Preference for Verifiability*

Rommeswinkel, Hendrik[1,2]

[1]Graduate School of Economics, Hitotsubashi University
2-chōme-1 Naka, Kunitachi City, Tokyo 186-0004, Japan

[2]Waseda Institute for Advanced Study, Waseda University
1-chōme-104 Totsukamachi, Shinjuku City, Tokyo 169-8050, Japan

April 2024

*I thank presentation participants at the Hitotsubashi Summer Institute, at the Normative Economics and Public Policy Seminar, at the Karlsruhe Institute of Technology, at the Zukunftskolleg, University of Konstanz, at National Taiwan University, and at Kobe University for helpful comments.

Abstract

Decision makers may face situations in which they cannot observe the consequences that result from their actions. In such decisions, motivations other than the expected utility of consequences may play a role. The present paper axiomatically characterizes a decision model in which the decision maker cares about whether it can be ex post verified that a good consequence has been achieved. Preferences over acts uniquely characterize a set of events that the decision maker expects to be able to verify in case they occur. The decision maker chooses the act that maximizes the expected utility across verifiable events of the worst possible consequence that may have occurred.

For example, a firm choosing between different carbon emission reduction technologies may find some technologies to leave ex post more uncertainty about the level of emission reduction than other technologies. The firm may care about proving to its stakeholders that a certain amount of carbon reduction has been achieved and may employ privately obtained evidence to do so. It may choose in expectation less efficient technologies if the achieved carbon reduction is better verifiable using the expected future evidence.
KEYWORDS: Verifiability, uncertainty, maxmin utility, greenwashing, principal-agent
JEL CLASSIFICATION:

1 Introduction

In real-life decision making, we frequently make choices in which the final consequences we care about are only partially observable to us. For example, when deciding about different methods of CO2 compensation or emission reduction, the crucial consequence, the amount of CO2 emitted to the atmosphere, is not directly observed by most decision makers. Thus, the consequence the decision maker cares about remains more or less uncertain even after making a choice. In this case, it is natural if decision makers perceive a tradeoff between achieving good consequences in expectation and being able to verify that good consequences have been obtained.

This paper provides a first attempt at a decision theory with potentially unobservable consequences. Most models in decision theory implicitly assume that after making a decision, decision makers can observe the states and consequences of their actions. In decisions under uncertainty, commonly only ex ante uncertainty about consequences is permitted; after an act is chosen, there is no ex post uncertainty about the resulting consequence. While we employ the standard primitive of decision theory, preferences over acts, we dispose of the implicit assumption that there is no ex post uncertainty. Removing this assumption and assuming decision makers care about ex post verifiability makes deviations from expected utility maximization plausible that can be naturally understood in the language of models from the ambiguity literature (e.g., Chandrasekher et al., 2022; Gilboa & Schmeidler, 1989; Schmeidler, 1989).

Despite only requiring preferences over acts, our model yields a rich interaction between an (unobserved) stakeholder and the decision maker. From the behavior of the decision maker alone we can infer that the decision maker is in a principal-agent relationship with the stakeholder: First, the decision maker chooses an act that maps states into consequences. Second, either the stakeholder or the decision maker receive various proofs about which state of the world may have obtained. A proof comes in the form of an event; a proof of a particular event is only obtained if the true state is indeed within the event. Third, whoever received the proofs chooses from the available proofs and confronts the other party with this evidence. Lastly, the one confronted with the evidence points to a state consistent with the proven event (and that is in their interest) and claims that this consequence is likely to have realized.

More concretely, consider the following example from corporate carbon responsibility (CCR). A firm chooses between different carbon emission offsetting programs. The efficacy of the different offsetting programs +and thus the reduction of CO_2 in the atmosphere is uncertain. Therefore, a carbon emission offsetting program is an act that maps uncertain states of the world (efficacy of methods, costs, etc.) into consequences (the amount of CO_2 emitted to the atmosphere). Crucially, the exact state of the world (and thus the consequence) may remain unknown after the choice is made. The absence of objective probabilities and presence of ex post uncertainty may lead to greenwashing behavior: in their communication with stakeholders, firms may exaggerate the likelihood of states of the world in which their chosen action led to a large reduction in CO_2. Alternatively, pressure from stakeholders may lead other firms to choose overly cautious CO_2 offsetting programs because they want to be able to unambiguously prove that they have offset a certain amount of CO_2. Both behaviors may lead to deviations from subjective expected utility maximization (and thus, welfare losses). Both behaviors also depend on the degree of ex post uncertainty; if additional information will become available pertinent to some actions but not others, firms may deliberately choose acts for which the information is more or less relevant. Firms potentially choose ex-ante inferior programs depending on if for this program ex post more information will be available about the achieved emission reduction. The present paper axiomatically characterizes such behavior.

The decision model has the functional form of certain ambiguity aversion models. The decision maker is an expected utility maximizer over events that they expect to be able to verify. We call these events *verifiable events* and the set of verifiable events is subjective. For all other events, the decision maker exhibits extreme ambiguity aversion and evaluates such events by the worst possible consequence that may have resulted. The functional form is similar to the dual-self model of (Chandrasekher et al., 2022) except that utility is ambiguity neutral across verifiable events and the minimization step allows for all possible priors given a verifiable event. The present paper therefore provides a conceptually novel motivation for ambiguity aversion decision models: imperfectly observable consequences.

From the decision theoretic results we can derive policy implications in the context of CCR. Policy makers increasingly rely on CCR while focusing on regulating unverified environmental claims: for example, European Parliament (2023) plan to "proscribe [...] generic environmental claims [...] without proof" and "claims based on emissions offsetting schemes." The comparative statics and welfare analysis of our models show that both greenwashing behavior as well as verification seeking behavior may lead to welfare losses and these losses are nonmonotone in the degree of lack of verifiability. In the principal agent relationshipe of CCR, increased transparency requirements only unambiguously improve welfare if full ex post verifiability is induced. If some ex post uncertainty remains, increased transparency may lead to a lower welfare, depending on the set of available actions of firms. This puts some doubt on whether CCR combined with transparency requirements can suitably address environmental policy issues.

The paper proceeds as follows: Section 2 introduces the example application of carbon offsetting programs in more detail. Section 3 introduces the notation before Section 4 introduces the two decision models. The axioms that characterize these models are introduced in Section 5. Comparative statics for the models can be found in Section 6. The relation to the literature is discussed in Section 7.

2 Introductory Example

This section introduces an example of how the model presented in this paper can capture certain stylized facts about how firms interact with their stakeholders in corporate social responsibility issues where consequences are imperfectly observable. In such settings, subjective beliefs (for example about the efficacy of the firm's actions) may greatly differ between stakeholders and firms and there might not exist an objective truth to relate to.

Suppose a firm decides between different carbon offsetting programs. One option is to purchase certificates of nature-based carbon capture, or more poignant, to plant Trees. Another option is to purchase renewable energy certificates, RECs, and the third option is to improve the production processes to increase Efficiency. For simplicity, we assume that the total budget to be spent on the offsetting programs is fixed.

Suppose the set of consequences $\mathcal{X} = \mathbb{R}_+$ is the amount of CO2 emissions avoided in megatons. Let the set of states $\mathcal{S} = \{s, t, u\}$ consist of three mutually exclusive states. In state u, there is a low supply of emission-reduction certificates (both RECs and nature-based carbon removal). This makes purchasing such certificates costly and the budget is only sufficient to offset a small amount of CO2 emissions. In states s and t, the supply of such certificates is sufficient to offset a large amount. In state s, purchasing RECs leads consumers in the energy market to substitute from consuming a mix of sustainably and unsustainably produced electricity to a higher share of unsustainably produced energy. In state t, no such shift in behavior occurs. The options are shown in Figure 1.

Act	s	t	u
Trees	70	70	10
RECs	60	100	10
Efficiency	40	40	40

Figure 1: Carbon Reduction Programs

Which of the three programs should a firm choose? According to expected utility, this would depend on the risk preferences and the subjective beliefs of the decision maker. A sufficiently high subjective probability of s, t, or u can induce any of the three acts Trees, RECs, or Efficiency, respectively, to be optimal.

Suppose for a moment that there is no way of observing (or proving) the exact amount of CO2 reduction achieved by the different acts because only the total amount of emissions from many polluters can be measured in the atmosphere. In the absence of objective probabilities over states and when consequences are completely unverifiable, we can think of two stereotypical behaviors of firms when interacting with their stakeholders about carbon emissions:

A greenwashing firm may choose to purchase RECs and may exaggerate in their communication to stakeholders the likelihood that state t obtains. For example, Institute (2023) claim that for most firms joining the U.N. race to zero campaign (UNFCCC, 2023), the CO2 emission reductions are not sufficiently verified. Firms may choose to purchase RECs even if they subjectively believe state t to be very unlikely. Such behavior can be detected by an analyst from preferences that treat uncertain acts indifferent to their best possible consequence.

A verification seeking firm may choose Efficiency to eliminate any ex-post doubt of their stakeholders that they have achieved a certain amount of CO2 emission reductions. For example, a firm may try to prove to their stakeholders that their products are carbon neutral. Such firms may choose Efficiency even if they believe state u to be very unlikely because ex post they cannot exclude this state. Such behavior can be detected from preferences that treat uncertain acts indifferent to their worst possible consequence.

In the following, we provide two models –and axiomatic characterizations thereof– that correspond to these two types of behavior. Compared to the extreme cases discussed above, allow for some events, for example $\{s, t\}$, to be ex-post verifiable and subjectively determined. From an applied perspective, the main contribution is to provide conditions on preferences over emission reduction methods from which greenwashing behavior can be distinguished from verification-seeking behavior. In addition, for any firm following either preference, we can identify what information they expect to become available in the future.

3 Notation

Let $\mathcal{X}$ be a set of *consequences* and $\mathcal{S}$ a finite set of *states of the world*. $\mathcal{E} \subseteq 2^{\mathcal{S}}$ denotes the sigma algebra of *events*. A *simple act* is a measurable function $a : \mathcal{S} \to \mathcal{X}$ with a finite image. If E is an event and a, b are acts, then $a_E b$ is the act that agrees with a on all states $s \in E$ and that agrees with b on all states $s \in \overline{E} \equiv \mathcal{S} - E$. If $x \in X$, then x also denotes the constant act that implements x in all states.

We analyze preferences over acts. A *preference relation* is a binary relation $\succsim$ on $\mathcal{A}$. A function $U : \mathcal{A} \to \mathbb{R}$ *represents* $\succsim$ if $U(a) \geq U(b) \Leftrightarrow a \succsim b$. A representation U is *monotonic* if $U(a) \geq U(b)$ whenever for all $s \in \mathcal{S}$, $a(s) \succsim b(s)$. The *certainty equivalent* of an act a, denoted $[a] \in \mathcal{X}$, is a consequence such that $[a] \sim a$.

An event E is said to be *null* if $\beta \sim \gamma E \beta$ for all $\gamma, \beta \in X$ such that $\gamma \succ \beta$. An event E is said to be *universal* if $\gamma \sim \gamma E \beta$ for all $\gamma, \beta \in X$ such that $\gamma \succ \beta$. An event E is said to be *essential* if $\gamma \succ \gamma E \beta \succ \beta$ for some $\gamma, \beta \in \mathcal{X}$.

Unlike in expected utility, in our axiomatization what happens on null events may in principle still be preference relevant. We therefore introduce the stronger notion of irrelevant events. An event E is said to be *irrelevant* if $\gamma_E a \sim a$ for all $\gamma \in \mathcal{X}$ and $a \in \mathcal{A}$. A state of the world is said to be irrelevant if it is an element of an irrelevant event. The set of events that are not irrelevant is denoted $\mathcal{E}^*$ and the set of states that are not irrelevant is denoted $\mathcal{S}^*$. An event that is null need not be irrelevant. In our model, only if an event is irrelevant, then it never matters for preferences what consequences are acquired on this event.

If $\mathcal{F}$ is a set of events then $f : \mathcal{F} \to \mathbb{R}$ is called a *set function*. A set function is *grounded* if $f(\emptyset) = 0$. A set function is *normalized* if $f(\mathcal{S}) = 1$. A set function that is both normalized and grounded is called a *capacity*. A set function is *additive* if for all $E, F \in \mathcal{F}$ such that $E \cap F = \emptyset$, $f(E) + f(F) = f(E \cup F)$. A *probability measure* is a grounded, normalized, additive set function. A set function is called *supermodular (submodular)* on a set of events $\mathcal{G}$ if for all $E, F \in \mathcal{G}$, if $\eta(E) + \eta(F) \leq (\geq) \eta(E \cup F) + \eta(E \cap F)$. A set function is *modular* if it is both supermodular and submodular.

4 Decision Models

We analyze two decision models. In the first decision model, the decision maker has the desire to verify that a particular utility level has been reached. In the context of our example, this may arise because the decision maker wants to prove to a stakeholder that a certain benefit of taking the action has been achieved.

Definition 1 (Expected Verification Utility)**.** A preference relation $\succsim$ on $\mathcal{A}$ is an expected verification utility if there exists a set of verifiable events $\mathcal{V} \subseteq \mathcal{E}$, closed under intersections and containing $\mathcal{S}^*$, a probability measure $\mu : \mathcal{E} \to [0, 1]$, and a utility function $u : \mathcal{X} \to \mathbb{R}$ such that

$$U(a) = \int_{s \in \mathcal{S}^*} \max_{E \in \mathcal{V} : s \in E} \min_{t \in E} u(a(t)) d\mu \tag{1}$$

represents $\succsim$.

$\mathcal{V}$ is called the set of *verifiable events*. If $\mathcal{V} = \mathcal{E}$, then every state is verifiable and the decision maker maximizes expected utility. The verifiable events do not necessarily form a partition but are closed under intersections. It follows that the set of verifiable events is a π-system that contains $\mathcal{S}^*$. The interpretation of a verifiable event E is that if the event obtains (i.e., if the true state of the world is within E), then the decision maker receives a proof that they can use to prove that this event obtained. They then use this proof to prove to the stakeholder that at least the consequence $\min_{t \in E} u(a(t))$ has been achieved.[1]

[1]The model assumes the verifiable events to be act-independent. The act-dependent case is not treated in the current paper for two reasons: first, if such verifiable events would be objectively given for every act, then this would make the analysis almost trivial. Second, if

The decision model is a special case of an ambiguity aversion model within the intersection of Choquet expected utility and maxmin expected utility. However, the motivation for deviating from expected utility is not ambiguity aversion but the lack of ex-post verifiability of events: across verifiable events, the decision maker follows expected utility but whenever consequences arise on events that are not verifiable, the decision maker is extremely pessimistic.

Example. Firm A would like to be able to prove to its stakeholders that a high level of reduction of CO2 emissions has been achieved. The firm believes that state s and u are unlikely, though not impossible. In case the event $\{s, t\}$ obtains, the firm believes it is able to prove this to its stakeholders. For example, after purchasing any amount of certificates of any kind, the firm can show these certificates to its stakeholders. However, the firm is unable to prove that $\{u\}$ obtains if certificate prices are not publicly observable – it cannot use a low number of purchased certificates to prove that it is impossible to buy more certificates. The firm is also unable to prove that $\{s\}$ or that $\{t\}$ obtains since proving this would require observing the hypothetical behavior of how other market participants would behave given a purchase of RECs or no purchase of RECs. The firm can however verify that $\{s, t, u\}$ obtains since there is no other nonnull state. Firm A therefore considers two events when making a decision, $\{s, t\}$ and $\{s, t, u\}$. With probability $\mu(\{u\})$ it can only prove that $\{s, t, u\}$ obtains and that the emission reduction is at least equal to the worst consequence of an act. With probability $\mu(\{s, t\})$ the firm can verify that $\{s, t\}$ obtains and that the emission reduction is at least equal to the worst possible consequence on states s and t. The firm therefore multiplies the probability of u with the utility of the worst possible consequence on $\{s, t, u\}$ and the probability of $\{s, t\}$ with the worst possible consequence on $\{s, t\}$. If the probability of state u is sufficiently low, the firm will choose TREES. If u is sufficiently likely, the firm will choose EFFICIENCY. □

In the second decision model, the decision maker has the desire to obfuscate that a bad consequence might have resulted. In the context of our example, this may arise because a stakeholder may confront the firm with evidence about the state of the world (e.g., that $\{s\}$ obtains after having chosen RECs).

Definition 2 (Expected Obfuscation Utility)**.** A preference relation $\succsim$ on $\mathcal{A}$ is an expected obfuscation utility if there exists a set of verifiable events $\mathcal{V}$, closed

there are subjectively determined act-dependent verifiable events, then almost any behavior is possible and the verifiable events are unlikely to have interesting uniqueness properties.

under intersections, a unique probability measure $\mu : \mathcal{E} \to [0,1]$, and a unique utility function $u : \mathcal{X} \to \mathbb{R}$ such that

$$U(a) = \int_{s \in \mathcal{S}^*} \min_{E \in \mathcal{V}: s \in E} \max_{t \in E} u(a(t)) d\mu \tag{2}$$

represents $\succsim$.

The interpretation of this model is that after a stakeholder confronts the decision maker with the proof that an event E obtains, the decision maker points to the best possible consequence that might have been achieved on E. We assume that the stakeholder has adversarial preferences to the decision maker, i.e., the stakeholder tries to prove that at most a certain amount of utility has been obtained. The stakeholder always uses all available verifiable information (chooses the smallest event in $\mathcal{V}$ that contains the true state) and the decision maker then always points to the best consequence on this event.

Example. Firm F wants to evade negative publicity about its CO2 emissions. It worries that after making their choice, some stakeholder confronts them with evidence which state of the world obtains and proving to them that a bad consequence has resulted from their chosen act. Firm F wants to make sure that given the evidence they might be confronted with, there is a state of the world consistent with this evidence in which their chosen action yields a good result. The firm expects that only state $\{s, t\}$ can be proven by the stakeholder (for example by low market prices for RECs.) However, the firm does not expect anyone to be able to obtain definitive proof that either $\{s\}$ or $\{t\}$ obtains. If $\{s, t\}$ is sufficiently likely, the firm chooses RECs in order to be able to point to the good state t in which 100 megatons of CO2 have been reduced. However, if state $\{u\}$ is sufficiently likely, the firm will choose to improve its EFFICIENCY. □

Both of these decision models are extreme cases and unobservable consequences may generate a rich variety of decision models that are less extreme. The present decision models only represent a starting point for the exploration of decision theories in which consequences cannot be observed.

5 Axioms

This section provides a set of descriptive axioms that characterize expected verification and obfuscation utilities. Such axioms allow us to test whether decision makers' behavior is consistent with our decision models.

We assume the existence of a biseparable utility representation. Biseparable preferences allow us to identify the decision weights of individuals with respect to binary acts while making minimal assumptions about their behavior. The axioms for biseparable utility are provided in Ghirardato and Marinacci (2001) and for convenience restated in appendix G.

Axiom 0 (Biseparable Preference)**.** $\succsim$ is a *biseparable preference* if there exists a monotonic representation $U : \mathcal{A} \to \mathbb{R}$, an essential event $E \in \mathcal{E}$, a set function $\mu : \mathcal{E} \to [0, 1]$, such that for all $\gamma \succsim \beta$ and all events $F \in \mathcal{E}$:

$$U(\gamma F \beta) = \mu(F)U(\gamma) + (1 - \mu(F))U(\beta) \tag{3}$$

The image $U(\mathcal{X})$ is a convex set.

Thus, in the biseparable model, decision makers' preferences are only meaningfully restricted on binary acts. Under the assumption that certainty equivalents exist and preferences are biseparable, we can define preference averages of consequences and acts. These preference averages, first introduced in Ghirardato et al. (2003), allow us to use Anscombe and Aumann (1963) style axioms without resorting to objective probabilities being available. While objective probability lotteries can easily be implemented in a laboratory setting, for applications in the context of greenwashing behavior, such lotteries are often unavailable.

Definition 3 (Preference Average)**.** For all $x, y \in \mathcal{X}$ with $x \succsim y$, z is a *preference average* of x and y if $xEy \sim [xEz]E[zEy]$. z is denoted by $1/2x \oplus 1/2y$. For two acts $a, b \in \mathcal{A}$, c is a *pointwise preference average* of a and b if for all $s \in \mathcal{S}$, $c(s) = 1/2a(s) \oplus 1/2b(s)$. c is denoted by $1/2a \oplus 1/2b$.

Example. Suppose Trees and RECs refer to investing a fixed amount of money into the relevant carbon offset certificates. Suppose further that there are no returns to scale to either technology. If the decision maker is risk neutral (in terms of the curvature of u) over the amount of CO2 emissions reduced, then $1/2$ Trees $\oplus 1/2$ RECs refers to the act in which the carbon emissions in every state of the world are the arithmetic average of Trees and RECs and can be thought of as investing half the money into each technology. However, if the decision maker does not have a linear utility over the achieved reduction, this will generally not be the case. Instead, in each state $1/2$ Trees $\oplus 1/2$ RECs would offer the certainty equivalent of a (perfectly verifiable) equal probability lottery of the reduction achieved by Trees and RECs on that state. □

Definition 4 (Comonotonic Acts)**.** Acts $a, b \in \mathcal{A}$ are *comonotonic* if for all $s, s' \in \mathcal{S}$,

— $a(s) \succ a(s') \Rightarrow b(s) \succsim b(s')$, and

— $b(s) \succ b(s') \Rightarrow a(s) \succsim a(s')$.

Thus, two acts are comonotonic, if they agree on the ranking of states according to whether they achieve more desirable consequences. Two acts a, b are not comonotonic if we can find two states s, s' such that according to a, state s yields a strictly better consequence than s' and act b yields a strictly worse consequence on s than s'. Notice that constant acts are comonotonic with all acts.

Axiom 1 (Comonotonic Independence)**.** $\succsim$ fulfills *comonotonic independence* if for all comonotonic a, b, c, $a \succsim b$ if and only if $1/2a \oplus 1/2c \succsim 1/2b \oplus 1/2c$.

If the acts a, b, c are comonotonic then for every verifiable event the state that achieves the worst consequence is preserved by pointwise preference averages.

In the context of preference for verifiability, comonotonic independence has a natural interpretation. If we assume that the decision maker cares about which consequences can be proven to have obtained, comonotonic independence emerges as a natural condition. Pointwise mixing two acts a, b with a third c may affect which consequences can be proven to have obtained. However, as long as all three acts are comonotonic, the effect will be symmetric and no preference reversal should occur.

Example. In our example, all acts are comonotonic and thus decision makers with an expected verification utility or obfuscation utility fulfill independence with respect to these three acts. However, once we also allow for acts that yield a uniquely best consequence on u (or s) a decision maker with an expected verification or obfuscation utility would violate independence. Such acts could for example be the possibility to purchase an insurance on high prices of RECs which allow the decision maker to achieve a higher carbon reduction in case state u obtains. □

Axiom 2 (Supermodularity)**.** $\succsim$ fulfills *supermodularity (submodularity)* if for all events E, F, and all consequences $\gamma, \beta \in \mathcal{X}$,

$$1/2[\gamma E \cup F\beta] \oplus 1/2[\gamma E \cap F\beta] \quad \succsim (\precsim) \quad 1/2[\gamma E\beta] \oplus 1/2[\gamma F\beta]. \tag{4}$$

Supermodularity effectively states that given events E and F, the value of getting a good consequence on $E \cup F$ consists at least of the value of getting a good consequence on E and on F net the value of getting a good consequence on $E \cap F$. It may be strictly greater in case there are verifiable events that are neither a subset of E nor F but are contained in $E \cup F$. In this case, the decision maker is able to prove on some additional states that a good consequence has been achieved. Submodularity states the reverse, the value of getting a good consequence on $E \cap F$ is at most the value of getting a good consequence on E and F net the value of the intersection.

Example. A firm trying to obfuscate that their CO2 compensation program is ineffective has an interest to choose a compensation program that on any ex-post verifiable event is at least in some state consistent with a high reduction of CO2. This generates submodularity since achieving a good consequence on E may also be helpful if event F is verifiable. If the stakeholder proves that event F (or, E) arises, the firm can then point to a state in $E \cap F$ and argue that the subjective probability of this state is highly likely. □

Definition 5 (Min-increasing event)**.** Let $\gamma \succ \beta$. An event E is *min-increasing* if $\gamma_E\beta \succ \gamma_{E-F}\beta$ for all nonnull events $F \subset E$.

In our verification utility, min-increasing events are the events E for which the decision maker has a cover of verifiable events V that are all subsets of E. On a min-increasing event, obtaining at least γ on every subevent is critical or the utility will decrease. If an event E is not critical, then we can make $\gamma_E\beta$ worse on some state and the decision maker is indifferent to such a change.

Definition 6 (Max-increasing event)**.** An event E is *max-increasing* if $\beta_{E-F}\gamma \succ \beta_E\gamma$ for all nonnull events $F \subset E$ and some consequences $\gamma \succ \beta$.

An event E is *max-increasing* if $\beta_{E\cup F}\gamma \succ \beta_E\gamma$ for all nonnull events $F \subset \overline{E}$ and some consequences $\gamma \succ \beta$.

In our obfuscation utility, max-increasing events are the events E for which adding a better consequence on any state strictly increases the utility.

Max-increasing events play in obfuscation utility exactly the same role as min-increasing events play in verification utility. We therefore refer by the term *critical events* to *either* the min-increasing or the max-increasing events, depending on the context.

Axiom 3 (Critical event modularity)**.** $\succsim$ fulfills *critical event modularity* with respect to a set of events $\mathcal{C}$ if for all $E, F \in \mathcal{C}$, and any event $A \subseteq E \cup F$,

1. $E \cap F$ is a critical event,
2. $E \cup F$ is a critical event, and
3. $1/2[\gamma A \beta] \oplus 1/2[\gamma A \cap E \cap F \beta] \sim 1/2[\gamma A \cup E \beta] \oplus 1/2[\gamma A \cap F \beta]$

The first two conditions are intuitive. For example, if E and F are verifiable, then their intersection is also verifiable (by providing both the proof that E and the proof that F). Their union is then critical because no longer achieving a good consequence on every state of both events means that it is no longer possible on E or F, or both, to prove that a good consequence has been obtained.

Example. If E and F are critical, then the decision maker expects to be able to prove this in case these events actually obtain. Notice that in this case $E \cap F$ is also critical and so is $E \cup F$. Suppose for the moment that $A = E \cup F$. Critical event modularity states that the value obtained from obtaining a good consequence on both E and F can be separated into the value from obtaining a good consequence on E and obtaining a good consequence on F. If A is a subset of their union, then critical event modularity generalizes this separability even for non-critical $A \subset E \cup F$. □

5.1 Characterization of Expected Verification Utility

The above stated axioms contain necessary and sufficient conditions on preferences for an expected verification utility. More precisely, we obtain the following characterization result:

Theorem 1 (Representation Theorem)**.** *The following statements are equivalent:*

1. *$\succsim$ is a biseparable preference fulfilling comonotonic independence, supermodularity, and critical event modularity with respect to the min-increasing events.*
2. *$\succsim$ is an expected verification utility.*

The representation has standard uniqueness properties with respect to U and u. The interesting uniqueness properties are with respect to the set of verifiable events and subjective beliefs μ. Can we identify from behavior which events the decision maker expects to be able to verify? It turns out that while the set $\mathcal{V}$ is not unique, its smallest elements are unique and once we close $\mathcal{V}$ under unions, the resulting set is unique and so are the beliefs over elements of this set and complements thereof:

Proposition 1 (Uniqueness)**.** *Suppose $\succsim^1$ and $\succsim^2$ are expected verification utilities. Then $a \succsim^1 b \Leftrightarrow a \succsim^2 b$ for all $a, b \in \mathcal{A}$ if and only if:*

— $U^1 = \theta U^2 + \phi$,

— $u^1 = \theta u^2 + \phi$,

— $cl_\cup(\mathcal{V}^1) = cl_\cup(\mathcal{V}^2)$,

— $\forall E \in cl_\cup(\mathcal{V}^1) : \mu^1(E) = \mu^2(E)$ *and* $\mu^1(\overline{E}) = \mu^2(\overline{E})$.

for some $\theta \in \mathbb{R}_+$ and some $\phi \in \mathbb{R}$

Example. In our example, the verifiability of $\{s, t\}$ and $\{s, t, u\}$ can be inferred from behavior of the decision maker and need not be exogeneously given. Moreover, the subjective probabilities of $\{s, t\}$ and $\{u\}$ are uniquely determined from behavior. □

Effectively, if preferences over acts are identical then if one decision maker can prove that a certain event obtains, the other decision maker is able to prove that this event obtains or even able to prove that a strict subset of this event obtains. Thus, it could be the case that decision maker 1 is able to prove that $E, F, E \cup F, E \cap F$ obtain (whenever they actually obtain) while decision maker 2 is only able to prove that $E, F, E \cap F$ obtain. In terms of behavior, we would not see a difference because the decision maker will always want to provide the most precise proof possible.

5.2 Characterization of Expected Obfuscation Utility

We also obtain a dual theorem that characterizes obfuscation utility. Since an expected obfuscation utility is equivalent to an expected verification utility with a reverse order of preferences, supermodularity is replaced by submodularity and critical event modularity holds with respect to max-increasing instead of min-increasing events.

Theorem 2 (Representation Theorem)**.** *The following statements are equivalent:*

1. *$\succsim$ is a biseparable preference fulfilling comonotonic independence, submodularity, and critical event modularity with respect to the max-increasing events.*

2. *$\succsim$ is an expected obfuscation utility.*

Naturally, the expected obfuscation utility has the same uniqueness properties as expected verification utility. Expected obfuscation utility can be interpreted as the decision maker minimizing the utility of a stakeholder with expected verification utility who has a reverse order of preference of consequences. Similarly, expected verification utility can be seen as minimizing the utility of a stakeholder who wants to "find a hair in the soup", i.e., who wants to point at the possibility that a bad consequence may have been obtained by the decision maker.

6 Comparative Statics

We first define comparative risk preferences in a standard way. This will help us to contrast comparative risk aversion with other comparative results.

Definition 7 (Comparative Risk Preference)**.** $\succsim^2$ is at least as risk averse than $\succsim^1$ if whenever $1/2x \oplus 1/2y \succsim^1 z$, then $1/2x \oplus 1/2y \succsim^2 z$. Two decision makers are equally risk averse if each is at least as risk averse as the other.

The following result is standard.

Proposition 2 (Comparative Risk Aversion)**.** *Suppose $\succsim^1$ and $\succsim^2$ are expected verification utilities or expected obfuscation utilities. Then the following statements are equivalent:*

1. *$\succsim^2$ is at least as risk averse as $\succsim^1$.*

2. *$u^2 = t \circ u^1$ for some continuous, concave function t.*

The following result shows that risk attitude is unrelated to the attitude towards verifiable events. Thus, the concern for verifiability cannot be captured by risk aversion and vice versa. We can see risk aversion as an aversion to ex-ante uncertainty of perfectly verifiable consequences and preference for verifiability as an aversion to ex-post uncertainty. Moreover, we only require identical preferences over two consequences in order to compare two decision makers with respect to their verifiable events.

Proposition 3 (Comparative Statics). *Suppose $\succsim^1$ and $\succsim^2$ are expected verification utilities with $\gamma \succ^1 \beta$ and $\gamma \succ^2 \beta$ and identical null events. Then the following statements are equivalent:*

1. $cl_{\cup}(\mathcal{V}^1) \subseteq cl_{\cup}(\mathcal{V}^2)$.
2. $\gamma E \beta \sim^2 \gamma(E-F)\beta$ *implies* $\gamma E \beta \sim^1 \gamma(E-F)\beta$.

The result shows why critical events are at the core of preference for verifiability. If one decision maker's set of critical events is a subset of the other decision maker's, then it is also necessarily the case that their set of verifiable events is a subset.

In the context of our example, we may care about how inefficient a decision maker's choice is when following an expected verifiability utility or an expected obfuscation utility. A plausible benchmark is expected utility maximization because the agent is probabilistically sophisticated over verifiable events.

Definition 8 (Welfare Loss). Let $\succsim$ be an expected verification (obfuscation) utility with representation $U, u, \mu, \mathcal{V}$. Let $A \subset \mathcal{A}$ be a finite set over which $\succsim$ is strict. The welfare loss is defined as

$$L_{\succsim,\mu}(A) = \max_{a \in A} \int u \circ a d\mu - \int u \circ \left(\arg\max_{a \in A} U \right) d\mu \tag{5}$$

It is noteworthy that from observing expected verification preferences alone we cannot uniquely determine the welfare loss because the probabilities of non-verifiable events are not uniquely determined. If some consequences are not perfectly verifiable, we need to complement the preference information with probability information of non-verifiable events. However, it turns out that even if we make the assumption that a policy maker picks the "correct" μ, policy interventions that change the set of verifiable events or that change firms' behavior from obfuscation-seeking to verification-seeking do not necessarily reduce the welfare loss. The following examples show that (holding fixed the verifiable events) a decision maker acting according to expected verification utility may indeed face a larger welfare loss than a decision maker acting according to expected obfuscation utility and that increasing the set of verifiable events need not lead to a welfare improvement. This is noteworthy because this implies that calls for better verifiability of CO2 reduction and/or incentives against greenwashing need to be carefully evaluated for whether they will achieve the desired purpose.

Corollary 1 (Verification and Obfuscation Welfare Loss Comparison)**.** *Let $\succsim$ be an expected verification utility and $\succsim^\dagger$ an expected obfuscation utility with a nontrivial set of verifiable events $\mathcal{V}$. Let the risk preferences of $\succsim$ and $\succsim^\dagger$ be identical. Then either one, but not both, of the following statements is true:*

1. *For all μ and all decision problems A, $L_{\succsim,\mu}(A) = L_{\succsim^\dagger,\mu}(A) = 0$.*

2. *There exist μ and decision problems A and $A^\dagger$ such that $L_{\succsim,\mu}(A) > L_{\succsim^\dagger,\mu}$ and $L_{\succsim^\dagger,\mu}(A^\dagger) > L_{\succsim,\mu}(A^\dagger)$.*

Proof. If the welfare loss is always equal to zero, then the set of verifiable events must contain all singletons of $\mathcal{S}$. If the welfare loss is nonzero for some μ and some decision problem A, then there must exist states s, t that are null but not irrelevant. Let E be a critical event containing s, t.

Suppose A contains an act a^*_μ that is optimal under expected utility and $a^\dagger_\mu$ be the act that is optimal according to $\succsim^\dagger$. Then it must be the case that for some verifiable event E, $a^\dagger$ implements different consequences on at least two states in E and $\mu(E)\max_{s\in E} u(a^\dagger(s)) - u(a^*(s)) > \int_E u(a^\dagger(s)) - u(a^*(s))d\mu \geq \mu(E)\min_{s\in E} u(a^\dagger(s)) - u(a^*(s))$. Notice that we can change μ just on the event E such that the central term becomes identical to the LHS. Moreover, without loss of generality we can make a^* and $a^\dagger$ identical on all states outside E. Doing so for all verifiable events eliminates the welfare loss for $\succsim^\dagger$. Then, as $\mu(s^\dagger) \to 1$, according to expected utility, $a^\dagger$ is strictly optimal and as $\mu(s^\dagger) \to 0$, a^* is optimal. □

Even though we may intuitively find that behavior according to expected verification utility is normatively more appealing than that of expected obfuscation utility, the expected utility loss may be larger.

Example. Suppose in the example decision problem states s and u are very unlikely. In this case, an expected utility maximizer would choose RECs, the same action as a maximizer of expected obfuscation utility. A maximizer of expected verifiability utility would choose Trees and incur a welfare loss. This is because the decision maker fears being blamed for an ex-post suboptimal action that was ex-ante optimal. □

Definition 9 (Comparative Loss from Intransparency)**.** Let $\succsim$ and $\succsim'$ be expected verification (obfuscation) utility with identical risk preferences but different sets of verifiable events $\mathcal{V} \subseteq \mathcal{V}'$. Then the welfare loss due to intransparency is defined as:

$$T_{\succsim,\mu}(A, \mathcal{V}') = L_{\succsim,\mu}(A) - L_{\succsim',\mu}(A) \tag{6}$$

where $\succsim'$ is the expected verification (obfuscation) utility with representation $U, u, \mu, \mathcal{V}'$.

The following result is obvious, but highlighted for its policy relevance:

Corollary 2 (Welfare Loss Compared to Perfect Information)**.** *The welfare loss compared to perfect information is always nonnegative,* $T_{\succsim,\mu}(A, 2^S) \geq 0$.

It follows that a policy maker who could influence the set of verifiable events would always want to implement ex-post certainty about the consequences. The question is whether more generally any information increase reduces the welfare loss. Again, it is possible to find cases in which an information increase may increase the welfare loss.

Corollary 3 (Indeterminacy of Welfare Loss)**.** *For all* $\mathcal{V} \subset \mathcal{V}' \neq 2^S$*, there exist decision problems* A*,* A' *such that* $T_{\succsim,\mu}(A, \mathcal{V}') < 0 < T_{\succsim,\mu}(A', \mathcal{V}')$

Example. In our example, suppose $\mathcal{V}' = \{\{s\}, \{s,t,u\}\}$ and $\mathcal{V} = \{\{s,t,u\}\}$. Let A contain only the act Trees and RECs with a slight payoff increase of RECs in state u. Suppose the beliefs are such that $\mu(\{t\}) = .99$, meaning that an expected utility maximizer would strictly prefer RECs. Under verifiable events $\mathcal{V}$, the decision maker would prefer RECs and there is zero welfare loss. In $\mathcal{V}'$ the decision maker is able to verify that event $\{s\}$ obtains, and now prefers Trees over RECs because in state s the act Trees yields a higher payoff. However, this leads to a strictly positive welfare loss despite providing more information. Together, this and the previous example show that policy decisions regarding the incentives for transparency and efficiency of carbon emission reduction are nontrivial. While transparency and verifiability are perhaps intrinsically desirable, they may come at the cost of firms choosing less efficient CO2 reduction strategies. This suggests that policies that are aimed at increasing ex-post verifiability need to be carfully examined whether the gain in transparency might be offset by a loss in efficiency — unless, ideally, the policy implements full ex-post transparency. □

7 Relation to Literature

Our model is a special case of the classic ambiguity models of maxmin expected utility (Gilboa & Schmeidler, 1989), and Choquet expected utility (Schmeidler, 1989). Most closely related to the present paper is the dual self model of ambiguity (Chandrasekher et al., 2022). In this model, decision makers maximize an objective $U(a) = \max_{P \in \mathcal{P}} \min_{p \in P} \int u \circ a dp$ where $\mathcal{P}$ is a set of sets of priors. The interpretation of the model is that there are two selves, a pessimistic one and an optimistic one. The pessimistic one evaluates the acts according to the maxmin expected utility model (Gilboa & Schmeidler, 1989) by choosing the worst possible prior. However, the optimistic self determines the set of available priors that the pessimist can choose from. The present model is a special case of the dual self model. In a dual self interpretation of our model, the maximizer chooses a minimal cover of verifiable events. Each such cover corresponds to a set of priors via the restriction that a prior p needs to fulfill $p(E) = \mu(E)$ for all events E in the cover. In other words, to evaluate an act, the minimizer can choose to assign the entire probability of a verifiable event to the worst state of the verifiable event. Relative to the characterization of (Chandrasekher et al., 2022), our model provides conditions under which the maximization and minimization steps can be written inside the expectation, i.e., we provide a way to determine ambiguity-neutral events.

If the set of verifiable events is exogeneously given, a characterization using cominimum additivity (Kajii et al., 2007, 2009) instead of comonotonic independence is possible. Cominimum additivity with respect to a set of events $\mathcal{V}$ requires that all acts that agree on the worst states in all of the events in $\mathcal{V}$ have to fulfill the independence axiom. Cominimum additivity provides intermediate cases between full independence and comonotone additivity.

Preference for verifiability is a form of information preference since the decision maker cares about how much information is ex post available about which consequence has been achieved. Information preferences are often modeled using a two stage approach (Dillenberger & Raymond, 2019; Kreps & Porteus, 1978; Segal, 1990). Given our decision model, the first stage would correspond to the verifiable events and the second stage to the final consequences and only the first stage is observed. If verifiable events $\mathcal{V}$ are objectively given and both ex-ante and ex-post probability distributions over consequences are objective, it is possible to perform our analysis using preferences over $\Delta\mathcal{V} \times \Delta\mathcal{X}^{\mathcal{V}}$. However, this way of modeling would severely restrict the explanatory power of our model. It does not allow to infer the verifiable events from behavior but requires the analyst to have data on the ex-post beliefs of the decision maker. In the context of our example, this would defeat the main purpose of the paper since ex-post the stakeholder and decision maker would agree on the probability distribution of the CO_2 emissions.

The present paper also provides a decision theoretic foundation for the standard notion of verifiability employed in contract theory starting with Bull and Watson (2004): verifiable events are closed under intersection but not necessarily under relative complements. Our application shows that this notion can be productively used to model greenwashing.

Our research also relates to definitions (de Freitas Netto et al., 2020) and formal models of greenwashing (Wu et al., 2020) and of green products (Groening et al., 2018). Unlike previous modeling attempts, the present paper provides a purely behavioral definition of greenwashing versus verification-seeking behavior. We do not require unobservable model components such as incorrect consumer beliefs due to deception or detailed information about the interaction between consumers and firms. This yields a very parsimonious model of greenwashing which does not require any information about the behavior of consumers of greenwashed products – preference data of the firm over "green" policies fully identifies the model.

8 Discussion

When consequences are not directly observable to decision makers, it is plausible that these decision makers deviate from expected utility. The present paper provides a starting point for the analysis of such deviations. Specifically, decision makers may for example care about verifying or obfuscating what consequence has been achieved in the presence of verifiable information. As our axiomatic analysis shows, unobservable consequences may provide a rationalization for an extreme form of ambiguity preferences. Crucially, such behavior can be identified from preferences over acts alone and it is not necessary to introduce additional primitives. In particular, we neither need to enrich the outcome space by ex-post beliefs over consequences nor do we need to impose any information structure of what is verifiable.

Starting with UNFCCC (2023), there have been attempts to increasingly rely on CCR to achieve CO_2 emission reductions. These attempts have been criticised as "inappropriately verified" (Institute, 2023, p.6). Our comparative statics results suggest that unless perfect ex-post verifiability is implemented, policies aimed at increasing verifiability of consequences may not be suitable to reduce the welfare losses from verification-seeking and obfuscation-seeking behavior. In such cases, traditional environmental policies that directly incentivize expected CO2 emission reductions could be more suitable than a combination of CCR and transparency policies.

Appendix A Proof of Theorem 1

Proof. By comonotonic independence and the biseparable utility representation we obtain (see Proposition 2 of Ghirardato et al., 2003) a representation $U(a) = \int (u \circ a) d\mu$ where μ is a capacity and the integral is in the sense of Choquet. Let m^μ be the Möbius inverse of μ, i.e., $m^\mu(E) = \sum_{A \subseteq E} (-1)^{|A|-1} \mu(A)$. The Choquet integral expressed in terms of the Möbius inverse (see Grabisch, 2016, p. 235) is given by

$$\int (u \circ a) d\mu = \sum_{E \in \mathcal{E}} m^\mu(E) \min_{s \in E} u(a(s)). \tag{A.7}$$

We now prove the key lemma which relates properties of m^μ to whether a set is critical.

Lemma 1. *For all $S \subseteq \mathcal{S}^*$:*

- $m(S) \geq 0$*, and*
- *if $m(S) > 0$ then S is critical and there does not exist a cover of critical sets that are all strict subsets of S.*

Proof. We prove this by induction on the cardinality of S, n.

Case $n = 1$: Since μ is a capacity, $m(S) + m(\emptyset) = \mu(S) \geq \mu(\emptyset) = m(\emptyset) = 0$. Thus $m(S) \geq 0$. If $m(S) = \mu(S) > 0 = \mu(\emptyset)$, then since S contains a single element, S is critical and has no critical subsets.

Case $n > 1$: Suppose for all sets S of size $n-1$ or smaller the induction hypothesis holds. We distinguish the case that $m(S) < 0$ from the case that $m(S) > 0$ and there exists a cover of S of critical events that are strict subsets of S and derive a contradiction for each case.

- Suppose for sake of contradiction that $m(S) < 0$.

We first show that S is critical and there exists a cover of critical subsets of S. For every $s \in S$ we have by monotonicity of the capacity that $\mu(S) \geq \mu(S - \{s\})$. It follows from the definition of m that $m(S) + \mu(S - \{s\}) + \sum_{E \subset S : s \in E} m(E) = \mu(S) \geq \mu(S - \{s\})$ and thus $m(S) + \sum_{E \subset S : s \in E} m(E) \geq 0$. Therefore for some $E \subset S$ containing s we have that $m(E) > 0$ and by the induction hypothesis E is critical. It follows that every $s \in S$ is contained in a critical event and since by critical event modularity critical events are closed under unions, S is also critical.

Since $n \geq 2$ and there exists a cover of critical subsets of S and critical subsets are closed under unions, we can find a cover $\{A, B\}$ of two critical subsets of S with $A - B$ and $B - A$ nonempty. Then by critical event modularity, $\mu(S) + \mu(A \cap B) = \mu(A) + \mu(B)$. Thus, by the definition of m, $\sum_{E \subseteq S} m(E) + \sum_{E \subseteq A \cap B} m(E) = \sum_{E \subseteq A} m(E) + \sum_{E \subseteq B} m(E)$. This is equivalent to: $m(S) + \sum_{E \subset S : E \not\subseteq A, B} m(E) = 0$. Because by assumption $m(S) < 0$, this is only possible if $m(E) > 0$ for some $E \subset S$, $E \not\subseteq A, B$. But from critical event modularity follows that $\mu(E) + \mu(E \cap A \cap B) = \mu(A \cap E) + \mu(B \cap E)$ and thus $\sum_{F \subset E : F \not\subseteq A, F \not\subseteq B} m(F) < 0$, contradicting the induction hypothesis.

— Suppose for sake of contradiction that $m(S) > 0$ and there exists a cover of critical subsets of S. Then $\mu(S) + \mu(A \cap B) = \mu(A) + \mu(B)$ by critical event modularity. Thus, $\sum_{E \subseteq S} m(E) + \sum_{E \subseteq A \cap B} m(E) = \sum_{E \subseteq A} m(E) + \sum_{E \subseteq B} m(E)$ which is equivalent to: $m(S) + \sum_{E \subset S : E \not\subseteq A, B} = 0$. But this is only possible if $m(E) < 0$ for some $E \subset S$, contradicting the induction hypothesis.

□

Lemma 2. *If $E, F \in \mathcal{E}$, $m^{\mu}(E) > 0$, $m^{\mu}(F) > 0$ and $s \in E \cap F$, then there exists $G \in \mathcal{E}$ such that $G \subseteq E \cap F$, $m^{\mu}(G) > 0$, and $s \in G$.*

Proof. By Lemma 1, E and F are critical. It follows from critical event modularity that $E \cap F$ is critical. Suppose $m(G) = 0$ for all $G \subseteq E \cap F$ such that $s \in G$. Then $\mu(E \cap F) = \sum_{G \subseteq E \cap F} m^{\mu}(G) = 0 + \sum_{G \subseteq (E \cap F) - \{s\}} = \mu((E \cap F) - \{s\})$, contradicting that $E \cap F$ is critical. Thus, there exists some $G \subseteq E \cap F$ such that $m^{\mu}(G) > 0$. □

Lemma 3. *For every $s \in \mathcal{S}^*$ there exists a unique event $\phi(s) \in \mathcal{E}$ such that $m^{\mu}(\phi(s)) > 0$ and if $s \in F$ and $\mu(F) > 0$, then $\phi(s) \subseteq F$.*

Proof. Since $s \in \mathcal{S}^*$, it must be included in at least one event E with $m^\mu(E) > 0$. Since $\mathcal{S}$ and $\mathcal{E}$ are finite, there exists some E such that $m^\mu(E) > 0$ and $m^\mu(F) = 0$ for all $F \subset E$ such that $s \in F$. To see that there exists only one such event, suppose $s \in E \cap F$, $m^\mu(E) > 0$ and $m^\mu(F) > 0$. Then for some $G \subseteq E \cap F$, $m^\mu(G) > 0$. If neither E nor F have strict subsets on which m^μ is strictly positive, then by Lemma 2 it must be the case that $E = F = G$. □

Let $\phi : \mathcal{S}^* \to \mathcal{E}$ be the function that maps states s into the smallest event $E \ni s$ such that $m^\mu(E) > 0$. Let $\phi^{-1}(E) = \{t \in E | \phi(t) = E\}$ be the set of all states that ϕ maps into event E. By Lemma 3, ϕ and ϕ^{-1} are well defined.

We define a probability measure η inductively by: $\eta(\{s\}) = 0$ if $s \notin \mathcal{S}^*$, $\eta(\{s\}) = m^\mu(\phi(s))/|\phi^{-1}(\phi(s))|$ if $s \in \mathcal{S}^*$, and $\eta(E \cup \{s\}) = \eta(E) + \eta(\{s\})$ for all $E \in \mathcal{E}$ and $s \notin E$. Thus, for every state s we find the probability mass of the states $\phi^{-1}(E)$ and divide it evenly among these states.

Denote $\mathcal{V} = \{E \in \mathcal{E} | m^\mu(E) > 0\}$, then:

$$
\begin{aligned}
&\int_{s \in \mathcal{S}} \max_{E \in \mathcal{V}: s \in E} \min_{t \in E} u(a(t)) d\eta \\
&\sum_{s \in \mathcal{S}^*} \eta(s) \max_{E \in \mathcal{V}: s \in E} \min_{t \in E} u(a(t)) \\
&= \sum_{s \in \mathcal{S}^*} \eta(s) \min_{t \in \phi(s)} u(a(t)) \\
&= \sum_{E \in \mathcal{E}} \sum_{s \in \phi^{-1}(E)} \eta(s) \min_{t \in E} u(a(t)) \\
&= \sum_{E \in \mathcal{E}} m^\mu(E) \min_{t \in E} u(a(t)) = U(a). \qquad \text{(A.8)}
\end{aligned}
$$

The first equality sign follows from $\eta(s) = 0$ for all $s \notin \mathcal{S}^*$. The second equality sign follows since if $F \ni s$ and $F \in \mathcal{V}$, then by the Lemma 3, $\phi(s) \subseteq F$ and thus $\min_{t \in \phi(s)} u(a(t)) \geq \min_{t \in F} u(a(t))$. The third equality sign follows since $\phi : \mathcal{S}^* \to \mathcal{E}$ is a well defined function and thus each state appears exactly once in the summation $\sum_{E \in \mathcal{E}} \sum_{s \in \phi^{-1}(E)}$. The fourth equality sign follows by definition of η, since $m^\mu(E) = \sum_{s \in \phi^{-1}(E)} \eta(s)$. □

Appendix B Proof of Proposition 1

Proof. $\Leftarrow$ is trivial, we prove $\Rightarrow$. Suppose $\succsim^1 = \succsim^2$.

From the uniqueness properties of the biseparable preferences follows that $U^1 = \theta U^2 + \phi$ and $u^1 = \theta u^2 + \phi$. What is left to show is the relation between the verifiable sets $\mathcal{V}^1$ and $\mathcal{V}^2$. Note that the critical events in the representation are exactly the sets E for which $m^{\mu}(E) > 0$ or for which there exists a cover of sets $E_1, \ldots, E_n \subseteq E$ such that $\forall i : m^{\mu}(E_i) > 0$.

$\mathcal{V}^1 \subseteq \mathcal{V}^2$: If $DV = cl_{\cup}(\mathcal{V}^1) - cl_{\cup}(\mathcal{V}^2)$ is nonempty, then for some $V \in DV$, V is critical in U^1. We show that V cannot be critical in U^2: If V is critical in U^2, then either $m^{\mu^2}(V) > 0$ and thus $V \in \mathcal{V}^2$ or there exists a cover $V_1, \ldots, V_n$ such that $\forall i : m^{\mu^2}(V_i) > 0$. If the latter is the case, then $V_1, \ldots, V_n \in \mathcal{V}^2$ and thus $V \in cl_{\cup}(\mathcal{V}^2)$. It follows that $V \in DV$ cannot be critical in U^2. But if V is critical in U^1 but not in U^2, then the preferences cannot be identical.

$\mathcal{V}^2 \subseteq \mathcal{V}^1$ follows by a symmetric argument.

$\forall E \in cl_{\cup}(\mathcal{V}^1) : \mu^1(E) = \mu^2(E)$: Suppose for some $E \in cl_{\cup}(\mathcal{V}^1)$, $\mu^1(E) > \mu^2(E)$. Then, for some $\gamma \succ \beta \succ \alpha$, $\gamma E \beta \succ_1 \alpha$ and $\alpha \succ_2 \gamma E \beta$, yielding a contradiction. □

Appendix C Proof of Theorem 2

Proof. We define a dual preference $\succsim^{\dagger}$ by: $a \succsim^{\dagger} b \Leftrightarrow b \succsim a$.

Lemma 4 (Comonotonic Independence Equivalence). *$\succsim$ fulfills comonotonic independence if and only if $\succsim^{\dagger}$ fulfills comonotonic independence.*

Proof. Let a, b, c be comonotonic acts. Then,

$$\begin{array}{ccc} a \succsim b & \Leftrightarrow^{def.} & b \succsim^{\dagger} a \\ \Leftrightarrow^{com.ind.} & & \Leftrightarrow^{com.ind.} \\ \alpha a \oplus (1-\alpha)c \succsim \alpha b \oplus (1-\alpha)c & \Leftrightarrow^{def.} & \alpha a \oplus (1-\alpha)c \succsim^{\dagger} \alpha b \oplus (1-\alpha)c \end{array} \tag{C.9}$$

□

The next lemma is trivial.

Lemma 5 (Supermodularity-Submodularity Correspondence). *$\succsim$ fulfills submodularity if and only if $\succsim^{\dagger}$ fulfills supermodularity.*

Lemma 6 (Critical Event Modularity Correspondence). *The max-increasing events fulfill critical event modularity in $\succsim$ if and only if the min-increasing events fulfill critical event modularity in $\succsim^{\dagger}$.*

Proof. $\overline{E}$ is min-increasing if $\beta E\gamma \succ^{\dagger} \beta E \cup F\gamma$ for all nonnull events $F \subset \overline{E}$. E is max-increasing if $\beta E \cup F\gamma \succ \beta E\gamma$ for all nonnull events $F \subset \overline{E}$. Thus, $\overline{E}$ is min-increasing in $\succsim^{\dagger}$ if and only if E is max-increasing $\succsim$.

Suppose E and F are min-increasing in $\succsim^{\dagger}$. Then $\overline{E}$ and $\overline{F}$ are max-increasing in $\succsim$. Thus, $\overline{E} \cup \overline{F}$ and $\overline{E} \cap \overline{F}$ are max-increasing in $\succsim$. It follows that their complements $E \cap F$ and $E \cup F$, respectively, are min-increasing in $\succsim^{\dagger}$ and thus $\succsim^{\dagger}$ fulfills critical event modularity. □

Then $\succsim^{\dagger}$ fulfills comonotonic independence, supermodularity, and critical event modularity. It follows that $\succsim$ can be represented by:

$$\begin{aligned} U(a) &= -\int \max_{E \in \mathcal{V}} \min_{s \in E} u(a(s)) d\mu \\ &= -\int \max_{E \in \mathcal{V}} - \max_{s \in E} -u(a(s)) d\mu \\ &= \int \min_{E \in \mathcal{V}} \max_{s \in E} -u(a(s)) d\mu \end{aligned} \tag{C.10}$$

□

Appendix D Proof of Proposition 2

Appendix E Proof of Proposition 3

Proof. $1 \Rightarrow 2$: If $\gamma E\beta \sim^{2} \gamma(E-F)\beta$, then E is not critical. Since the critical events of decision maker 2 contain all critical events of decision maker 1, E cannot be critical for decision maker 1. Thus, $\gamma E\beta \sim^{1} \gamma(E-F)\beta$.

$2 \Rightarrow 1$: If E is critical for decision maker 1 but not critical for decision maker 2, then for decision maker 2 there exists some set $F \subset E$ such that $\gamma E\beta \sim^{2} \gamma(E-F)\beta$. But then also $\gamma E\beta \sim^{1} \gamma(E-F)\beta$ and thus E cannot be critical for decision maker 1. Thus, the critical sets of decision maker 1 are a subset of the critical sets of decision maker 2. □

Appendix F Proofs of Corollaries 1-3

Appendix G Axioms for Biseparable Preferences

In the main text, a biseparable preference is assumed. Here we restate the axioms used by Ghirardato and Marinacci (2001) to characterize biseparable preferences. The axioms are slightly adjusted to fit the notation of the present paper.

Axiom 4 (Preference Relation). $\succsim$ is a *nontrivial preference relation* if it is a complete, transitive, nontrivial binary relation.

Definition 10 (Dominant Acts). An act f *dominates* an act g if for all $s \in \mathcal{S}$, $f(s) \succsim g(s)$.

Axiom 5 (Dominance). $\succsim$ fulfills *dominance* if for all $f, g \in \mathcal{A}$, whenever f dominates g, then $f \succsim g$.

Axiom 6 (Eventwise Monotonicity). $\succsim$ fulfills *eventwise monotonicity* if for all $E \in \mathcal{E}$ and $x, y, z \in \mathcal{X}$ such that $y \succsim z$,

— if E is nonnull, then $x \succ y$ implies $xEz \succ yEz$, and

— if E is nonuniversal, then $x \succ z$ implies $yEx \succ yEz$.

Let $\mathcal{X}$ be a topological space and let $\mathcal{X}^{\mathcal{S}}$ be endowed with the product topology.

Axiom 7 (Continuity). $\succsim$ fulfills *continuity* if for all nets $\{f_\alpha\}_{\alpha \in D} \subseteq X^{\mathcal{S}}$ such that f_α and f are measurable with respect to the same finite partition, whenever $f_\alpha \succsim g$ ($g \succsim f_\alpha$) for all $\alpha \in D$, then $f \succsim g$, ($g \succsim f$).

Continuity ensures the existence of certainty equivalents and therefore the subjective mixtures (Nakamura, 1990) are well defined:

Definition 11 (Subjective Mixture). A *subjective mixture* is an act $f \oplus_E g : s \mapsto [f(s)Eg(s)]$.

Axiom 8 (Binary Comonotonic Act Independence). $\succsim$ fulfills *binary comonotonic act independence* if for all essential events $E \in \mathcal{E}$, every $F \in \mathcal{E}$, and for all pairwise comonotonic acts $xEy, x'Ey', x''Ey''$, if $x''Ey''$ is dominated by both xEy and $x'Ey'$ or dominates both xEy and $x'Ey'$, then

$$xEy \succsim x'Ey' \quad \Rightarrow \quad (xEy) \oplus_B (x''Ey'') \succsim (x'Ey') \oplus_B (x''Ey''). \tag{G.11}$$

Theorem 3 (Biseparable Utility Representation (Ghirardato & Marinacci, 2001)). *Suppose $\succsim$ is a relation on $\mathcal{A}$ and there is at least one essential event $E \in \mathcal{E}$, then the following statements are equivalent:*

1. *$\succsim$ is a nontrivial preference relation that fulfills dominance, eventwise monotonicity, continuity, and binary comonotonic act independence.*

2. *$\succsim$ is a biseparable preference.*

References

Anscombe, F. J., & Aumann, R. J. (1963). A Definition of Subjective Probability. *The Annals of Mathematical Statistics*, *34*(1), 199–205.

Bull, J., & Watson, J. (2004). Evidence disclosure and verifiability. *Journal of Economic Theory*, *118*(1), 1–31.

Chandrasekher, M., Frick, M., Iijima, R., & Le Yaouanq, Y. (2022). Dual-self representations of ambiguity preferences. *Econometrica*, *90*(3), 1029–1061.

de Freitas Netto, S. V., Sobral, M. F. F., Ribeiro, A. R. B., & da Luz Soares, G. R. (2020). Concepts and forms of greenwashing: A systematic review. *Environmental Sciences Europe*, *32*(19), 1–12.

Dillenberger, D., & Raymond, C. (2019). Additive-Belief-Based-Preferences.

European Parliament. (2023). *Eu to ban greenwashing and improve consumer information on product durability* [https://www.europarl.europa.eu/news/en/press-room/20230918IPR05412/eu-to-ban-greenwashing-and-improve-consumer-information-on-product-durability].

Ghirardato, P., & Marinacci, M. (2001). Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of operations research, 26(4), 864-890.*, *26*(4), 864–890.

Ghirardato, P., Maccheroni, F., Marinacci, M., & Siniscalchi, M. (2003). A subjective spin on roulette wheels. *Econometrica*, *71*(6), 1897–1908.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, *18*(2), 141–153.

Grabisch, M. (2016). *Set functions, games and capacities in decision making*. Springer.

Groening, C., Sarkis, J., & Zhu, Q. (2018). Green marketing consumer-level theory review: A compendium of applied theories and further research directions. *Journal of Cleaner Production*.

Institute, N. C. (2023). *Corporate climate responsibility monitor* [https://newclimate.org/resources/p climate-responsibility-monitor-2023].

Kajii, A., Kojima, H., & Ui, T. (2007). Cominimum additive operators. *Journal of Mathematical Economics*, *43*, 218–230.

Kajii, A., Kojima, H., & Ui, T. (2009). Coextrema additive operators. In S. K. Neogy, A. K. Das, & R. B. Bapat (Eds.), *Modeling, computation and optimization* (pp. 73–95). World Scientific.

Kreps, D. M., & Porteus, E. L. (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica*, *46*(1), 185. https://doi.org/10.2307/1913656

Nakamura, Y. (1990). Subjective expected utility with non-additive probabilities on finite state spaces. *Journal of Economic Theory*, *51*, 346–366.

Schmeidler, D. (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica*, *57*(3), 571–587.

Segal, U. (1990). Two-Stage Lotteries without the Reduction Axiom. *Econometrica*, *58*(2), 349. https://doi.org/10.2307/2938207

UNFCCC. (2023). *Race to zero progress report* [https://climatechampions.unfccc.int/system/race-to-zero/].

Wu, Y., Zhang, K., & Xie, J. (2020). Bad greenwashing, good greenwashing: Corporate social responsibility and information transparency. *Management Science*, *66*(7), 3095–3112.