Responsibility for regime shifts in managed ecosystems

Stefan Baumgärtner*

University of Freiburg, Germany

August 14, 2019

Abstract: I develop a quantitative measure of a manager's responsibility for a regime shift in a managed ecosystem with stochastic dynamics. I build on an established and clearly defined concept of responsibility, which I operationalize in a simple generic model. *Causal responsibility* is the degree of causation of an outcome due to the manager's action, which is in contrast to chance influences ("good luck" or "bad luck") that may also have caused the outcome. *Normative responsibility* is the manager's obligation to see to it that the system does, or does not, undergo a regime shift. It implies a particular management action. *Virtuous responsibility* is the degree to which the manager lives up to her normative responsibility when taking a management action. The quantitative measurement of responsibility is relevant to judge the quality of different management actions, to reward or punish the manager based on the extent of her (ir)responsibility, and to design institutions that enable and encourage responsible management of ecosystems with potential regime shifts.

JEL-Classification: Q57

Keywords: accountability, ascription, ecosystem management, causation, norms, regime shift, responsibility

*Correspondence: Chair of Environmental Economics and Resource Management, University of Freiburg, Tennenbacher Str. 4, D-79106 Freiburg, Germany, phone: +49.761.203-3753, email: stefan.baumgaertner@ere.uni-freiburg.de.

1 Introduction

Many ecosystems that are managed for the ecosystem services they provide to people can potentially be in different locally stable system states that are separated by critical thresholds (Holling 1973, Levin 1999, Muradian 2001, Walker and Meyers 2004). Regime shifts from one state of the system into another one are often catastrophic: they are abrupt, (quasi-)irreversible, and imply a significant decrease in human welfare (Arrow et al. 1995, Levin et al. 1998, Scheffer et al. 2001). Examples include fish stocks, rangelands, agricultural soils, groundwater tables, or the Earth-climate system at large (Gunderson and Pritchard 2002, Folke et al. 2004). In general, such ecosystems exhibit a stochastic dynamics. Hence, regime shifts are essentially stochastic events which may happen due to natural circumstances beyond human control, such as e.g. weather, diseases or pests. Yet, (mis)management by human actors has a large impact on the state of the system and, thus, may also cause regime shifts.

One major and systematic challenge for assessing and incentivicing good management of such systems is the following. In a managed stochastic ecosystem the outcome – whether a regime shift actually occurs or not – is not just a matter of good or bad management; it is also essentially influenced by pure chance. Normally, good management results in a good system state and bad management results in a bad system state. However, a bad state of the system may occur even under good management ("bad luck"), and a good state of the system may occur even under bad management ("good luck"). Roughly speaking, in a stochastic system even for a given management everything can still happen with some probability. Therefore, the actually realized system state is not a good indicator of the quality of the management.

This raises the question of what is the responsibility of the manager of such a system for a regime shift? More exactly: To what extent is an actually occurring regime shift caused by management, and to what extent is it simply due to pure chance? Also: To what extent does a manager who employs a particular management action actually live up to an obligation to manage the system according to a given norm, say, to keep the system in the good state and prevent a regime shift into the bad state?

Here, I develop a quantitative measure of a manager's responsibility for a regime shift in a managed ecosystem with stochastic dynamics. To that end, I build on the well established and clearly defined concept of responsibility (Baumgärtner et al. 2018), which I operationalize in a simple and generic ecological-economic model. The notion of responsibility comes with several meanings, both descriptive and normative. *Causal responsibility* is the degree of causation of an outcome due to the manager's action, which is in contrast to chance influences ("good luck" or "bad luck") that may also have caused the outcome. *Normative responsibility* is the manager's obligation to see to it that the system does, or does not, undergo a regime shift. It implies a particular management action. *Virtuous responsibility* is the degree to which the manager lives up to her normative responsibility when taking a given management action.

The quantitative measurement of responsibility is relevant to judge the quality of different management actions, to reward or punish the manager based on the extent of her (ir)responsibility, and to design institutions that enable and encourage responsible management of ecosystems with potential regime shifts.

The paper is organized as follows. In Section 2, I review the concept of responsibility, which is the conceptual basis for the subsequent formalization and measurement. In Section 3, I introduce a simple generic model of a managed stochastic ecosystem with potential regime shifts. In Section 4, I introduce and discuss the quantitative measures of the three layers of responsibility – causal, normative, and virtuous – in terms of the model. In Section 5, I develop a number of extensions and generalizations. Section 6 concludes.

2 The concept of responsibility[1]

Responsibility is a core concept in morals, philosophical ethics, law, organisations and politics. The concept links abstract and general norms to the concrete and specific action context in a situation where the actor has discretionary scope for acting freely.

[1]This section is based on Baumgärtner et al. (2018: Section 3.1). The exposition here is in parts literally taken from this source.

In such instances, the concept of responsibility establishes an architecture of argument to assess and guide actions with regard to given norms.

Responsibility is a multi-layered notion. At least three different, yet related, aspects of the notion have been distinguished (e.g. Hart 1968: 211–230, Jonas 1979: 172, Bovens 1998: Chap. 3).

(1) Causal responsibility

The notion of "responsibility" in the phrase "someone is responsible for something" is often used to describe a causal relation. If used in this sense, it has the following precise meaning: "Person A is [causally] responsible for X if and only if A's action is the cause of X."

Causal responsibility ascribes the consequences of some action to the perpetrator ("Who did it?").[2] This requires that some change X in the state of the system is in fact a causal effect of the person's action. For example, one may say that an ecosystem manager [person A] is responsible for a regime shift in an ecological-economic system [system state X] to state that the regime shift was caused by this manager's action, rather than by some other person's actions or by natural circumstances which may, potentially, also have caused the regime shift.

Causal responsibility is purely descriptive. It has, as such, no moral content or implication: it "is a precondition of morality, but not morality itself" (Jonas 1979: 179, translation by Baumgärtner et al. 2018: 14). Hence, causal responsibility exists and can be assessed independently of any norms. It is a necessary condition of moral assessment, as a person can only be morally praised or blamed for an action, and its consequences, that can be ascribed to her.

(2) Normative responsibility

The notion of "responsibility" in the phrase "someone is responsible for something" is also often used to describe an obligation. Then, it has the following meaning (Goodin

[2]Here, for clarity and simplicity, I speak of "causal" responsibility to denote what Baumgärtner et al. (2018: Section 3.1) call "ascriptive" responsibility.

1986: 50, similarly Bovens 1998: 25): "Person A is responsible for X [in the *normative* sense] if and only if A ought to see to it that X".

To say that the acting person is responsible in the normative sense means that she ought to see to it that certain consequences result from her actions and that these consequences meet certain well-defined standards. This imposes an imperative on the person to act accordingly. For example, one may say that an ecosystem manager [person A] is responsible for preventing a regime shift in an ecological-economic system [system state X] to state that the manager ought to act such as to prevent a regime shift.

In the definition of normative responsibility, the normative condition is specified in terms of the potential consequence X of A's action rather than in terms of the action itself, and X is related to A's actions. This means, for normative responsibility the desired result is specified and several alternative ways exist to achieve the result. So, there is a discretionary freedom, and an obligation, for the responsible person to choose an appropriate action, that is, an action the consequences of which actually meet the normative standard.[3]

Causal responsibility is a precondition for normative responsibility. For, a person can only have obligations for those consequences of which she is the perpetrator, that is, which can be ascribed to her. One can generally only be responsible in the normative sense for what is in one's power, where power is the ability to actually bring about some intended effect through one's action. If one does not have the power to bring about, or to prevent, a certain effect, one cannot be obliged to bring about, or to prevent, this effect.[4]

[3] This distinguishes a normative responsibility from a duty. As Goodin (1986: 50) points out, also a duty can be formalized in the form "A ought to see to it that X". However, in the case of a duty, X takes the form "A does or refrains from doing Y", where Y is some specific action.

[4] It is a long-standing and generally accepted principle, which has its origin in Roman law, that no one is obligated beyond what she is able to do ("nemo ultra posse obligatur"). In modern ethics, this has come to be known as the Ought-Implies-Can Principle, which has been introduced into modern ethical discussion by Immanuel Kant (1999[1781]: A548/B576, p. 473; 1998[1793]: 6:50, p. 94). It means that normative responsibility presupposes the possibility, at least in principle, of compliance, that is, one can only be obliged to do what one actually can at least in principle do.

If a person carries normative responsibility and, therefore, has an obligation that the consequences of her actions meet certain standards, she must be ready to give account of their fulfilment or non-fulfilment ("accountability"). That is, she must be ready to respond to questions about how she has used her discretionary freedom when acting such as to meet the obligation.

(3) Virtuous responsibility

The action of a person is often characterized as being "responsible" or "not responsible". This characterization refers to whether the person actually lives up to her normative responsibility. One says: "Person A acts *responsibly* [in the virtuous sense] for given X, if and only if A actually sees to it that X."[5]

For the assessment of whether person A acts responsibly [in the virtuous sense], the success in terms of the actual realization of the normatively desired system state X is not the primary criterion. Rather, whether person A lives up to her normative responsibility depends on her actual efforts to pursue the goal – whether she *has actually seen to it* that X.

Whether one acts responsibly or not is the basis for praise and blame. If one judges a person as acting "responsibly" or "irresponsibly", one refers to responsibility in this virtuous sense. If one says that a person acts "irresponsibly" this means that the person's actions do not meet the obligations imposed by the underlying normative responsibility: this person does not live up to her normative responsibility. However, it does *not* mean that she is not the originator of this action and its consequences (causal responsibility), nor does it mean that normative standards do not apply (normative responsibility): to be irresponsible is not to be not-responsible.

What is the right action to live up to one's (normative) responsibility cannot be derived unambiguously from the underlying normative responsibility itself. Rather,

[5]Williams (2008: 459) defines this third meaning of responsibility in an even farther reaching way: "Responsibility represents the readiness to respond to a plurality of normative demands" (similarly Bovens 1998: 26 and Baumgärtner et al. 2018: 17).

there are typically several options to live up to one's (normative) responsibility.[6] Also, it is often not a binary distinction, whether someone has acted responsibly or not, but one may act responsibly to some extent (Goodin 1986: 53–55).

3 Model

Consider an ecosystem that may potentially be in either one of two different locally stable states $s \in \{s_i, s_o\}$. In the initial state s_i, an exogenous stochastic event may trigger the system to flip into the other potential state s_o. Such a regime shift occurs with an objective probability p with $0 \leq p \leq 1$. Conversely, the system remains in the initial state with probability $1-p$. For example, think of a shallow lake that may flip from an oligotrophic to a eutrophic state due to a storm stirring up the nutrient sediments from the lake ground, a fish stock that may collapse due to the outbreak of a disease, a savanna rangeland that may flip from grass-dominated to bush-dominated after a long time period without any natural fire, or an agricultural soil that may salinate due to a rising groundwater table. Initially, the regime-shift probability is $p = p_0$ with $0 \leq p_0 \leq 1$.

While the potential regime shift between the initial and the other state may be due to natural circumstances and is insofar essentially stochastic, management of the system has an effect, too, and may even dominate the natural stochasticity. In particular, management affects the probability of a regime shift. For example, nutrient inflow from nearby agriculture increases the risk that a storm flips a shallow lake from an oligotrophic to a eutrophic state, overharvesting reduces the genetic variability of a fish stock and thus increases its susceptibility to disease, suppression of fire in a savanna rangeland increases the risk of a flip to a bush-dominated state, and irrigation of an agricultural land will raise the ground water table and thus increase the risk of salination.

The system manager autonomously chooses a management action under uncertainty,

[6] This is a fundamental difference between a (normative) responsibility and a duty (see Footnote 3). While a duty directly determines the action that is necessary for its fulfilment, a (normative) responsibility asks for a certain result, leaving it open how to reach this result through action Goodin (1986).

that is, before the uncertainty is resolved and one of the two potential system states is actually realized. The management action is described by a continuous and normalized variable $a \in [0,1]$. Here, $[0,1]$ is the set of all feasible actions, which contains infinitely many different actions; that is, the manager can choose from a set of many different alternatives. Action a affects the probability of the regime shift from the initial into the other system state, such that the resulting regime-shift probability is given by a continuously differentiable and strictly monotonically increasing function of action:[7]

$$\begin{aligned} p = p(a) \qquad &\text{with} \qquad p'(a) > 0 \text{ for all } a \\ &\text{and} \qquad p(0) = p_{\min},\ p(1) = p_{\max},\ \text{with } 0 \leq p_{\min} \leq p_0 \leq p_{\max} \leq 1 \ . \end{aligned} \tag{1}$$

In words, the stronger the action, i.e. the larger a, the higher the probability that the system flips into the other state due to exogenous disturbance; with no action, the regime-shift probability is minimal but may still be greater than zero; and with maximal action, the regime-shift probability is maximal but may still be smaller than one. For example, a may be the intensity of fertilizer use in agriculture, the harvest intensity of a fish stock, the grazing intensity of a savanna rangeland, or the extraction rate from a groundwater reservoir.

The minimal regime-shift probability $p_{\min}$ is smaller than the initial regime-shift probability p_0, and the maximal regime-shift probability $p_{\max}$ is larger than the initial regime-shift probability p_0. That is, the action may change the regime-shift probability in both ways: it may increase or decrease the regime-shift probability. Hence, there is exactly one feasible management action $a_0 \in [0,1]$ that does not change the regime-shift probability from its initial level. This *probabilistically neutral action* a_0 is defined through[8]

$$p(a_0) = p_0 \ . \tag{2}$$

[7]The assumption of monotonicity is for simplification. It avoids the hassle of dealing with case distinctions which would be necessary if $p(a)$ was non-monotonic. Then, it is without loss of generality to assume that $p(a)$ is increasing in a.

[8]Equation (2) has a unique solution a_0 because of strict monotonicity of $p(\cdot)$ and $p_{\min} \leq p_0 \leq p_{\max}$ (Property 1).

If the manager takes the probabilistically neutral action a_0, she does not affect the probabilistic condition of the system. The managed system is still in the initial probabilistic condition, and no observer of the system could tell that there is some management affecting the system. In that sense, this action is neutral. In many systems and for many initial probabilistic conditions, this probabilistically neutral action will be to simply not act at all. Yet, it is imaginable that the initial regime-shift probability is only maintained through some real action. In such a case, not acting at all would not maintain the initial probabilistic condition of the system.

While the probabilistically neutral action does not change the regime-shift probability from its initial level, it does not necessarily preserve the initial state of the system. For, even with the probabilistically neutral action there may actually still be a regime shift into the other state of the system – namely with probability p_0. All actions $a < a_0$ lead to a regime-shift probability $p(a)$ which is smaller than the initial regime-shift probability p_0; all actions $a > a_0$ lead to a regime-shift probability $p(a)$ which is larger than the initial regime-shift probability p_0.

For graphical and numerical illustration, I employ the following specific relationship between the action a and the regime-shift probability p (cf. Figure 1):

$$p(a) = p_{\min} + (p_{\max} - p_{\min})\, a \; . \tag{3}$$

This specific model has the fundamental properties (1). In addition, it assumes a linear relationship between the action a and the resulting regime-shift probability $p(a)$. As the action variable a is a stylized variable and not an empirically meaningful and measurable variable, any statement about the curvature of $p(a)$ would not be meaningful either. Therefore, assuming a linear relationship is not more restrictive than the model (1) itself. In the linear model (Equation 3), the probabilistically neutral action is

$$a_0 = \frac{p_0 - p_{\min}}{p_{\max} - p_{\min}} \; . \tag{4}$$

The conditional probability that the system eventually is in state s when the manager has taken action a is denoted by $P(s\,|\,a)$. With model (1), it is given by

$$P(s\,|\,a) = \begin{cases} p(a) & \text{for } s = s_o \\ 1 - p(a) & \text{for } s = s_i \end{cases} \; . \tag{5}$$

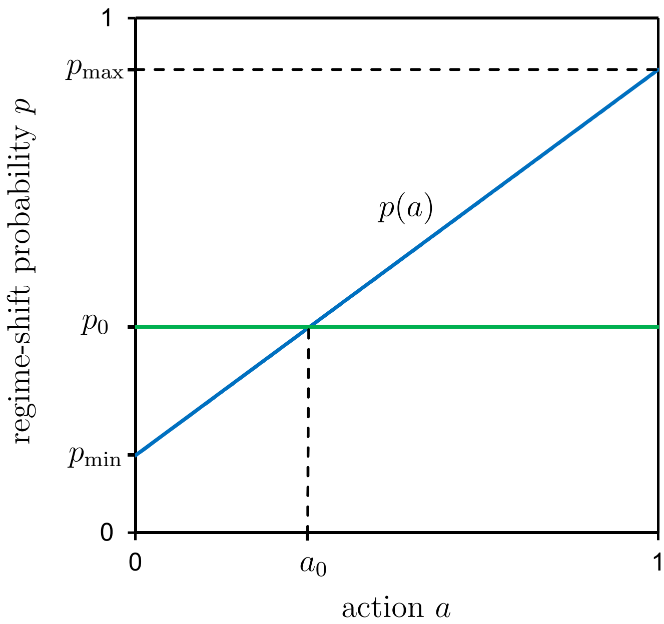

Figure 1: Initial regime-shift probability p_0 (green line), regime-shift probability p as a function of management action a (blue line) according to the linear model, $p(a) = p_{\min} + (p_{\max} - p_{\min})\, a$ (Equation 3), and probabilistically neutral action a_0 (Equation 4). Parameter values: $p_0 = 0.4$, $p_{\min} = 0.15$, $p_{\max} = 0.9$.

At the time of taking action, the manager acts under uncertainty. That is, she does not know for certain which outcome – system state s_i or s_o – will actually result from her action. The best that she can know in this probabilistic setting is the objective probability distribution over these two potential outcomes, and how any of her feasible actions affects this probability distribution. I assume that the manager at the time of action has complete and true knowledge of this probability distribution and how it is affected by any of her feasible actions, that is, she knows the probability functions (1) and (5).

After the manager has taken action, uncertainty is resolved and one of the two potential system states, s_i or s_o, is actually realized.

Note that, to simplify matters and establish a clear conceptual focus in terms of responsibility, the model considered here makes a number of restrictive (and in part unrealistic) assumptions. It assumes that (1) the manager acts autonomously and could have acted otherwise, as alternative actions are feasible; (2) the manager completely and truly knows what she is doing, that is, she knows the objective probability distribution over all potential system states and how any of her feasible actions affects this probability distribution; and (3) there is neither prior history nor subsequent future to the action considered here, that is, neither the manager's set of feasible actions nor her knowledge are influenced by her previous actions, and the action actually taken has no other than the immediate consequences on the system state described here. Deviation from any of these restrictive assumptions opens up a suite of deep philosophical problems in the concept of responsibility. Here, I want to avoid all those and focus on just one problem in the concept of responsibility – namely that of acting under outcome risk.

4 Measuring responsibility

I will now formally specify and measure the three different concepts of the manager's responsibility for the state of the system – causal, normative, and virtuous – in terms of the model.

4.1 Causal responsibility

My measure of causal responsibility essentially builds on the idea of Vallentyne (2008) of how to attribute "partial responsibility" to an agent in a probabilistically uncertain situation where both agency and brute luck play a role in bringing about an outcome. The core of Vallentyne's (2008) idea is that the agent's partial responsibility for an outcome is the change in the outcome's probability which is due – directly and indirectly – to agency. Let me apply this idea in the situation considered here – a potential regime shift of a managed system.

Initially – that is, prior to the manager's interference with the system – the system is in state s_i. Which system state $s \in \{s_i, s_o\}$ will eventually occur is uncertain. There is a probability of $P(s \mid a_0)$ for state s to occur. In this situation, the manager takes action $a \in [0, 1]$, which affects the probability of state s occurring to $P(s \mid a)$. Eventually, uncertainty is resolved and the actual state s of system is realized, which may be either the initial state (no regime shift), $s = s_i$, or the other state (regime shift), $s = s_o$. With this sequence of events, what is the manager's causal responsibility for the realization of the actual system state s? In other words, to what extent is it the manager's action a that has caused the actual system state s to occur, and – conversely – to what extent is this outcome due to brute luck?

4.1.1 Ex-post causal responsibility

To start the treatment of causal responsibility, I take an ex-post perspective. I take the eventual realization of the actual state $s \in \{s_i, s_o\}$ of the system as given and, from this result, look back at the chain of events that has lead to this *explanandum state* ("the state to be explained") – starting from state s having a probability of zero all the way to this state actually occurring with probability one. There are three steps in this chain, each of which has potentially affected the uncertainty and the outcome. For each of these steps, I consider the (shift in) probability of the outcome, i.e. state s occurring, and whether this is to be attributed to brute luck to or the manager's action (Table 1 presents an overview):

(1) Initial situation. There is an initial probability of $P(s \mid a_0)$ for state s to occur.[9] If the explanandum state is s_o (regime shift), this initial probability is p_0; if the explanandum state is s_i (no regime shift), this initial probability is $1 - p_0$ (from Equations 5 and 1). This is the situation prior to the manager's interference with the system. Whatever would have resulted from this initial situation of uncertainty, obviously, cannot be caused by the manager's action. Hence, this probability of $P(s \,|\, a_0)$ of state s, and what resulted from it, cannot be attributed to the manager's action. It is entirely due to brute luck, and therefore entirely attributed to brute luck.

(2) Manager's action. In this initial situation the manager takes action a. This directly affects the probability of explanandum state s: it adds an amount $P(s \,|\, a) - P(s \,|\, a_0)$ to the probability of this state. This amount may be positive or negative. That is,

[9] Here, I use the equivalence-by-definition of Equation (2) to express the initial regime-shift probability p_0 through the conditional probability $P(s \,|\, a)$ and the probabilistically neutral action a_0. This serves to have the complete probability accounting in terms of $P(s \,|\, a)$.

action a may make the occurrence of state s more or less likely than in the initial situation. If the explanandum state is s_o (regime shift) this probability shift is $p(a) - p_0$, which is positive for $a > a_0$ and negative for $a < a_0$; if the explanandum state is s_i (no regime shift) this probability shift is $p_0 - p(a)$, which is positive for $a < a_0$ and negative for $a > a_0$ (from Equations 5 and 1). Hence, taking an action stronger than the probabilistically neutral action, $a > a_0$, makes the regime shift to the other state more likely, and the remaining in the initial state less likely, than in the initial situation. And conversely, taking an action weaker than the probabilistically neutral action, $a < a_0$, makes the regime shift to the other state less likely, and the remaining in the initial state more likely, than in the initial situation.

Obviously, this shift in the probability of the explanandum state s is directly and entirely due to the manager's action. As the manager has – by assumption – acted in full autonomy and with complete and true knowledge of the consequences of her actions on the probability distribution, the probability shift $P(s \,|\, a) - P(s \,|\, a_0)$ is entirely attributed to the manager's action.[10]

(3) Resolution of uncertainty. With the manager's action and its direct impact on the probability of state s (step 2), the outcome is still uncertain. The explanandum state s will occur with probability $P(s \mid a)$, and it will not occur with probability $1 - P(s \,|\, a)$. In particular, state s_o (regime shift) will occur with probability $p(a)$ and state s_i (no regime shift) will occur with probability $1 - p(a)$ (from Equation 5). Now, the third and final step in the chain that leads to the actual realization of the explanandum state s is the resolution of uncertainty. In this step, the probability of state s increases by such a positive amount that it becomes one, i.e. it is actually realized. Thus, the probability shift in step 3 is $1 - P(s \,|\, a)$. If the explanandum state is s_o (regime shift), this probability shift is $1 - p(a)$; if the explanandum state is s_i (no regime shift), this probability shift is $p(a)$ (from Equation 5).

This positive probability shift towards the actual realization of the explanandum

[10]If the manager had not acted autonomously, or had incomplete or false beliefs about the consequences of her actions, this could lead to attributing the probability shift $P(s \,|\, a) - P(s \,|\, a_0)$ not entirely, or even not at all, to the manager.

state s is "outcome luck". It is a third instance of probability shift in the chain of causality, in addition to the initial uncertainty (step 1) and the manager's action (step 2). Should this outcome luck be attributed to brute luck or to the manager's action? Both the initial situation, which is due to brute luck, and the manager's action have contributed to the probability $P(s \mid a)$, which establishes the outcome luck in the actual realization of state s. It is therefore plausible that the outcome luck is attributed to both brute luck and the manager's action. I follow Vallentyne's (2008) suggestion to attribute the outcome luck, i.e. the probability change $1 - P(s \mid a)$, to brute luck and to the manager's action in proportion to their contribution to establishing the outcome luck.

To quantitatively determine the relative shares of outcome luck attributed to either of these two causes in proportion, one needs to distinguish between two conceptually different situations: (i) action a *increases* the probability of the explanandum state s, $P(s \mid a) > P(s \mid a_0)$, which is the case for an action $a > a_0$ when explaining a regime shift to the other state s_o (Figure 2 bottom right), or for an action $a < a_0$ when explaining the remaining in the initial state s_i (Figure 2 top left); (ii) action a *decreases* the probability of the explanandum state s, $P(s \mid a) < P(s \mid a_0)$, which is the case for an action $a < a_0$ when explaining a regime shift to the other state s_o (Figure 2 bottom left), or for an action $a > a_0$ when explaining the remaining in the initial state s_i (Figure 2 top right). Vallentyne (2008) only considers situation (i), and only through a numerical example. Here, I formulate his argument about situation (i) in a general and formal manner, and I develop a suggestion for situation (ii) which is in parallel with Vallentyne's argument for situation (i).

In situation (i), the outcome luck exists because with the manager's action there is a probability of $P(s \mid a)$ for the explanandum state s, which implies a probability gap of $1 - P(s \mid a)$ to be closed by outcome luck for this state to actually occur. The outcome-luck-establishing probability $P(s \mid a)$ is the result of two contributions: the initial probability $P(s \mid a_0)$ of this state due to brute luck (step 1), and the probability shift $P(s \mid a) - P(s \mid a_0) > 0$ due to the manager's action which has made the occurrence of the explanandum state more likely (step 2). The first contribution has a relative

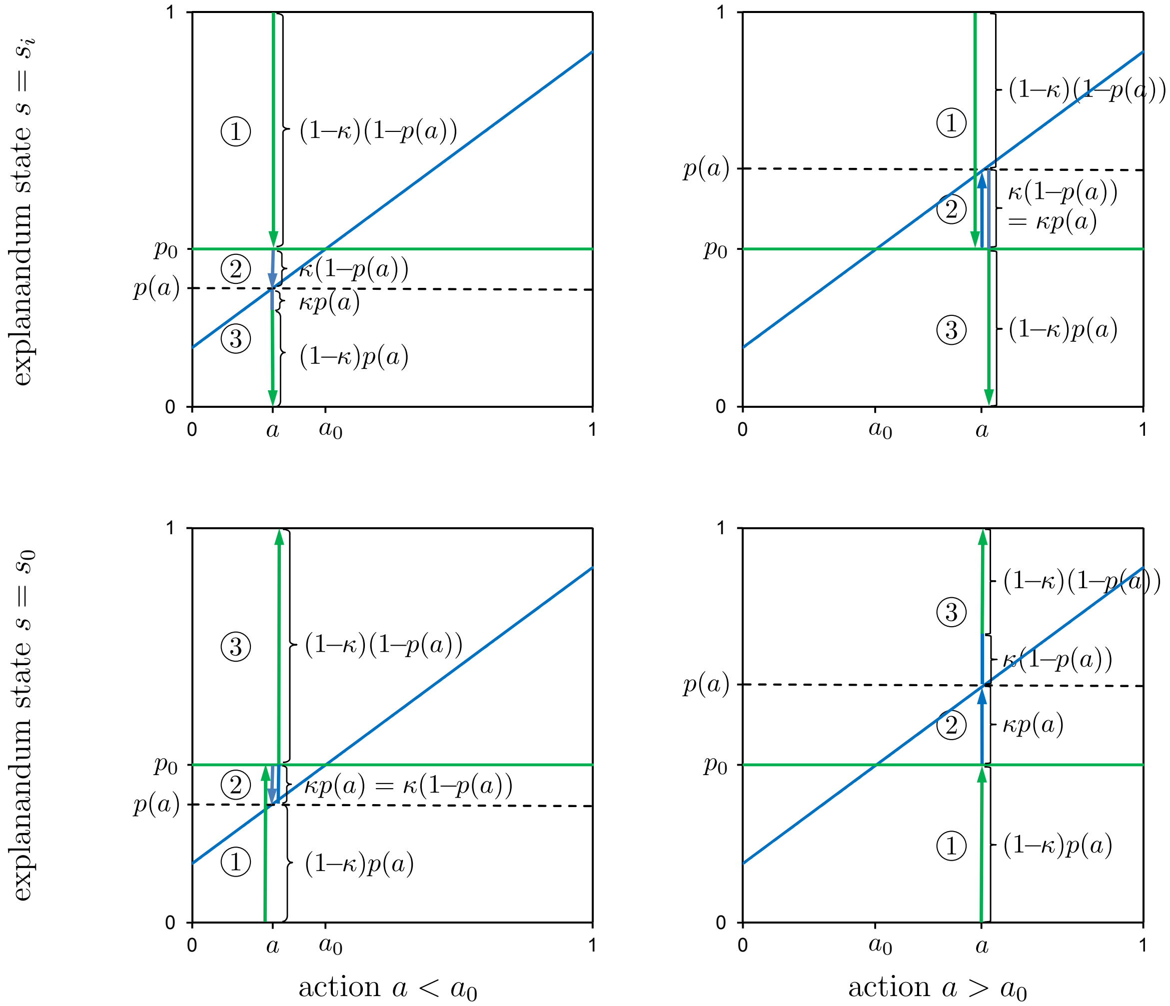

Figure 2: Probability shifts in steps (1), (2) and (3) for the explanandum state $s = s_i$ (top) and $s = s_o$ (bottom), and for an action a weaker (left) and stronger (right) than the probabilistically neutral action a_0. Arrows in blue denote probability shifts attributed to action; arrows in green denote probability shifts attributed to brute luck. Parameter values: $p_0 = 0.4$, $p_{\min} = 0.15$, $p_{\max} = 0.9$, $a = 0.2$ (left), 0.6 (right). Remember (from Figure 1) that each diagram shows the probability of a regime shift (green line: initially; blue line: with action). Hence, for the explanandum state $s = s_o$ (bottom) the chain of causality goes from $p = 0$ to $p = 1$ in this diagram, whereas for the explanandum state $s = s_i$ (top) the chain of causality goes from $1 - p = 0$, i.e. $p = 1$, to $1 - p = 1$, i.e. $p = 0$, in this diagram.

share of $P(s \mid a_0)/P(s \mid a)$, and the second contribution has a relative share of $(P(s \mid a) - P(s \mid a_0))/P(s \mid a)$, of the outcome-luck-establishing probability $P(s \mid a)$. Therefore, a relative share of $P(s \mid a_0)/P(s \mid a)$ of outcome luck is attributed to brute luck, and a relative share of $(P(s \mid a) - P(s \mid a_0))/P(s \mid a)$ of outcome luck is attributed to the manager's action. Both relative shares add up to one, so that the entire outcome luck is attributed to brute luck and to action. This implies that the larger the probability shift due to action, the larger the proportion of outcome luck also attributed to action and the smaller the proportion of outcome luck attributed to brute luck. At minimum, the proportion of outcome luck attributed to action is zero when the action a does not shift the probability of the explanandum state at all, i.e. for $a = a_0$. In this case, none of the outcome luck is attributed to action; it is all attributed to brute luck. At maximum, this proportion is one when the action a has solely established the probability of the explanandum state, which is the case when the initial probability of the explanandum state was zero, i.e. $P(s \mid a_0) = 0$. In this case, the entire outcome luck is attributed to action; none of it is attributed to brute luck.

In situation (ii), the outcome luck consists in the probability gap of $1 - P(s \mid a)$ to be closed for the explanandum state s to actually occur. This probability gap is the result of two contributions: the initial probability gap of $1 - P(s \mid a_0)$ due to brute luck (step 1) and the additional probability gap of $(1 - P(s \mid a)) - (1 - P(s \mid a_0)) = P(s \mid a_0) - P(s \mid a) > 0$ due to the manager's action which has made the occurrence of the explanandum state less likely (step 2). Therefore, it is immediately obvious that an absolute amount of $1 - P(s \mid a_0)$ should be attributed to brute luck, and an absolute amount of $P(s \mid a_0) - P(s \mid a) > 0$ to the manager's action. Both add up to the entire absolute amount of outcome luck. While this already answers the question what part of outcome luck should be attributed to action and what part to brute luck, for a general and consistent treatment one can also express this result in terms of relative shares, as in situation (i). The difference to situation (i) is the choice of base for calculating relative shares, which needs to be different because of the negative sign of the direct probability sift due to the manager's action, $P(s \mid a) - P(s \mid a_0) < 0$, in situation (ii).

Again, the probability gap which makes the outcome luck is the result of two contributions: the initial probability gap of $1 - P(s \,|\, a_0)$ due to brute luck (step 1) and the additional probability gap of $(1 - P(s \,|\, a)) - (1 - P(s \,|\, a_0)) = P(s \,|\, a_0) - P(s \,|\, a) > 0$ due to the manager's action which has made the occurrence of the explanandum state less likely (step 2). The first contribution has a relative share of $(1 - P(s \,|\, a_0))/(1 - P(s \,|\, a))$, and the second contribution has a relative share of $(P(s \,|\, a_0) - P(s \,|\, a))/(1 - P(s \,|\, a))$, of the total outcome luck of $1 - P(s \mid a)$. Therefore, a relative share of $(1 - P(s \mid a_0))/(1 - P(s \,|\, a))$ of outcome luck is attributed to brute luck, and a relative share of $(P(s \,|\, a_0) - P(s \,|\, a))/(1 - P(s \,|\, a))$ of outcome luck is attributed to the manager's action. Again, both relative shares add up to one, so that the entire outcome luck is attributed to brute luck and to action. And as in situation (i), the larger the probability shift due to action, the larger the proportion of outcome luck also attributed to action and the smaller the proportion of outcome luck attributed to brute luck. At minimum, the proportion of outcome luck attributed to action is zero when the action a does not shift the probability of the explanandum state at all, i.e. for $a = a_0$. In this case, none of the outcome luck is attributed to action; it is all attributed to brute luck. At maximum, this proportion is one when the action a has solely established the outcome luck, which is the case when the initial probability of the explanandum state was one, i.e. $P(s \,|\, a_0) = 1$. In this case, the entire outcome luck is attributed to action; none of it is attributed to brute luck.

Taking situations (i) and (ii) together, the relative share $\kappa(s, a)$ of outcome luck attributed to the manager's action a when state s is actually realized is given by

$$\kappa(s, a) = \begin{cases} \dfrac{P(s \,|\, a) - P(s \,|\, a_0)}{P(s \,|\, a)} & \text{for } P(s \,|\, a) > P(s \,|\, a_0) \\ 0 & \text{for } P(s \,|\, a) = P(s \,|\, a_0) \\ \dfrac{P(s \,|\, a_0) - P(s \,|\, a)}{1 - P(s \,|\, a)} & \text{for } P(s \,|\, a) < P(s \,|\, a_0) \end{cases} . \qquad (6)$$

Conversely, a relative share $1 - \kappa(s, a)$ of outcome luck is attributed to brute luck, where

$$1 - \kappa(s,a) = \begin{cases} \dfrac{P(s\,|\,a_0)}{P(s\,|\,a)} & \text{for } P(s\,|\,a) > P(s\,|\,a_0) \\ 1 & \text{for } P(s\,|\,a) = P(s\,|\,a_0) \\ \dfrac{1 - P(s\,|\,a_0)}{1 - P(s\,|\,a)} & \text{for } P(s\,|\,a) < P(s\,|\,a_0) \end{cases} . \tag{7}$$

From the considerations along the three steps it is apparent that the probability changes add up (Table 1). The initial probability of the explanandum state s (step 1) plus the probability change due to action a (step 2) plus the probability change due to outcome luck (step 3) yield the finally realized probability of this state s which is one (line 4): $P(s\,|\,a_0) + [P(s\,|\,a) - P(s\,|\,a_0)] + [1 - P(s\,|\,a)] = 1$.

To obtain the manager's ex-post causal responsibility for the actual realization of state s when taking action a, one must calculate – following the suggestion by Vallentyne (2008) – the total shift in probability of state s that is attributed to the manager's action. From the considerations above, there are two probability shifts to be taken into account: the direct probability shift due to action a, $P(s \mid a) - P(s \mid a_0)$ (step 2), and a share $\kappa(s,a)$ of the probability shift due to outcome luck, $1 - P(s\,|\,a)$ (step 3). With $\kappa(s,a)$ from Equation (6), this yields:

$$\begin{aligned} & P(s\,|\,a) - P(s\,|\,a_0) + \kappa(s,a)\,[1 - P(s\,|\,a)] \\ & = \begin{cases} \dfrac{P(s\,|\,a) - P(s\,|\,a_0)}{P(s\,|\,a)} & \text{for } P(s\,|\,a) > P(s\,|\,a_0) \\ 0 & \text{for } P(s\,|\,a) \leq P(s\,|\,a_0) \end{cases} . \end{aligned} \tag{8}$$

Note that there is one important qualitative difference between situations (i) and (ii), where $P(s \mid a) > P(s \mid a_0)$ and $P(s \mid a) \leq P(s \mid a_0)$, respectively. In situation (i), the positive absolute amount of outcome luck attributed to action in step (3) adds to the positive direct probability shift due to action in step (2), thus making the total probability shift attributed to action larger than the direct probability shift due to action (Figure 2 bottom right and top left). In contrast, in situation (ii) the positive absolute amount of outcome luck which is due to, and therefore attributed to, action, $P(s\,|\,a_0) - P(s\,|\,a) > 0$, is exactly the same absolute amount, yet of opposite sign, as the direct negative probability shift due to action in step (2), $P(s\,|\,a) - P(s\,|\,a_0) < 0$

(Figure 2 bottom left and top right). Thus, the two exactly cancel. In situation (ii), therefore, the total absolute amount of probability shift attributed to action is zero.

One can also calculate the total shift in probability of state s that is attributed to brute luck. From the considerations above, there are two probability shifts to be taken into account: the initial probability, $P(s \mid a_0)$ (step 1), and a share $1-\kappa(s,a)$ of the probability shift due to outcome luck, $1-P(s \mid a)$ (step 3). With $1-\kappa(s,a)$ from Equation (7), this yields:

$$P(s|a_0) + [1-\kappa(s,a)]\,[1-P(s|a)] \;=\; \begin{cases} \dfrac{P(s|a_0)}{P(s|a)} & \text{for } P(s|a) > P(s|a_0) \\ 1 & \text{for } P(s|a) \leq P(s|a_0) \end{cases} . \quad (9)$$

Obviously, the total probability shift attributed to the manager's action (Equation 8) and the total probability shift attributed to brute luck (Equation 9) add up to one – the final probability of the explanandum state s. Hence, the considerations so far in this section (summarized in Table 1) establish a complete and consistent accounting of probability shifts and their attribution to either the manager's action or to brute luck. This motivates the following definition of a measure of the manager's ex-post causal responsibility for the actually realized state s of the system.

	probability (in lines 1, 4) and shift in probability (in lines 2, 3) of state s		attributed to	
(1)	P initial	$P(s\|a_0)$		brute luck
(2)	$+\,\Delta P$ due to action a	$P(s\|a) - P(s\|a_0)$		action
(3)	$+\,\Delta P$ due to outcome realization	$1 - P(s\|a)$	$1-\kappa(s,a)$	brute luck
			$\kappa(s,a)$	action
(4)	$=\;\; P$ final (actually realized)	1		

Table 1: Probability accounting and attribution. The fractions $\kappa(s,a)$ and $1-\kappa(s,a)$ of outcome luck attributed to the manager's action and to brute luck, respectively, are given by Equations (6) and (7).

Definition 1 (ex-post causal responsibility)

The manager's *ex-post causal responsibility* for the actually realized state s of the system when taking a feasible action $a \in [0,1]$ is measured by:

$$R^c(s\,|\,a) \;:=\; P(s\,|\,a) - P(s\,|\,a_0) + \kappa(s,a)\,[1 - P(s\,|\,a)] \tag{10}$$

$$= \begin{cases} \dfrac{P(s\,|\,a) - P(s\,|\,a_0)}{P(s\,|\,a)} & \text{for } P(s\,|\,a) > P(s\,|\,a_0) \\ 0 & \text{for } P(s\,|\,a) \leq P(s\,|\,a_0) \end{cases}, \tag{11}$$

where the probabilistically neutral action a_0 is defined through Equation (2) and $\kappa(s,a)$ is given by Equation (6).

This expression answers the question: "If the manager takes action a and if the system eventually is actually in state s, to what extent is this outcome caused by the manager's action – rather than: by brute luck?" This is the manager's causal responsibility – in the sense of: ascribed causality – for the occurrence of state s. It is an ex-post measure as it takes the actual realization of state s as given and looks back to the chain of causality that has brought about this outcome.

It is obvious from the defining Equation (11) that the manager's ex-post causal responsibility for the system state depends on three constituents: the actually realized state s, which may be the initial state ($s = s_i$) or the other state ($s = s_o$), this state's initial probability prior to action, $P(s\,|\,a_0)$, and the action a that the manager has taken.

In general, $R^c(s \mid a)$ ranges between zero and one. It is zero for $a = a_0$. In this case, when the manager takes the probabilistically neutral action and, hence, does not influence at all the probability of the explanandum state s to occur, her causal responsibility for the actual occurrence of the explanandum state is zero. One says that she is "not causally responsible at all" for the outcome. The same goes for any action that decreases the probability of the explanandum state s from the initial probability, $P(s\,|\,a) < P(s\,|\,a_0)$. This is the case for an action $a < a_0$ when explaining a regime shift to the other state s_o (Figure 2 bottom left), or for an action $a > a_0$ when explaining the remaining in the initial state s_i (Figure 2 top right). This means, if the manager takes an action which decreases the probability of the explanandum state she is not causally responsible for this outcome at all; it is entirely due to brute luck – whatever action she took.

This result may seem surprising at first. (Why does it make no difference at all for measuring responsibility what action the manager takes? Why does acting not imply any causal responsibility?) Yet, it is indeed a straight-forward implication of Vallentyne's (2008) concept of "partial responsibility" employed here. In this concept, there are two probability shifts attributed to the manager's action: the direct probability shift of the explanandum state due to the manager's action, and a share of outcome luck which is in proportion to the contribution of action in establishing this outcome luck. If the manager's action makes the explanandum state *less* likely, the first of these two probability shifts is negative and the second one is positive and has exactly the same absolute amount. Hence, the two exactly cancel and the resulting causal responsibility is zero.

In contrast, the manager has a positive causal responsibility for the explanandum state s if her action a has increased the probability of the explanandum state from the initial probability, $P(s \mid a) > P(s \mid a_0)$. This is the case for an action $a > a_0$ when explaining a regime shift to the other state s_o (Figure 2 bottom right), or for an action $a < a_0$ when explaining the remaining in the initial state s_i (Figure 2 top left). This causal responsibility is given by $(P(s \mid a) - P(s \mid a_0))/P(s \mid a)$ (Equation 11). It is simply the direct increase in the explanandum state's probability due to the action, $P(s \mid a) - P(s \mid a_0) > 0$, normalized to the explanandum state's probability resulting from the action, $P(s \mid a)$. As the latter is smaller than one, the fraction is greater than the direct increase in the explanandum state's probability due to the action. The normalization factor results from measuring causal responsibility, again according to Vallentyne's (2008) concept of "partial responsibility", as a sum of two contributions – the direct probability shift of the explanandum state due to the action and that share of outcome luck which is in proportion to the action's contribution to establishing the outcome luck. The latter has the explanandum state's probability, $P(s \,|\, a)$, as a base and this carries over to the sum.

The maximal value of $R^c(s \,|\, a)$ can be as large as one, which means that the manager is "fully responsible" for causing the explanandum state s. In general, $R^c(s \mid a)$ will take on a value in between zero and one, and one says that the manager is "partially responsible" for causing the explanandum state s or, more exactly, that the occurrence of the explanandum state s "is to $R^c(s\,|\,a) \cdot 100\,\%$ caused by the manager's action a, and to $(1 - R^c(s\,|\,a)) \cdot 100\,\%$ by brute luck".

The measure $R^c(s\,|\,a)$ in the defining Equation (11) is generic for any given state s. It can be specified for the two potential states of the system as follows.

Proposition 1

If the manager takes a feasible action $a \in [0, 1]$ and there is actually a regime shift into the other state s_o, her ex-post causal responsibility for this regime shift is

$$R^c(s_o\,|\,a) = \begin{cases} 0 & \text{for } a \leq a_0 \\ \dfrac{p(a) - p_0}{p(a)} & \text{for } a > a_0 \end{cases} . \tag{12}$$

Conversely, if there is actually no regime shift and the system remains in the initial state s_i, her ex-post causal responsibility for the remaining in the initial state is

$$R^c(s_i\,|\,a) = \begin{cases} \dfrac{p_0 - p(a)}{1 - p(a)} & \text{for } a < a_0 \\ 0 & \text{for } a \geq a_0 \end{cases} . \tag{13}$$

Proof. In Equation (11), employ $P(s \mid a)$ (from Equation 5) and $p(a_0) = p_0$ (from Equation 1). □

The manager's ex-post causal responsibility for a regime shift, $R^c(s_o\,|\,a)$ (Equation 12, Figure 3 right), is zero for all actions $a < a_0$, which decrease the probability of a regime shift. From the probabilistically neutral action, $a = a_0$, it increases monotonically until it reaches its maximal value of $(p_{\max} - p_0)/p_{\max}$ for the strongest action, $a = 1$. In general, this maximal value can be as large as one, which means that the manager is "fully responsible" for causing the regime shift. This is only possible for an extreme parameter value $p_0 = 0$, which means that prior to any action the regime shift was impossible. But even the maximal value could be as small as zero. This is only possible for an extreme parameter value $p_{\max} = p_0$, which means that even the strongest action does not increase the probability of a regime shift beyond the initial probability.

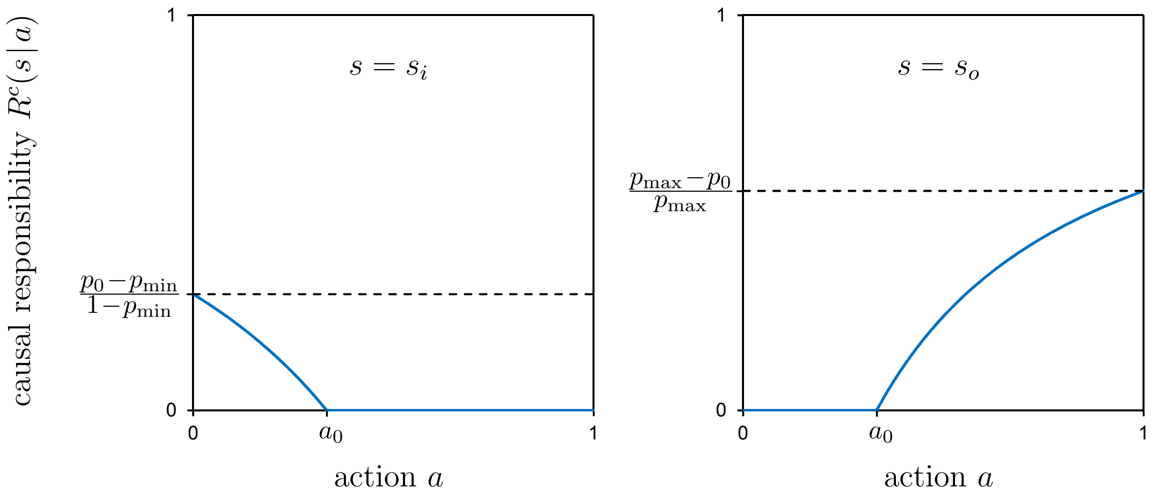

Figure 3: Ex-post causal responsibility for actually remaining in the initial state, $R^c(s_i\,|\,a)$ (left, Equation 13), and for an actual regime shift into the other state, $R^c(s_o \mid a)$ (right, Equation 12), as a function of action, according to the linear model, $p(a) = p_{\min} + (p_{\max} - p_{\min})\,a$ (Equation 3). Parameter values: $p_0 = 0.4$, $p_{\min} = 0.15$, $p_{\max} = 0.9$.

The manager's ex-post causal responsibility for remaining in the initial situation, $R^c(s_i\,|\,a)$ (Equation 13, Figure 3 left), assumes its maximal value of $(p_0 - p_{\min})/(1 - p_{\min})$ for the weakest action, $a = 0$. It decreases monotonically until it reaches the value of zero for the probabilistically neutral action, $a = a_0$. For all stronger actions $a > a_0$, which decrease the probability of remaining in the initial state, it remains zero. In general, the maximal value can be as large as one, which means that the manager is "fully responsible" for remaining in the initial situation. This is only possible for an extreme parameter value $p_0 = 1$, which means that prior to any action remaining in the initial situation was impossible $(1 - p_0 = 0)$. But even the maximal value could be as small as zero. This is only possible for an extreme parameter value $p_{\min} = p_0$, which means that even the weakest action does not increase the probability of remaining in the initial state beyond the initial probability.

4.1.2 Ex-ante expected causal responsibility

So far, I have discussed the manager's causal responsibility from an ex-post perspective. That is, I have taken the actual realization of state s of the system as given and measured the manager's causal responsibility for bringing about this outcome. But at the time of action, that is ex-ante to the resolution of uncertainty, the manager does not yet know for certain which one of the two potential system states will actually be realized. With probability $P(s_i \,|\, a) = 1 - p(a)$ it will be state s_i (no regime shift), in which case her ex-post causal responsibility for bringing about this outcome would be $R^c(s_i \mid a)$; and with probability $P(s_o \mid a) = p(a)$ it will be state s_o (regime shift), in which case her ex-post causal responsibility for bringing about this outcome would be $R^c(s_o \mid a)$. Hence, in this situation of uncertainty about the eventually resulting state, also the manager's causal responsibility for the outcome, $R^c(s \,|\, a)$, is still uncertain. Given the probability distribution over the two potential outcomes that results from her action, and given the ex-post causal responsibilities for each of the two outcomes, the manager faces an ex-ante expected causal responsibility for the uncertain state s of the system when taking action a.

Definition 2 (ex-ante expected causal responsibility)
The manager's *ex-ante expected causal responsibility* for the state s of the system when taking a feasible action $a \in [0, 1]$ is measured by:[11]

$$R^c(a) := \mathcal{E}_{p(a)} \left[R^c(s \,|\, a) \right] = P(s_i \,|\, a) R^c(s_i \,|\, a) + P(s_o \,|\, a) R^c(s_o \,|\, a) \ , \tag{14}$$

where $\mathcal{E}_{p(a)}$ denotes the expectation operator with the probability distribution that results from action a and the ex-post causal responsibility for state s, $R^c(s \,|\, a)$, is defined in Definition 1.

This expression answers the question: "If the manager takes action a, to what extent

[11]In slight abuse of notation, I denote this function also R^c, just like the ex-post causal responsibility. The difference is that the ex-ante expected causal responsibility does not depend on the state s. As I use these functions throughout the manuscript with an explicit specification of their arguments, there should not be any misunderstanding: $R^c(s, a)$ is the ex-post and $R^c(a)$ the ex-ante expected responsibility.

can she expect at the time of action – that is, prior to the resolution of uncertainty – to be causally responsible for the resulting state of the system?" In general, $R^c(a)$ ranges between zero and one. It is zero for the probabilistically neutral action $a = a_0$, and non-vanishing for all other actions $a \neq a_0$.

The manager's ex-ante expected causal responsibility when taking action a, $R^c(a)$, can be specified from the defining Equation (14) by using the specific probabilities for both potential system states, $P(s\,|\,a)$, and the specific expressions for the ex-post causal responsibility for each state, $R^c(s\,|\,a)$.

Proposition 2

If the manager takes a feasible action $a \in [0, 1]$ her ex-ante expected causal responsibility for the state of the system is

$$R^c(a) = |\; p(a) - p_0 \;| = \begin{cases} p_0 - p(a) & \text{for } a < a_0 \\ 0 & \text{for } a = a_0 \\ p(a) - p_0 & \text{for } a > a_0 \end{cases} . \tag{15}$$

Proof. In the defining Equation (14), employ $P(s\,|\,a)$ from Equation (5) and the specific expressions for $R^c(s\,|\,a)$ from Equations (12) and (13). □

This is the extent of direct and indirect causation that the manager can expect to have for the state of the system which is still uncertain at the time of action. It is based on two aspects: the ex-post causal responsibility for each of the two potential system states and the probabilities of each of them, given the manager's action. The measure depends on both the initial regime-shift probability prior to action, p_0, and the action a that the manager takes.

Equation (15) states that the manager's ex-ante expected causal responsibility when taking action a is simply the absolute amount of the direct change in the regime-shift probability from the action over the initial probability p_0 (Figure 4). It is zero if the manager takes the probabilistically neutral action, $a = a_0$. As a increases from a_0, a regime shift becomes more likely than initially and this linearly increases the manager's ex-ante expected causal responsibility, up to a maximum of $p_{\max} - p_0$ for the strongest action, $a = 1$. As a decreases from a_0, a regime shift becomes less likely than initially

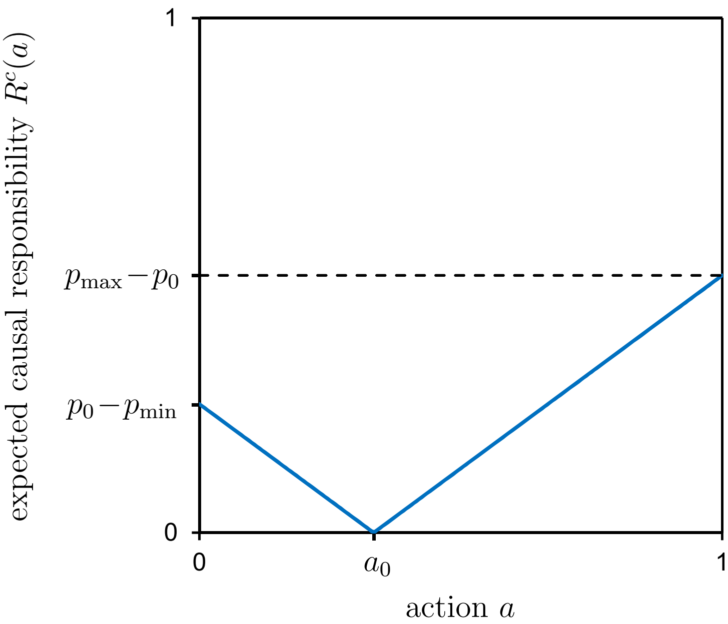

Figure 4: Ex-ante expected causal responsibility for the system state as a function of action, $R^c(a)$ (Equation 15), according to the linear model, $p(a) = p_{\min} + (p_{\max} - p_{\min})\, a$ (Equation 3). Parameter values: $p_0 = 0.4$, $p_{\min} = 0.15$, $p_{\max} = 0.9$.

and this also linearly increases the manager's ex-ante expected causal responsibility, up to a maximum of $p_0 - p_{\min}$ for the weakest action, $a = 0$. As $R^c(0) = p_0 - p_{\min}$ and $R^c(1) = p_{\max} - p_0$, and the former is larger (smaller) than the latter for $p_0 > (<)\ (p_{\max} - p_{\min})/2$, the maximum ex-ante expected causal responsibility R^c is reached at $a = 0$ $(a = 1)$ for $p_0 > (<)\ (p_{\max} - p_{\min})/2$. This means, if the prior regime shift probability p_0 is relatively large, the maximum ex-ante expected causal responsibility for the system state is reached for taking the weakest action, $a = 0$, which makes the remaining in the initial state most likely. And if the prior regime shift probability p_0 is relatively small, the maximum ex-ante expected causal responsibility for the system state is reached for taking the strongest action, $a = 1$, which makes a regime shift most likely. For extreme parameter values, either $p_0 = 1$ and $p_{\min} = 0$, or $p_0 = 0$ and $p_{\max} = 1$, this maximum ex-ante expected causal responsibility is one, which means that the manager, when taking the respective action, can expect to have full causal responsibility for the system state that actually occurs.

It is interesting to observe that, while the manager's ex-post causal responsibility depends on the direct probability shift due to action a in a non-linear manner (Equations 12 and 13, Figure 3), her ex-ante expected causal responsibility depends on the direct probability shift due to action a in a (piecewise) linear manner (Equation 15, Figure 4). Technically, the reason is in two properties of the ex-post causal responsibility: (1) The expressions for the ex-post responsibility for each explanandum state have in their denominator the probability of this very state, which is then used again as a weight for the respective ex-post responsibility when calculating the expected responsibility. So, for each state the probability of the state cancels out and it remains the direct probability shift due to action. (2) These expressions for the ex-post responsibility hold for the range of actions where the action increases the probability of the explanandum state. For those actions that decrease the probability of the explanandum state, the ex-post responsibility is zero. So, from both ex-post responsibilities – one for each explanadum state – only that part enters the expected value where the action increases the probability of the explanandum state. For the initial state, this is for $a < a_0$ and for the other state this is for $a > a_0$, so that the entire range of a is covered and for each value of a by exactly one ex-post responsibility. Conceptually, this neatly fitting interplay of several technical properties is due to Vallentyne's (2008) concept of partial ex-post responsibility (as detailed and formalized in Section 4.1.1).

To put this measure of ex-ante expected responsibility into perspective, compare it to the following simplest aggregate measure of how action a changes the probabilistic conditions of the system compared to the initial situation. This is the l^1-norm of $P(s \,|\, a) - P(s \,|\, a_0)$.

Definition 3

The *aggregate probabilistic impact* of a feasible action $a \in [0, 1]$ on the system is measured by the normalized aggregate absolute amount of probability shifts from the initial situation:

$$I(a) := \frac{1}{2} \sum_{s=s_i,s_o} \mid P(s \,|\, a) - P(s \,|\, a_0) \mid \quad . \tag{16}$$

$I(a)$, is calculated as the sum of the absolute amounts of the direct probability shifts of both potential system states due to this action. As a strong action $a > a_0$ will make the regime shift to the other state more likely and the remaining in the initial state less likely, one has $P(s_o \mid a) - P(s_o \mid a_0) = p(a) - p_0 > 0$ and $P(s_i \mid a) - P(s_i \mid a_0) = p_0 - p(a) < 0$. The two probability shifts have identical absolute amounts and opposite signs. Likewise, as a weak action $a > a_0$, will make the regime shift to the other state less likely and the remaining in the initial state more likely, one has $P(s_o \,|\, a) - P(s_o \,|\, a_0) = p(a) - p_0 < 0$ and $P(s_i \,|\, a) - P(s_i \,|\, a_0) = p_0 - p(a) > 0$. Again, the two probability shifts have identical absolute amounts and opposite signs. The measure $I(a)$ simple adds the absolute amounts of the direct probability shifts. Without taking the absolute amounts, summing the two would yield zero as the two just offset. The factor of 1/2 in front of the sum serves to normalize the maximum amount of $I(a)$ to one, as by summing over the absolute amounts one counts each probability shift twice.

The probabilistic impact of an action on the system, $I(a)$, ranges between zero and one. It is zero if and only if action a does not change the probability of any state of the system, $P(s \,|\, a) = P(s \,|\, a_0)$ for both $s \in \{s_i, s_o\}$. And it is one if and only if any state that has a vanishing initial probability occurs with certainty with the action a and, vice versa, any state that has an initial probability of one has a vanishing probability with the action a. Hence, the probabilistic impact of an action on the system is maximal if and only if it completely and certainly reverses the state of the system. In general, the larger $I(a)$ the larger the aggregate impact of action a on the probabilities of the potential system states.

The measure $I(a)$ (Equation 16) has been suggested by Krysiak (2008, 2011, 2012) and called "responsibility" (without any qualification, though), to differentiate between two kinds of uncertainty in a system – the uncertainty that exists in the system prior to, and independently of, present action; and the uncertainty in the system that is due to present action. The latter, according to Krysiak, is measured by $I(a)$.

Employing the specific probabilities $P(s \,|\, a)$ from Equation (5) in the defining Equation (16), one obtains

$$I(a) = |\ p(a) - p_0\ | = \begin{cases} p_0 - p(a) & \text{for } a < a_0 \\ 0 & \text{for } a = a_0 \\ p(a) - p_0 & \text{for } a > a_0 \end{cases} . \tag{17}$$

From Equation (17), it becomes apparent that there is a simple relationship between the aggregate probabilistic impact of an action on the system, and the ex-ante expected causal responsibility for the state of the system when taking this action, and also that $I(a)$ has the same properties as $R^c(a)$.[12]

Proposition 3

The ex-ante expected causal responsibility for the state of the system when taking action a is, for all feasible actions $a \in [0,1]$, equivalent to the aggregate probabilistic impact of action a on the system:

$$R^c(a) \equiv I(a) \quad \text{for all } a . \tag{18}$$

Proof. Compare Equations (15) and (17). □

Although simple, this is a non-tautological and non-trivial result. The two measures, $I(a)$ and $R^c(a)$, are defined in conceptually different and independent ways. Whereas the aggregate probabilistic impact $I(a)$ is calculated as a simple sum of the two direct

[12] $I(a)$ ranges between zero and one. It is zero for $a = a_0$. This means, taking the probabilistically neutral action does not change the probabilities of the two potential system states at all. As $I(a)$ increases linearly as a either decreases or increases from a_0, the maximum aggregate change of probabilities is obtained for either the weakest action, $a = 0$, or for the strongest action, $a = 1$. As $I(0) = p_0 - p_{\min}$ and $I(1) = p_{\max} - p_0$, and the former is larger (smaller) than the latter for $p_0 > (<)\ (p_{\max} - p_{\min})/2$, the maximum probabilistic impact I is reached at $a = 0$ $(a = 1)$ for $p_0 > (<)\ (p_{\max} - p_{\min})/2$. This means, if the prior regime shift probability p_0 is relatively large, the maximum aggregate probabilistic impact on the system is reached for taking the weakest action, $a = 0$, which makes the remaining in the initial state most likely. And if the prior regime shift probability p_0 is relatively small, the maximum aggregate probabilistic impact on the system is reached for taking the strongest action, $a = 1$, which makes a regime shift most likely. For extreme parameter values, either $p_0 = 1$ and $p_{\min} = 0$, or $p_0 = 0$ and $p_{\max} = 1$, this maximum impact is one, which means that the respective action completely and certainly reverses the state of the system.

probability shifts, taken as absolute amounts, the ex-ante expected causal responsibility $R^c(a)$ is calculated as an expectation value, using different probabilities as weights, over ex-post causal responsibilities, which include not only the direct probability shift but also the indirect ones (as a proportional share of outcome luck). The equivalence between the two holds because I employ Vallentyne's (2008) notion of causal responsibility and take it as an ex-post measure, which is then subject to an ex-ante expectation.[13] If one employed a different responsibility concept, that did not attribute a proportional part of outcome luck to the manager's action, there would be no equivalence in general between the manager's ex-ante expected causal responsibility and the action's aggregate probabilistic impact.

Proposition 3 means that one can indeed, as Krysiak does, think of the aggregate probabilistic impact of the manager's action a on the system as the manager's (ex-ante expected) causal responsibility for the state of the system. The equivalence between the two measures, $I(a)$ and $R^c(a)$, therefore gives additional plausibility to using Vallentyne's (2008) concept of partial causal responsibility under outcome risk. It shows that this notion is consistent with other plausible ways of thinking about responsibility.

4.2 Normative responsibility

In general, a normative responsibility is an imperative on the manager with regard to some normative state of the system. It puts an obligation on the manager to act accordingly.

Definition 4 (normative responsibility)
The manager's *normative* responsibility for the state of the system is to see to it that $s \overset{!}{=} s^n$, where $s^n \in \{s_i, s_o\}$ is given.

[13]Technically, the result comes about because in Vallentyne's notion of responsibility the indirect probability shift is considered as a fraction of outcome luck, where the probability $P(a \mid s)$ of the explanandum state s shows up in the denominator. When calculating the expectation value over ex-post responsibilities the very same probability $P(a \mid s)$ shows up as a weight for the ex-post responsibility, so that it cancels out, and the expectation value in R^c becomes a simple sum over probability changes, like in I.

For example, the manager's normative responsibility could simply be to keep the system in the initial state, $s^n = s_i$, which is a plausible norm if the system is initially in a good state and the other state was worse. Or the manager's normative responsibility could simply be to shift the system into the other state, $s^n = s_o$, which is a plausible norm if the system is initially in a bad state and it would be an improvement to shift it into a good state (as it is the case, e.g., with a collapsed fishery). More generally, the manager's normative responsibility could be to act such that the system is in a state that fulfills some more complex condition. For example, the manager's normative responsibility could be to act such that the system is in the state that maximizes expected net benefits, taking into account the potential benefits in each state, the (opportunity) costs of taking action, and how action affects the regime-shift probability. Even under such a more complex norm it is evident – albeit only indirectly and after some considerations – which one of the two potential states of the system is the normative one, $s^n \in \{s_i, s_o\}$

Given the norm that the manager ought to fulfill, a^n denotes the action that meets the normative imperative in the best possible way, that is, it maximizes the conditional probability of $s = s^n$:

$$a^n := \text{argmax } P(s^n \,|\, a) \; . \tag{19}$$

For example, if the manager's normative responsibility is to keep the system in its initial state, $s^n = s_i$, the action that best meets this imperative is (employing $P(s \,|\, a)$ from Equation 5 and Equation 1)

$$a^n = \text{argmax } P(s_i \,|\, a) = \text{argmin } p(a) = 0 \; . \tag{20}$$

In words, if the manager's normative responsibility is to keep the system in its initial state, the responsible action is the weakest feasible action. In contrast, if the manager's normative responsibility is to shift the system into the other state, $s^n = s_o$, the action that best meets this imperative is (again employing $P(s \,|\, a)$ from Equation 5 and Equation 1)

$$a^n = \text{argmax } P(s_o \,|\, a) = \text{argmax } p(a) = 1 \; . \tag{21}$$

In words, if the manager's normative responsibility is to shift the system into the other state, the responsible action is to act maximally. Either way, if $p_{\min} > 0$ and $p_{\max} < 1$

even taking the responsible action a^n will not make the intended normative system state s^n to actually occur for certain, but there is a non-vanishing probability of the unintended system state to actually occur.

4.3 Virtuous responsibility

We now know (from Section 4.2) what action the manager should take to best live up to her normative responsibility, $s \stackrel{!}{=} s^n$, namely action a^n. We also know (from Section 4.1) what is, ex post, the extent of causation of an eventual state s of the system if the manager takes some action a, namely $R^c(s \,|\, a)$. From these two building blocks we can develop a measure of the manager's virtuous responsibility, R^v, when taking action a. I suggest to use the following one.

Definition 5 (virtuous responsibility)
The manager's *virtuous responsibility* for meeting the norm to see to it that $s \stackrel{!}{=} s^n$ when taking a feasible action $a \in [0, 1]$ is measured by

$$R^v(a, s^n) := 1 - \frac{B\left(\mathcal{E}_{p(a)} \mid R^c\left(s \,|\, a^n\right) - R^c\left(s \,|\, a\right) \mid\right)}{B\left(\max\limits_a \left\{\mathcal{E}_{p(a)} \mid R^c\left(s \,|\, a^n\right) - R^c\left(s \,|\, a\right) \mid\right\}\right)} \,, \tag{22}$$

where $\mathcal{E}_{p(a)}$ is the expectation operator with the probability distribution that results from action a, $R^c\left(s \,|\, a\right)$ is given by Equation (11), a^n is given by Equation (19) from s^n, and $B : [0, 1] \to \mathbb{R}_+$ may be any function with $B(0) = 0$, $B'(\cdot) > 0$ and $B''(\cdot) \geq 0$.

This expression answers the question: "If the manager has the normative responsibility to see to it that $s \stackrel{!}{=} s^n$ and takes action a, to what extent is she living up to fulfilling the norm?" Essentially, expression (22) compares how much the action actually taken, a, deviates from the one that should have been taken to fulfill the norm, a^n. The larger this deviation, the less does the manager live up to her normative responsibility. Taking one minus the measure of deviation is in line with measuring responsibility by adding (or subtracting) probabilities, as I already did for measuring causal responsibility in Section 4.1.

Precisely, the measurement of the gap in responsibility according to expression (22) is as follows. It does not simply use the difference between the actual action a and the responsible action a^n, but the difference between the ex-post causal responsibilities associated with these two actions, $R^c\,(s\,|\,a^n) - R^c\,(s\,|\,a)$. This is the difference in causal effect on the system that the actual action a makes compared to the responsible action a^n. Measuring the deviation of a particular action from a norm in terms of causal effect is in line with consequentialist ethics according to which one must judge a person's actions based on the consequences of these actions. Expression (22) takes the absolute amount of this responsibility gap because it makes no difference for the judgement in what way the manager deviates from the responsible action: doing too little is as wrong as doing too much if the norm implies a unique responsible action. At the time of action, the outcome is still uncertain. That is, the manager does not yet know for certain which state will actually occur, and by how much the ex-post causal effect of any particular action will deviate from that of the responsible action. The best that she can do ex-ante is to form an expectation of the outcome based on the probability distribution that her action implies. Therefore, also the judgment of her action must be based on the state of (probabilistic) knowledge at the time of action and, hence, on the ex-ante expected causal effects of the actual and the responsible action. For this reason, expression (22) builds the expectation value over the (absolute amount) of the ex-post causal responsibility gap. The ex-ante expected gap in causal responsibility thus calculated, $\mathcal{E}_{p(a)} \mid R^c\,(s\,|\,a^n) - R^c\,(s\,|\,a) \mid$, is the core element of the measure of virtuous responsibility in Equation (22).

The function B is introduced to allow for different ways of how to take into account such an expected gap in causal responsibility when measuring virtuous responsibility. One can think of B as a *blame function* that specifies by how much an increase in the ex-ante expected gap in causal responsibility decreases the manager's virtuous responsibility. Specifying this function is a normative issue, as it conveys a value statement about how wrong a given ex-ante expected gap in causal responsibility is considered to be. I refrain from making such a normative choice here, but leave it open what blame function to use. Formally, the simplest possibility is that B is the linear function $B(x) = x$. In this case, each marginal increase in the expected-causal-responsibility-gap decreases the measure of virtuous responsibility by the same absolute amount. In general, it seems plausible to assume that B is a convex function (although an ethics in which B is concave is imaginable, too). This means that the marginal effect of the expected-causal-responsibility-gap on the measure of virtuous responsibility is increasing with the extent of the gap: a larger deviation from the norm has a larger relative effect than a smaller deviation. A simple example of a strictly convex function would be the quadratic function $B(x) = x^2$.

Taking in the denominator the maximal value of the nominator, i.e. the blame for the maximal possible gap in ex-ante expected causal responsibility, serves to normalize the fraction to one. Obviously, if the manager takes the responsible action, $a = a^n$, there is no gap in causal responsibility and the fraction becomes zero. Then, $R^v(a^n, s^n) = 1$, which means that the virtuous responsibility is 100 %. One says, "the manager fully lives up to her normative responsibility", or "the manager acts fully responsibly", with respect to the given norm. In contrast, if the actual action a deviates from the responsible action a^n in the maximal possible way, the fraction becomes one. Then, the virtuous responsibility $R^v(a, s^n)$ is zero. One says, "the manager does not live up to her normative responsibility at all", or "the manager does not act responsibly at all", with respect to the given norm. In general, R^v ranges between zero and one by construction.

The measure of the manager's virtuous responsibility, $R^v(a, s^n)$ (Equation 22), depends on the normative system state, s^n, that is, the state that the manager ought to attain under the given norm, and the action a that she actually takes. It does *not* depend on the actually realized state s of the system. This means, which state of the system actually occurs, s_i or s_o, is irrelevant for assessing to what extent the manager has lived up to fulfilling her normative responsibility. The rationale behind this argument is that the actual realization of a system state is fundamentally a matter of chance, and the manager can only be praised or blamed for what she did, but not for the arbitrary influence of chance. Therefore, to assess to what extent the manager lives up to fulfilling her normative responsibility to attain a given system state, one must focus on the causal influence on the system that she actually exerted to make the normative state more likely, and compare this to the causal influence that she ought to have exerted to best meet the given normative imperative.

The measure of the manager's virtuous responsibility, $R^v(a, s^n)$ (Equation 22), also depends on the limits of how the feasible actions a can affect the probabilities of the potential system states, $P(s \,|\, a)$, which enter the expression through the max-function in the denominator. For judging the gap in causal responsibility, measure (22) takes the maximal extent of the gap as a reference. In the linear model of Equation (3), this is determined by the two parameters $p_{\max}$ and $p_{\max}$ which parameterize the minimal and the maximal regime-shift probability that is attainable through any feasible action $a \in [0, 1]$.

The measure of normative responsibility, $R^v(a, s^n)$, in the defining Equation (22) is generic for any given normative state s^n. It can be specified for the two potential normative system states, $s^n = s_i$ (no regime shift) or $s^n = s_o$ (regime shift), by using the specific expressions for the ex-post causal responsibility for each state, $R^c(s \,|\, a)$, and by using a particular blame function F.

Proposition 4

(1) For a linear blame function $B(x) = x$ the following holds:
If the manager has a normative responsibility to see to it that $s \stackrel{!}{=} s_i$, that is, to keep the system in the initial state, and takes a feasible action $a \in [0, 1]$, her virtuous responsibility for meeting the norm is

$$R^v(a, s_i) = \frac{p_{\max} - p(a)}{p_{\max} - p_{\min}} \ . \tag{23}$$

Conversely, if the manager has a normative responsibility to see to it that $s \stackrel{!}{=} s_o$, that is, to shift the system into the other state, and takes feasible action $a \in [0, 1]$, her virtuous responsibility for meeting the norm is

$$R^v(a, s_o) = \frac{p(a) - p_{\min}}{p_{\max} - p_{\min}} \ . \tag{24}$$

(2) For a quadratic blame function $B(x) = x^2$ the following holds:
If the manager has a normative responsibility to see to it that $s \stackrel{!}{=} s_i$, that is, to keep the system in the initial state, and takes feasible action $a \in [0, 1]$, her virtuous responsibility for meeting the norm is

$$R^v(a, s_i) = 1 - \left[\frac{p(a) - p_{\min}}{p_{\max} - p_{\min}}\right]^2 \ . \tag{25}$$

Conversely, if the manager has a normative responsibility to see to it that $s \overset{!}{=} s_o$, that is, to shift the system into the other state, and takes feasible action $a \in [0, 1]$, her virtuous responsibility for meeting the norm is

$$R^v(a, s_o) = 1 - \left[\frac{p_{\max} - p(a)}{p_{\max} - p_{\min}}\right]^2 . \tag{26}$$

Proof. In Equation (22), employ $P(s \mid a)$ (from Equation 5), $R^c(s \mid a)$ (from Equations 12 and 13), $a^n = 0$ for $s^n = s_i$ (from Equation 20) and $a^n = 1$ for $s^n = s_o$ (from Equation 21), as well as $p(0) = p_{\min}$ and $p(1) = p_{\max}$ (from Equation 1). □

Figure 5 illustrates this result. If the norm is to keep the system in the inital state

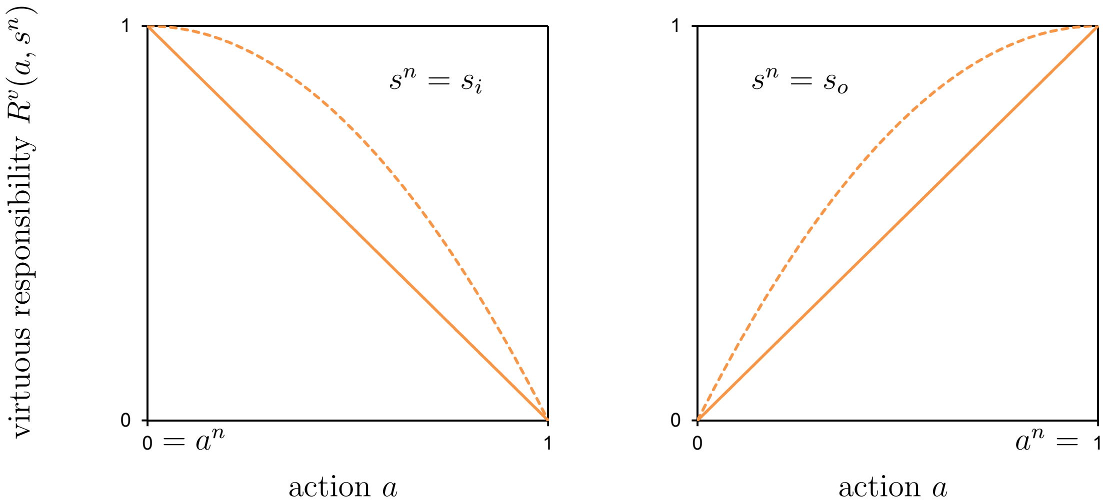

Figure 5: Virtuous responsibility as a function of action, $R^c(a, s^n)$ (Equations 23–26), for the norm to keep the system in the initial state, $s^n = s_i$ (left), and for the norm to shift the system into the other state, $s^n = s_o$ (right), according to the linear model, $p(a) = p_{\min} + (p_{\max} - p_{\min})\,a$ (Equation 3), and for the linear blame function $B(x) = x$ (solid curves) as well as for the quadratic blame function $B(x) = x^2$ (dashed curves). Parameter values: $p_{\min} = 0.15$, $p_{\max} = 0.9$.

(Figure 5 left), the responsible action is the weakest feasible action, $a = 0$, as this action maximizes the probability of remaining in the initial state. If the manager takes this action, her virtuous responsibility is $R^v(0, s_i) = 1$. That is, when taking the weakest feasible action she fully lives up to fulfilling the norm. In other words, she acts fully responsible with respect to the given norm. As a increases, her virtuous responsibility $R^v(a, s_i)$ decreases monotonically up to the strongest feasible action, $a = 1$, where the probability of remaining in the initial state becomes zero and also the manager's virtuous responsibility becomes zero, $R^v(1, s_i) = 0$. That is, when taking the strongest feasible action the manager does not live up at all to fulfilling the norm. In other words, she acts fully irresponsible with respect to the given norm. With a linear blame function, the decrease in virtuous responsibility when a increases is linear, too. Each unit of increase in action a leads to the same absolute amount of decrease in virtuous responsibility. With a quadratic blame function, the decrease in virtuous responsibility when a increases is concave: the steepness increases with a. Each additional unit of increase in action a leads to a larger absolute amount of decrease in virtuous responsibility.

If the norm is to shift the system into the other state, the situation is exactly reversed (Figure 5 right).

5 Discussion

6 Conclusions

Measuring the degree of causation that a manager has for a potential regime shift by taking a given management action ("causal responsibility", Section 4.1) is a crucial and important first step in assessing the quality of the management action. For, in a stochastic ecological-economic system the outcome – whether a regime shift occurs or not – is also influenced by pure chance. Roughly speaking, in a stochastic system everything can happen with some probability. Therefore, the actual outcome is not a good indicator of the quality of the management action. To assess the quality of management, the influence of chance and the influence of the management action on the outcome must be separated, and the measure developed in Section 4.1 achieves this.

With such a measure of causal responsibility, and taking into account the given norm that the manager ought to fulfill, one can then, in a next step, measure to what extent a manager has acted responsibly with respect to the given norm ("virtuous responsibility", Section 4.3). One should praise or blame the manager for her action depending on the extent of her having acted responsibly – rather than based on the actual outcome which may be due to pure chance. In economic terms, one should reward or punish the manager depending on the extent of her virtuous responsibility. More generally, one should construct institutions that react in a differentiated manner to a manager's action depending on her virtuous responsibility.

In ethics, there is a position that it is *fair* to attribute reward or punishment in proportion to virtuous responsibility. In contrast, economic analysis of liability for damaging a common-pool resource has revealed that it is, in general, *inefficient* if reward or punishment is attributed in strict proportion to the actor's virtuous responsibility, because it may create incentives for free-riding. More research is needed to study how to design institutions that enable and encourage responsible management of ecological-economic systems with potential regime shifts.

Acknowledgments

I am grateful to Frank Krysiak and Martin Quaas for discussion and comments, and to the German Federal Ministry of Education and Research (BMBF) for financial support under grant no. 01LC1826C.

References

Arrow, K.J., B. Bolin, R. Costanza, P. Dasgupta, C. Folke, C.S. Holling, B.-O. Jansson, S. Levin, K.-G. Mäler, C. Perrings and D. Pimentel (1995), Economic growth, carrying capacity, and the environment, *Science* 268(5210): 520–521.

Baumgärtner, S., T. Petersen and J. Schiller (2018), The concept of responsibility: Norms, actions, and their consequences, *SSRN Working Paper*, doi:10.2139/ssrn3157667, available from https://www.ssrn.com/abstract=3157667.

Braham, M. and M. van Hees (2009), Degrees of causation, *Erkenntnis* 71: 323–344.

Braham, M. and M. van Hees (2012), An anatomy of moral responsibility, *Mind* 121(483): 601–634.

Bovens, M. (1998). *The Quest for Responsibility: Accountability and Citizenship in Complex Organizations.* Cambridge: Cambridge University Press.

Folke, C., S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson and C.S. Holling (2004), Regime shifts, resilience, and biodiversity in ecosystem management, *Annual Review in Ecology, Evolution and Systematics* 35: 557–581.

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

Goodin, R.E. (1986), Responsibilities, *The Philosophical Quarterly* 36(142): 50–56.

Gunderson, L.H. and L. Pritchard Jr. (eds) (2002), *Resilience and the Behavior of Large-Scale Systems*, Island Press, Washington DC.

Hart, H.L.A. (1968), *Punishment and Responsibility: Essays in the Philosophy of Law*, New York.

Holling, C.S. (1973), Resilience and stability of ecological systems, *Annual Review of Ecology and Systematics*, 4: 1–23.

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

Kant, I. (1999)[1781], *Critique of Pure Reason*, in: The Cambridge Edition of the Works of Immanuel Kant, Cambridge: Cambridge University Press.

Kant, I. (1998)[1793], *Religion Within the Boundaries of Mere Reason*, in: The Cambridge Edition of the Works of Immanuel Kant, Cambridge: Cambridge University Press.

Krysiak, F.C. (2008), Sustainability, risk, and responsibility, Contribution to Workshop *Conceptualizing Sustainability under Uncertainty*, Camp Reinsehlen, 2008.

Krysiak, F.C. (2011), Nachhaltigkeit, Risiko und Diskontierung, in: M. Held, G. Kubon-Gilke, R. Sturn (eds), *Normative und institutionelle Grundfragen der Ökonomik, Jahrbuch 9 – Institutionen ökologischer Nachhaltigkeit*, Metropolis Verlag, Marburg, pp. 133–156.

Krysiak, F.C. (2012), Responsibility and learning on a slippery slope, Contribution to Workshop *Responsibility for Sustainability*, Camp Reinsehlen, 2012.

Levin, S.A. (1999), *Fragile Dominion: Complexity and the Commons*, Perseus Books, Reading, MA.

Levin, S.A., S. Barrett, S. Aniyar, W. Baumol, C. Bliss, B. Bolin, P. Dasgupta, P. Ehrlich, C. Folke, I.-M. Gren, C.S. Holling, A.-M. Jansson, B.-O. Jansson, D. Martin, K.-G. Mäler, C. Perrings and E. Sheshinsky (1998). Resilience in natural and socioeconomic systems, *Environment and Development Economics*, 3(2): 222–234.

Muradian, R. (2001), Ecological thresholds: a survey, *Ecological Economics* 38: 7–24.

Scheffer, M., S. Carpenter, J.A. Foley, C. Folke and B. Walker (2001), Catastrophic shifts in ecosystems, *Nature* 413: 591–596.

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

Walker, B.H. and J.A. Meyers (2004), Thresholds in ecological and social–ecological systems: a developing database, *Ecology and Society* 9(2): 3, available from http://www.ecologyandsociety.org/vol9/iss2/art3.

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