a^{\{2,3\}}) = .15$	$p(a^{\{2,3\}}) - p(a^{\{3\}}) = .65$	$p(a^{\{3\}}) - p(a^\emptyset) = .1$
$R_i(a)$	.45	.325	.125

Table 1: Marginal probability changes for the different sequences of building all possible coalitions, and individual causal responsibilities  $R_i(a)$ , for each actor  $i$

Note that the sum of individual probabilities, which is the actors' collective responsibility  $R(a_1, a_2, a_3)$ , is 0.9. This is in line with (4) which does not attribute  $p_0$  to the actors.

## 4.2 Linear additive probabilities

The, in general, complicated formula (5) for individual responsibility reduces to a simple and plausible one if the probability function  $p$  has a simple structure.

**Proposition 3.** *For all  $i \in \mathcal{N}$  and  $a \in A$  it holds that*

$$R_i(a) = \frac{a_i}{\sum_{i \in \mathcal{N}} a_i} R(a) \quad (7)$$

*if and only if*

$$p(a^C) = \beta \cdot \sum_{i \in C} a_i + p_0 \quad \text{with } \beta \in \mathbb{R} \text{ for all } C \in \mathcal{P}(\mathcal{N}). \quad (8)$$

**Proof:** See Appendix A3.

The responsibility attribution scheme (7) states that individual responsibility is given by a proportional share of collective responsibility, where the proportionality is in terms of the individual's share of action relative to the aggregate action. For example, responsibility for climate change is often attributed to individual countries according to the country's relative share of greenhouse gas emissions, where greenhouse gas emissions increase the probability

of climate change. Our Proposition 3 shows that such a simplified attribution scheme is only warranted if the actions impact on the probabilities as specified by (8), that is, if the probability function is linear and additive in the actions. In the case of greenhouse gas emissions causing climate change, however, this condition is not fulfilled.

The simplified responsibility attribution scheme (7), being a special case of the general scheme (5), has all the properties stated in Proposition 2. Some of these general properties are less ambiguous in this special case. For instance, 1. and 2. of Proposition 2 simplify as follows: For all  $i, j \in \mathcal{N}$  and all  $a \in A$ :

1.  $R_i(a) \leq R_i(\tilde{a})$  if and only if  $a_i \leq \tilde{a}_i$ .
2.  $R_i(a) \geq R_j(a)$  if and only if  $a_i \geq a_j$ .

## 5 Discussion and conclusion

Equation (5) provides an attribution of collective causal responsibility in a stochastic nonlinear system to individual actors who act simultaneously, adopting Shapley's (1953) fundamental concept of how to divide a collective effect into individual marginal contributions.

By the Shapley mechanism the individual responsibility depends not only on the actor's own action, but on the action of others, too. This can give rise to results which one may think of as counterintuitive and which are discussed, e.g., in Braham and Van Hees (2009).

One restriction of our framework is the assumption that any action, *ceteris paribus*, increases the regime-shift probability. This was to avoid that some individual responsibilities could become smaller than zero and, consequently, other individual responsibilities greater than one. One possible extension of our framework is to probability-lowering actions.

We suggest to apply our measure (5) to assign strict liability for a damage to an actor in proportion to the actors' causal responsibility for the damage. Another application would be to calculate efficient user fees for exploiting a common-pool resource in proportion to the users' impact on the state of the resource. For such applications, and their efficiencies, both the regulator as well as the actors need to know the function  $p(a)$ . While this seems demanding, it is needed to attribute individual causal responsibility in a stochastic system.

## References

- Ang, A. and Timmermann, A. (2012). Regime changes and financial markets. *Annual Review of Financial Economics*, 4(1):313–337.
- Baumgärtner, S. (2020). Responsibility for regime shifts in managed ecosystems. *SSRN Discussion paper 3752006*. Available at: <https://ssrn.com/abstract=3752006>.
- Braham, M. and Van Hees, M. (2009). Degrees of causation. *Erkenntnis*, 71(3):323–344.
- Dehez, P. and Ferey, S. (2013). How to share joint liability: A cooperative game approach. *Mathematical Social Sciences*, 66(1):44–50.
- Ferey, S. and Dehez, P. (2016). Multiple causation, apportionment, and the shapley value. *The Journal of Legal Studies*, 45(1):143–171.
- Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L., and Holling, C. S. (2004). Regime shifts, resilience, and biodiversity in ecosystem management. *Annual Review of Ecology, Evolution, and Systematics*, 35:557–581.
- Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J. (2008). Tipping elements in the earth’s climate system. *Proceedings of the national Academy of Sciences*, 105(6):1786–1793.
- Otto, F. E. L. (2017). Attribution of weather and climate events. *Annual Review of Environment and Resources*, 42:627–646.
- Pfrommer, T., Goeschl, T., Proelss, A., Carrier, M., Lenhard, J., Martin, H., Niemeier, U., and Schmidt, H. (2019). Establishing causation in climate litigation: admissibility and reliability. 152(1):67–84.
- Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28):307–317.
- Steffen, W., Rockström, J., Richardson, K., Lenton, T. M., Folke, C., Liverman, D., Summerhayes, C. P., Barnosky, A. D., Cornell, S. E., Crucifix, M., et al. (2018). Trajectories of the earth system in the anthropocene. *Proceedings of the National Academy of Sciences*, 115(33):8252–8259.

Vallentyne, P. (2008). Brute luck and responsibility. *Politics, Philosophy & Economics*, 7(1):57–80.

Young, H. P. (1988). *Individual contribution and just compensation*, pages 267–278. Cambridge University Press.

# A1 Proof of Proposition 1

1.  $R_i(a)$  satisfies Axioms 1-3.

Axiom 1 is obvious.

Axiom 2.

$0 \leq R_i(a) \leq R(a)$ :

(i)  $0 \leq R_i(a)$ , as  $\frac{(n-1-|C|)! \cdot |C|!}{|\mathcal{N}|!} > 0$  and  $\frac{p(a^{C_i}) - p(a^C)}{p(a)} \geq 0$  for all  $C \in \mathcal{P}(\mathcal{N})$ .

(ii) Note that  $\frac{p(a^{C_i}) - p(a^C)}{p(a)} \leq \frac{p(a) - p_0}{p(a)} \leq 1$  for all  $a \in A$ ,  $C \in \mathcal{P}(\mathcal{N})$ . Further,

$$\sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} \frac{(n-1-|C|)! \cdot |C|!}{|\mathcal{N}|!} = \sum_{|C|=0}^{n-1} \frac{(n-1)!}{|C|!(n-1-|C|)!} \frac{(n-1-|C|)! \cdot |C|!}{n!} \quad (\text{A.1})$$

$$= n \frac{(n-1)!}{n!} = 1. \quad (\text{A.2})$$

Hence,

$$\sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} \frac{(n-1-|C|)! \cdot |C|!}{|\mathcal{N}|!} \frac{p(a^{C_i}) - p(a^C)}{p(a)} \leq \frac{p(a) - p_0}{p(a)}. \quad (\text{A.3})$$

Therefore,  $0 \leq R_i(a) \leq R(a)$ . □

$\sum_{i=1}^n R_i(a) = R(a)$ :

(i) Consider that actors' actions are added to a coalition  $C$  according to an arbitrary sequence  $v$ , until  $a^C = a^{\mathcal{N}} = a$ . Then

$$\sum_{i=1}^n \frac{p(a^{C_i}) - p(a^C)}{p(a)} = \frac{p(a) - p_0}{p(a)} = R(a). \quad (\text{A.4})$$

(ii) Hence, the equal-probability average over all possible sequences of actors is

$$\frac{1}{n!} \sum_{v(C)} \sum_{i=1}^n \frac{p(a^{C_i}) - p(a^C)}{p(a)} = \frac{1}{n!} \cdot n! \cdot \frac{p(a) - p_0}{p(a)} = R(a) \quad (\text{A.5})$$

(iii) Note that  $\frac{(n-1-|C|)! \cdot |C|!}{|\mathcal{N}|!}$  gives the relative frequency with which the specific marginal probability change  $\frac{p(a^{C_i}) - p(a^C)}{p(a)}$  can occur for a specific actor  $i$  when building up each coali-

tion by sequentially adding one actor at a time. Hence,

$$\frac{1}{n!} \sum_{v(C)} \sum_{i=1}^n \frac{p(a^{C^i}) - p(a^C)}{p(a)} = \sum_{i=1}^n \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} \frac{(n-1-|C|)! \cdot |C|!}{|\mathcal{N}|!} \frac{p(a^{C^i}) - p(a^C)}{p(a)} \quad (\text{A.6})$$

$$= \sum_{i=1}^n R_i(a) = R(a). \quad (\text{A.7})$$

□

**Axiom 3.**

$$(i) R(a) = \sum_{s=1}^S w_s R^s(a) \Rightarrow R_i(a) = \sum_{s=1}^S w_s R_i^s(a):$$

$$R(a) = \sum_{s=1}^S w_s R^s(a) \stackrel{\text{Ax.2}}{=} \sum_{s=1}^S w_s \sum_{i=1}^n R_i^s(a) = \sum_{i=1}^n \sum_{s=1}^S w_s R_i^s(a) \stackrel{\text{Ax.2}}{\Rightarrow} \sum_{s=1}^S w_s R_i^s(a) \stackrel{!}{=} R_i(a). \quad (\text{A.8})$$

$$(ii) R_i(a) = \sum_{s=1}^S w_s R_i^s(a) \Rightarrow R(a) = \sum_{s=1}^S w_s R^s(a):$$

$$R_i(a) = \sum_{s=1}^S w_s R_i^s(a) \Rightarrow \sum_{i=1}^n R_i(a) = \sum_{i=1}^n \sum_{s=1}^S w_s R_i^s(a) \stackrel{\text{Ax.2}}{\Leftrightarrow} R(a) = \sum_{s=1}^S w_s R^s(a). \quad (\text{A.9})$$

□

**2. Axioms 1-3 imply  $R_i(a)$ .**

The remainder of this proof follows the determination of the Shapley Value (Shapley, 1953, section 3) and has the following structure:

**2.1** We define a set of (hypothetical) unanimity systems whose weighted probabilities decompose the probabilities of an actual system  $s \in S$ .

**2.2** The probabilities of the probability functions of the set of unanimity systems are a linear basis of the probabilities of  $s$ . That means, that the probabilities in  $s$ , given by the actions of all possible coalitions, can be represented by a unique set of values  $w_U$ , such that for all  $C \in \mathcal{P}(\mathcal{N})$ ,  $p(a^C)$  can be expressed by linear combination of values  $w_U$  and the probability values of the corresponding unanimity system. We prove this by showing that the vectors of probabilities of the unanimity systems are linearly independent.

**2.3** We show that the individual  $i$ 's responsibility resulting from (5) is a weight sum of  $i$ 's

responsibilities in the set of unanimity systems and therefore unique.

**2.1:** So far, we have shown that  $R_i(a)$  fulfills the stated axioms. To prove that there exists only one responsibility measure that fulfills the axioms for the general case, we consider (hypothetical) systems, that we call "unanimity systems".<sup>9</sup> A unanimity system  $s^U \in S$  is defined as a system for which the chosen actions  $a = (a_1, \dots, a_n)$  entail probabilities  $p^U$  as follows: For a corresponding coalition  $U \in \mathcal{P}(\mathcal{N})$ ,

$$p^U(a^C) = \begin{cases} 1, & \text{if } U \subseteq C \\ 0 & \text{else.} \end{cases} \quad (\text{A.10})$$

We assume that for any coalition  $U \in \mathcal{P}(\mathcal{N})$ , a respective unanimity system  $s^U$  exists, hence there are  $2^n$  unanimity systems (given a finite  $|\mathcal{N}|$ ). We denote the set of unanimity systems as  $(s^U)_{U \subseteq \mathcal{N}}$  and the corresponding set of probability functions  $(p^U)_{U \subseteq \mathcal{N}}$ . Note that for a coalition  $U$ ,  $p^U(a^{U \cup C}) = p^U(a^U) = 1$ , and  $p^U(a^{C \cap U}) = p^U(a^C)$  for any  $C \in \mathcal{P}(\mathcal{N})$ .

Due to Axioms 1 and 2 the attributed responsibility of an actor  $i$  in a unanimity system is

$$R_i^U(a) = \begin{cases} \frac{1}{|U|} & \text{if } i \in U \\ 0 & \text{else.} \end{cases} \quad (\text{A.11})$$

Note that  $R_i^U(a) = 0$  for all  $i$  if  $U = \emptyset$ . Further, unanimity system responsibilities can be weighted, such that

$$\bar{w}_U R_i^U(a) = \begin{cases} \frac{\bar{w}_U}{|U|} & \text{if } i \in U \\ 0 & \text{else.} \end{cases}$$

**2.2:**

**Lemma 1.** *The regime shift probability  $p(a)$  of a system  $s \in S$  can be uniquely derived from a linear combination of weighted probabilities of the set of unanimity systems, such that*

$$p(a) = \sum_{U \subseteq \mathcal{N}} w_U p^U(a). \quad (\text{A.12})$$

---

<sup>9</sup>The term unanimity system stems from "unanimity games" which have a similar structure in payoffs as the unanimity systems have in probability change, see e.g. Young (1988).



The resulting  $w_U$  do not only hold for  $p(a^{\mathcal{N}})$ , but any regime shift probability  $p(a^C)$ ,  $C \in \mathcal{P}(\mathcal{N})$ . Lemma 1 is true if the vectors of probabilities of the unanimity systems  $(p^U)_{U \subseteq \mathcal{N}}$  are a linear basis for the values of  $p(a)$ . As stated above, there exist  $2^n$  distinct unanimity systems and therefore also  $2^n$  distinct probability functions. Each probability function can be characterized by a vector with length  $2^n$ , with components being the values of  $p^U(a^C)$  for any possible coalition  $C \in \mathcal{P}(\mathcal{N})$ . For instance, the vector for  $p^\emptyset$  is a vector with 1 in every entry, and the vector for  $p^{\mathcal{N}}$  is a vector with 0 except for one 1-entry. For illustrative purposes, the case of  $n = 3$  looks as follows:

$p^U \backslash a^C$	$a^\emptyset$	$a^1$	$a^2$	$a^3$	$a^{1,2}$	$a^{1,3}$	$a^{2,3}$	$a^{1,2,3}$
$p^\emptyset(a^C)$	=(1	1	1	1	1	1	1	1)
$p^1(a^C)$	=(0	1	0	0	1	1	0	1)
$p^2(a^C)$	=(0	0	1	0	1	0	1	1)
$p^3(a^C)$	=(0	0	0	1	0	1	1	1)
$p^{1,2}(a^C)$	=(0	0	0	0	1	0	0	1)
$p^{1,3}(a^C)$	=(0	0	0	0	0	1	0	1)
$p^{2,3}(a^C)$	=(0	0	0	0	0	0	1	1)
$p^{1,2,3}(a^C)$	=(0	0	0	0	0	0	0	1)

**Lemma 2.** For any finite  $|\mathcal{N}|$ , the set of vectors for  $(p^U)_{U \subseteq \mathcal{N}}$  is linear independent, i.e.

$$\sum_{U \subseteq \mathcal{N}} \lambda_U p^U(a) = (0, \dots, 0) \Rightarrow \lambda_U = 0 \text{ for all } U \in \mathcal{P}(\mathcal{N}). \quad (\text{A.13})$$

**Proof:** By contradiction, assume

$$\sum_{U \subseteq \mathcal{N}} \lambda_U p^U(a) = (0, \dots, 0), \text{ where there exists at least one } \lambda_U \neq 0. \quad (\text{A.14})$$

In this case, define  $M \in \mathcal{P}(\mathcal{N})$  as minimal size coalition, such that  $\lambda_M \neq 0$ . As  $U$  can either be a subset of  $M$  or not, we rewrite

$$\sum_{U \subseteq \mathcal{N}} \lambda_U p^U(a) = 0 \Leftrightarrow \sum_{U \subseteq M} \lambda_U p^U(a^M) + \sum_{U \not\subseteq M} \lambda_U p^U(a^M) = 0. \quad (\text{A.15})$$

From the definition of unanimity systems, we know that the second part of the addition is equal 0, and  $p^U(a^M) = 1$  in the first part of the addition. Hence,

$$\sum_{U \subseteq M} \lambda_U = 0. \quad (\text{A.16})$$

Since  $M$  is a subset of minimal size for which  $\lambda_M \neq 0$ , we have that  $\lambda_U = 0$  for all real subsets  $U \subset M$ . Thus (A.16) is reduced to  $\lambda_M = 0$ . This is a contradiction! Hence, the set of vectors for  $(p^U)_{U \subseteq \mathcal{N}}$  is linear independent (Lemma 2) and consequently,  $p(a)$  can be uniquely derived from a linear combination (Equation A.12, Lemma 1).  $\square$

Note that from (A.12) results the general expression for  $w_U$ : For any  $U \subseteq \mathcal{N}$ ,

$$w_U = \sum_{V \subseteq U} (-1)^{|U|-|V|} \cdot p(a^V). \quad (\text{A.17})$$

For illustration, let us calculate  $w_{\{1,2,3\}}$ . From (A.12), we know that

$$\begin{aligned} p(a^\emptyset) &= w_\emptyset, \quad p(a^{\{1\}}) = w_\emptyset + w_{\{1\}}, \quad p(a^{\{2\}}) = w_\emptyset + w_{\{2\}}, \quad p(a^{\{3\}}) = w_\emptyset + w_{\{3\}}, \\ p(a^{\{1,2\}}) &= w_\emptyset + w_{\{1\}} + w_{\{2\}} + w_{\{1,2\}}, \quad p(a^{\{1,3\}}) = w_\emptyset + w_{\{1\}} + w_{\{3\}} + w_{\{1,3\}}, \\ p(a^{\{2,3\}}) &= w_\emptyset + w_{\{2\}} + w_{\{3\}} + w_{\{2,3\}} \text{ and} \\ p(a^{\{1,2,3\}}) &= w_\emptyset + w_{\{1\}} + w_{\{2\}} + w_{\{3\}} + w_{\{1,2\}} + w_{\{1,3\}} + w_{\{2,3\}} + w_{\{1,2,3\}}. \\ \Leftrightarrow w_{\{1,2,3\}} &= p(a^{\{1,2,3\}}) - w_{\{2,3\}} - w_{\{1,3\}} - w_{\{1,2\}} - w_{\{3\}} - w_{\{2\}} - w_{\{1\}} - w_\emptyset \\ &= p(a^{\{1,2,3\}}) - (p(a^{\{2,3\}}) + p(a^{\{1,3\}}) + p(a^{\{1,2\}})) - 2w_{\{3\}} - 2w_{\{2\}} - 2w_{\{1\}} - 3w_\emptyset \\ &\quad - w_{\{3\}} - w_{\{2\}} - w_{\{1\}} - w_\emptyset \\ &= p(a^{\{1,2,3\}}) - (p(a^{\{2,3\}}) + p(a^{\{1,3\}}) + p(a^{\{1,2\}})) + (w_{\{3\}} + w_{\{2\}} + w_{\{1\}}) + 2w_\emptyset \\ &= p(a^{\{1,2,3\}}) - (p(a^{\{2,3\}}) + p(a^{\{1,3\}}) + p(a^{\{1,2\}})) + (p(a^{\{3\}}) + p(a^{\{2\}}) + p(a^{\{1\}}) - 3w_\emptyset) + 2w_\emptyset \\ &= p(a^{\{1,2,3\}}) - (p(a^{\{2,3\}}) + p(a^{\{1,3\}}) + p(a^{\{1,2\}})) + (p(a^{\{3\}}) + p(a^{\{2\}}) + p(a^{\{1\}})) - p(a^\emptyset) \\ &= \sum_{V \subseteq \{1,2,3\}} (-1)^{3-|V|} p(a^V) \end{aligned}$$

**2.3:** From Axiom 3, we know that the overall individual responsibility  $R_i(a)$  can be composed from the responsibility of several systems  $R_i^s(a)$ . This also holds in the case of unanimity

systems such that:

$$R_i(a) = \sum_{\substack{U \subseteq \mathcal{N} \\ i \in U}} \bar{w}_U R_i^U(a). \quad (\text{A.18})$$

Let us set

$$\bar{w}_U = \frac{w_U}{p(a)}. \quad (\text{A.19})$$

Combining (A.11), (A.17), (A.18) and (A.19), we get

$$R_i(a) = \sum_{\substack{U \subseteq \mathcal{N} \\ i \in U}} \sum_{V \subseteq U} (-1)^{|U|-|V|} \cdot \frac{p(a^V)}{p(a)} \frac{1}{|U|} \quad (\text{A.20})$$

$$= \sum_{V \subseteq \mathcal{N}} \left[ \sum_{\substack{U \subseteq \mathcal{N} \\ V \cup \{i\} \subseteq U}} (-1)^{|U|-|V|} \cdot \frac{1}{|U|} \cdot \frac{p(a^V)}{p(a)} \right] \quad (\text{A.21})$$

Let us define

$$\gamma_i(V) := \sum_{\substack{U \subseteq \mathcal{N} \\ V \cup \{i\} \subseteq U}} (-1)^{|U|-|V|} \cdot \frac{1}{|U|}. \quad (\text{A.22})$$

(A.21) can be written as

$$R_i(a) = \sum_{\substack{V \subseteq \mathcal{N} \\ i \in V}} \gamma_i(V) \frac{p(a^V) - p(a^{V \setminus \{i\}})}{p(a)}. \quad (\text{A.23})$$

To see this, consider two coalitions  $V'$  and  $V = V' \cup \{i\}$ . Note that  $\gamma_i(V') = \gamma_i(V)$ , except for the cases in which  $i \notin V'$ , then  $\gamma_i(V') = -\gamma_i(V)$ .

With  $i \in V$ , there are  $\binom{|\mathcal{N}| - |V|}{|U| - |V|}$  coalitions  $U$  with  $|U|$  elements such that  $V \subseteq U$ . Thus,  $\gamma_i(V)$  can be written as

$$\gamma_i(V) = \sum_{|U|=|V|}^{|\mathcal{N}|} (-1)^{|U|-|V|} \binom{|\mathcal{N}| - |V|}{|U| - |V|} \frac{1}{|U|} \quad (\text{A.24})$$

Note that  $1/a = \int_0^1 x^{a-1} dx$ . Hence,

$$\gamma_i(V) = \sum_{|U|=|V|}^{|\mathcal{N}|} (-1)^{|U|-|V|} \binom{|\mathcal{N}| - |V|}{|U| - |V|} \int_0^1 x^{|U|-1} dx \quad (\text{A.25})$$

$$= \int_0^1 \sum_{|U|=|V|}^{|\mathcal{N}|} (-1)^{|U|-|V|} \binom{|\mathcal{N}| - |V|}{|U| - |V|} x^{|U|-|V|} x^{|V|-1} dx \quad (\text{A.26})$$

$$= \int_0^1 x^{|V|-1} \sum_{|U|=|V|}^{|\mathcal{N}|} (-1)^{|U|-|V|} \binom{|\mathcal{N}| - |V|}{|U| - |V|} x^{|U|-|V|} dx \quad (\text{A.27})$$

$$= \int_0^1 x^{|V|-1} \sum_{|U|=|V|}^{|\mathcal{N}|} \binom{|\mathcal{N}| - |V|}{|U| - |V|} (-x)^{|U|-|V|} dx \quad (\text{A.28})$$

The Binomial Theorem states that  $\sum_{j=0}^k \binom{k}{j} 1^{k-j} (-x)^j = (1-x)^k$ . Hence,

$$\gamma_i(V) = \int_0^1 x^{|V|-1} (1-x)^{|\mathcal{N}|-|V|} dx \quad (\text{A.29})$$

Note that this is an expression of the Euler integral of the first kind. For non-negative values of the exponents, the integral can be expressed as factorials as follows:

$$\int_0^1 x^{|V|-1} (1-x)^{|T|-1} dx = \frac{(|V|-1)! (|T|-1)!}{(|V|+|T|-1)!}. \text{ Substituting } |T| = |\mathcal{N}| - |V| + 1 \text{ gives}$$

$$\gamma_i(V) = \frac{(|V|-1)! (|\mathcal{N}| - |V|)!}{|\mathcal{N}|!}. \quad (\text{A.30})$$

Hence, inserting (A.30) in (A.23) results in

$$R_i(a) = \sum_{\substack{V \subseteq \mathcal{N} \\ i \in V}} \frac{(n - |V|)! \cdot (|V| - 1)!}{|\mathcal{N}|!} \frac{p(a^V) - p(a^{V \setminus \{i\}})}{p(a)} \quad (\text{A.31})$$

$$\Leftrightarrow R_i(a) = \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} \frac{(n - 1 - |C|)! \cdot |C|!}{|\mathcal{N}|!} \frac{p(a^{C \cup \{i\}}) - p(a^C)}{p(a)}, \quad (\text{A.32})$$

with  $V = C \cup \{i\}$ , because  $|V| = |C| + 1$  and  $p(a^V) - p(a^{V \setminus \{i\}}) \equiv p(a^{C \cup \{i\}}) - p(a^C)$ .  $\square$

## A2 Proof of Proposition 2

**Property 1** is obvious.

**Property 2.**

$$R_i(a) \geq R_j(a)$$

$$\begin{aligned}
&\Leftrightarrow \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} \frac{(n-1-|C|)! \cdot |C|! \cdot p(a^{C^i}) - p(a^C)}{|\mathcal{N}|! \cdot p(a)} \geq \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{j\})} \frac{(n-1-|C|)! \cdot |C|! \cdot p(a^{C^j}) - p(a^C)}{|\mathcal{N}|! \cdot p(a)} \\
&\Leftrightarrow \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} (n-1-|C|)! \cdot |C|! \cdot (p(a^{C^i}) - p(a^C)) \geq \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{j\})} (n-1-|C|)! \cdot |C|! \cdot (p(a^{C^j}) - p(a^C))
\end{aligned} \tag{A.33}$$

Note that all  $C \setminus \{i \cup j\}$  and all  $C \supset \{i, j\}$  occur on each side of the inequality with the same frequency. Hence,

$$\begin{aligned}
&\text{(A.33)} \Leftrightarrow \\
&\sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i, j\})} (n-2-|C|)! \cdot |C|! \cdot p(a^{C^i}) - \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i, j\})} (n-2-|C|)! \cdot |C|! \cdot p(a^{C^j}) \\
&\geq \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i, j\})} (n-2-|C|)! \cdot |C|! \cdot p(a^{C^j}) - \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i, j\})} (n-2-|C|)! \cdot |C|! \cdot p(a^{C^i}) \\
&\Leftrightarrow \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i, j\})} (n-2-|C|)! \cdot |C|! \cdot p(a^{C^i}) \geq \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i, j\})} (n-2-|C|)! \cdot |C|! \cdot p(a^{C^j}) \tag{A.34}
\end{aligned}$$

□

If (A.34) holds with equality, it also holds by equivalence of (A.33) that  $R_i(a) = R_j(a)$ . Hence, it is obvious that  $p(a^{C^i}) = p(a^{C^j})$  for all  $C \setminus \{i \cup j\} \Rightarrow R_i = R_j$ . One can easily find counterexamples to show that the reverse implication does not hold. □

**Property 3.**

(i) If  $p(a^{C^j}) - p(a^C) = 0$  for all  $C \in \mathcal{P}(\mathcal{N})$ , then by definition

$$R_j(a) = \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{j\})} \frac{(n-1-|C|)! \cdot |C|! \cdot p(a^{C^i}) - p(a^C)}{|\mathcal{N}|! \cdot p(a)} = 0 \tag{A.35}$$

(ii) By Axiom 2 it holds that  $\sum_{i=1}^n R_i(a) = R(a)$ .

(iii) Hence, for  $R_j(a) = 0$ ,  $\sum_{i \neq j} R_i(a) = R(a)$ , and consequently

$$R_i(a) = R_i(a_1, \dots, a_{j-1}, 0, a_{j+1}, \dots, n) \text{ for all } i. \tag{A.36}$$

□

**Property 4.**

(i) For collective action  $\tilde{a}$ ,  $R_i(\tilde{a}) = \frac{p(\tilde{a}) - p_0}{p(\tilde{a})} = R(\tilde{a})$  and  $R_j(\tilde{a}) = 0$  for all  $j \neq i$ . This follows from the property “irrelevance of dummy actions” and Axiom 2.

(ii) For any collective action  $a$  for which holds that  $p(a) = p(\tilde{a})$ , the collective responsibility  $R(a)$  is equal to  $R(\tilde{a})$ .

(iii) From Axiom 2 follows that  $R_i(a) \leq R(a) = R_i(\tilde{a})$ . □

### A3 Proof of Proposition 3

Note that  $p(a^C) = \beta \cdot \sum_{i \in C} a_i + p_0$  implies  $p(a^{C_i}) - p(a^C) = \beta \cdot a_i$ .

By (5):

$$R_i(a) = \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} \frac{(n-1-|C|)! \cdot |C|!}{|\mathcal{N}|!} \frac{p(a^{C_i}) - p(a^C)}{p(a)} \quad (\text{A.36})$$

$$= \sum_{C \in \mathcal{P}(\mathcal{N} \setminus \{i\})} \frac{(n-1-|C|)! \cdot |C|!}{|\mathcal{N}|!} \frac{p(a^i) - p_0}{p(a)} = \frac{|\mathcal{N}|!}{|\mathcal{N}|!} \frac{\beta a_i}{p(a)} \quad (\text{A.37})$$

$$\Leftrightarrow R_i(a) = \frac{\beta \sum_{i \in \mathcal{N}} a_i}{\beta \sum_{i \in \mathcal{N}} a_i} \frac{\beta a_i}{p(a)} = \frac{a_i}{\sum_{i \in \mathcal{N}} a_i} \frac{\beta \sum_{i \in \mathcal{N}} a_i}{p(a)} = \frac{a_i}{\sum_{i \in \mathcal{N}} a_i} \frac{p(a) - p_0}{p(a)} \quad (\text{A.38})$$

$$\Leftrightarrow R_i(a) = \frac{a_i}{\sum_{i \in \mathcal{N}} a_i} R(a) \quad (\text{A.39})$$

□