Title

Subjective complexity under uncertainty

Author
Valenzuela-Stookey, Quitzé

Publication Date
2023-11-01

DOI
10.1016/j.mathsocsci.2023.10.001

Subjective Complexity Under Uncertainty

Quitzé Valenzuela-Stookey*

October 3, 2023

Abstract

Complexity of the problem of choosing among uncertain acts is a salient feature of many of the environments in which departures from expected utility theory are observed. I study a class of Generalized Simple Bounds preferences in which acts that are complex from the perspective of the decision maker are bracketed by "simple" acts to which they are related by statewise dominance. I then study a refinement of the model in which the size of the partition with respect to which an act is measurable arises endogenously as a measure of subjective complexity. Finally, I consider choice behavior characterized by a "cautious completion" of Simple Bounds preferences, and discuss the relationship between this model and models of ambiguity aversion. I develop general comparative statics results, and explore applications to portfolio choice and insurance choice.

Keywords— Complexity, uncertainty, bounded rationality, ambiguity

Acts with uncertain outcomes are complicated, potentially infinite-dimensional, objects. The complexity of choosing between uncertain acts has been evoked as an explanation for many empirical observations at odds with the predictions of the subjective expected utility model, from ambiguity aversion (e.g. Gilboa and Schmeidler (1989)) to the so-called "equity premium puzzle" of Mehra and Prescott (1985). As early as Von Neumann and Morgenstern (1944) and Aumann (1962), it has been recognized that completeness of preferences over such acts may be a lot to ask.

This paper studies a choice-based notion of complexity in choice under uncertainty. I want to take seriously the idea that cognitive limitations may make certain pairs of acts

*Department of Economics, UC Berkeley; 530 Evans Hall, Berkeley, CA 94720, email: quitze@berkeley.edu. Many thanks to Eddie Dekel, Alessandro Pavan, and especially Marciano Siniscalchi for feedback on this project from its inception. I am also grateful to Edi Karni, Wojciech Olszewski, Francisco Poggi, Ludvig Sinander, Bruno Strulovici, Marie-Louise Vierø, participants at the Transatlantic Theory Workshop 2019, RUD 2020, and seminar participants at Northwestern University.

difficult to compare, which means I must allow the preference relation to be incomplete. This is in contrast to a notion of complexity in which the decision maker simply undervalues acts that are perceived as complex.

I take as primitive a reflexive and transitive, but not necessarily complete, binary preference relation. Preferences are assumed to be complete over the set of constant acts. Therefore, if preferences are incomplete, it must be that not all acts can be assigned a precise utility value, which really means a certainty equivalent among the constant acts. We can thus partition the set of acts into those for which the DM can and cannot identify a certainty equivalent. I refer to the former as "well-understood" acts, and denote the set of such acts by F_{CE}.[1]

Certain comparisons may also be easier to understand than others. At the very least, the decision maker should be able to conclude that act f is better than g if f gives a better consequence state-by-state (i.e. preferences should satisfy Monotonicity). When facing a choice between acts that are not well understood, the decision maker can gain traction by exploiting simple comparisons to acts that they understand well. In particular, consider the comparison between two acts f and g which are not well understood, and are not related by statewise dominance. Suppose there exists a well understood act h that is dominated statewise by f, and another well-understood act k which statewise dominates g. The DM can easily see that f is better than h, and that k is better than g. However since h and k are both well understood, meaning they can be ranked against all constant acts, they are necessarily comparable. If h is better than k then the DM can conclude by transitivity that f is better than g.

If on the other hand k is better than h in the above example, then the comparison is not immediately useful for assessing f versus g, without further assumptions. I would like to maintain minimal assumptions regarding the decision maker's ability to extrapolate from their preferences over well-understood acts. I therefore assume only that the decision maker can make comparisons involving an act that is not well understood if the pair of acts in question are related by statewise dominance. Of course, transitivity implies that

[1] The idea behind the label "well-understood" is that if a decision maker understands an act well, they should be able to calibrate its value precisely in relation to constant acts. Conversely, a certainty equivalent can be used as a simple proxy for an act f when comparing f to other acts, rendering such comparisons relatively easy.

further comparisons may be possible, via the type of reasoning described in the previous paragraph (formally, this means that $\succsim$ is the transitive and monotone closure of preferences over well-understood acts). This implies a representation of the following form: $f \succsim g$ iff either i) f statewise dominates g (write $f \geq g$), or ii)

$$\sup\{U(h) : h \in F_{CE}, f \geq h\} \geq \inf\{U(h) : h \in F_{CE}, h \geq g\},$$

where U is a utility representation of the preference over F_{CE} (assumed to be expected utility). I refer to these as *Generalized Simple Bounds* preferences.

Relative to the standard EU representation, the only new parameter is the set F_{CE} of well-understood acts. This is the parameter that defines the relevant notion of act complexity. One can consider a number of specifications for F_{CE}, characterized by different behavioral assumptions.

In Section 2 and Section 3 I characterize notions of complexity which depend on the richness of the coarsest partition with respect to which an act is measurable, which I refer to as *f's partition*. Endow the space of partitions with the usual refinement partial order. I focus on partition-based notions of act complexity, meaning preferences which satisfy the following axiom.

PARTITIONAL COMPLEXITY. If f is well understood and g has a (weakly) coarser partition than f, then g is also well understood.

In other words, Partitional Complexity says that the property of being well understood depends only on an act's partition, and that coarser partitions are easier to understand. Partitional Complexity is a natural axiom for characterizing complexity (see for example Saponara (2020)). In Appendix A I also show that under a set of standard Basic Conditions, this property is equivalent to a set of more primitive axioms (Proposition 11). Thus the results of Appendix A provide a more fundamental basis for the commonly used partition-based notion of complexity. I further motivate these conditions with a model of observational learning in Appendix B.

Characterizations based only on Partitional Complexity are intuitive, but intractable without further assumptions. I therefore study a representation under an additional set of axioms, which imply that an act's complexity is determined by the cardinality of its partition. I refer to these as Simple Bounds preferences. As discussed below, partition size is a common measure of complexity in both the theoretical and empirical literature. This notion captures the idea that the DM has trouble *i*) forming beliefs about many events and *ii*) combining a large number of potential outcomes to understand an act's value. The additional axioms imply the existence of an integer N, identified from choice behavior, such that an act is well understood if and only if it is N-simple, i.e. the cardinality of its partition is less than N (Theorem 2).

I then consider two models of complete preferences derived from Simple Bounds. These models complement the analysis of incomplete preferences by exploring how the DM handles their lack of understanding. Preferences are *Cautious* if the DM evaluates an act f according to its simple lower bound, i.e. the best act that has a certainty equivalent and is dominated statewise by f. Similarly preferences are *Reckless* if acts are evaluated according to their simple upper bound, defined analogously. I axiomatize these two preferences, relate them to the DM's attitude towards ambiguity, and compare them to existing models of ambiguity aversion.

I provide some general comparative statics results which are useful in applications, building on Tian (2015), Tian (2016), and Valenzuela-Stookey (2022). The fact that rich comparative statics can be obtained is an important strength of the model, and the derivation of these results is a key contribution.

Finally, I investigate two applications of the model. First, I study consumption savings problems and equilibrium asset prices. Cautious decision makers will save more, and allocate a greater portion of their savings to a safe rather than a risky asset, compared to fully rational individuals. In equilibrium these biases lead to higher prices for safe relative to risky assets, as observed in the "equity premium puzzle" of Mehra and Prescott (1985). Second, I study the choice of insurance plans. I show that the cautious model rationalizes many "behavioral" phenomena identified in the empirical literature, such as over-sensitivity to deductibles and coverage rates. It also helps explain the correlation, documented by Bhargava et al. (2017), between specific forms of dominated plan choice and both the degree of health risk and level of education. This application provides a good illustration of some of the general comparative statics results presented in Appendix D. In the working paper version, I also examine principal-agent contracting with a complexity constrained agent, identifying general features of optimal contracts.

Related Literature

Most closely related to the current paper are Gul and Pesendorfer (2014) and Echenique et al. (2020). Gul and Pesendorfer (2014) study a model in which there is a σ-algebra of "ideal" events $\mathcal{E}$, which can be thought of as well understood, and acts are bracketed by their upper and lower bounds among those measurable with respect to $\mathcal{E}$. In contrast, axiom C0 below implies that in the Simple Bounds model, all events are well-understood. Complexity here is about aggregation of payoffs across different events, rather than the events themselves. There is no σ-algebra $\mathcal{E}$ for which the set of N-simple acts is the set of $\mathcal{E}$-measurable acts (aside from the trivial case of $N = \infty$). Moreover, the way in which the bracketing acts are used is very different from the current paper: Gul and Pesendorfer (2014) consider an aggregation of the bracketing acts which generates complete preferences, whereas incompleteness is a key feature of the current paper.

The bracketing in the current paper is more in the spirit of Echenique et al. (2022), who study *twofold multiprior preferences* represented by $f \succsim g$ iff $\min_{\mu \in C} E_\mu[u \circ f] \geq \max_{\mu \in D} E_\mu[u \circ g]$. As in the Generalized Simple Bounds model, the decision maker is "twofold conservative", treating f pessimistically and g optimistically. However in Echenique et al. (2022), minimization/maximization occurs over fixed sets of beliefs, rather than bounding acts. In fact, it can be show that if the sets C and D in the twofold multiprior preference representation are allowed to depend on f and g respectively then the resulting class of preferences nests the Simple Bounds. In this sense, the Simple Bounds model extends Echenique et al. (2022).[2]

Other papers have built on the idea that acts with many outcomes may be difficult for a decision maker to evaluate. Neilson (1992) proposes a model of choice under risk in which the decision maker uses a different utility function when computing expectations for lotteries with different support sizes.

Puri (2020) axiomatizes a "Simplicity Representation" of choice under risk, in which a lottery p is evaluated according to $E_p[u(x)] - C(|support(p)|)$, for some increasing function

[2]The original draft of the current paper circulated prior to Echenique et al. (2020), which is the earlier of two papers subsumed by Echenique et al. (2022).

C.[3] While this model also relates support size to complexity, its empirical content is quite different; the Simplicity Representation and Simple Bounds are very far from being "dual" in the sense one might expect at first glance. Most importantly, the Simplicity Representation makes a sharp separation between the values on which a lottery is supported and its complexity, as measured by support size. As a consequence, the Simplicity Representation predicts potentially extreme preference reversals resulting from small changes; arbitrarily small perturbations of a lottery that increase the size of its support can dramatically change its complexity cost. This is not the case in the Simple Bounds model. Unsurprisingly, the Simplicity Representation also predicts violations of first order stochastic dominance, analogous to violations of Monotonicity, which cannot occur in the current setting.

Saponara (2020) axiomatizes a Revealed Reasoning model which is close in spirit to the Cautious model. In this model a decision maker is characterized by a set of partitions $\mathcal{P}$, and evaluates an act f according to the best act that is uniformly below f and measurable with respect to some partition $P \in \mathcal{P}$. As in Generalize Simple Bounds, $\mathcal{P}$ need not be characterized by a fixed number of elements. Instead the axioms impose that $\mathcal{P}$ satisfy a richness condition, which in some cases is in fact incompatible with $\mathcal{P}$ being equal to the set of N-element partitions, for some N. Despite these differences, at a high level, the relationship between the Generalized Simple Bounds and Revealed Reasoning models is similar to that between Bewley preferences and MEU (see Section 3.3), with the Cautious model providing more choice-based structure on the DM's behavior, relative to Revealed Reasoning. The current paper also complements Saponara (2020) by decomposing the Partitional Complexity axiom, which is implicit in that paper.

Ahn and Ergin (2010) also study preferences in which partitions play a central role. In their partition-dependent expected utility (PDEU) representation the decision maker uses a different belief to evaluate acts depending on the partition used to describe the state space. This can lead to preference reversals between acts f and g when different partitions (with respect to which f and g are measurable) are used to describe the state space.

There are formal similarities between this paper and Gilboa, Maccheroni, Marinacci, and Schmeidler (2010) and Lehrer and Teper (2014). As in Lehrer and Teper (2014), I begin by

[3] Related ideas appear in the menu choice literature. Ortoleva (2013) axiomatizes a model of preferences over lotteries of menus in which, similar to the model of Puri (2020), the decision maker attaches a cost to lotteries with more menus in their support.

characterizing preferences on a small set of acts and extend these to a larger subset, although not necessarily to all acts. As in GMMS, given a characterization of incomplete preferences, I show that under additional assumptions a complete preferences relation consistent with it exists, and takes a specific form. In GMMS the incomplete and complete preference relations admit representations à la Bewley (2002) and à la Gilboa and Schmeidler (1989) respectively. The incomplete preferences arising from the partition size notion of complexity do not admit a Bewley representation. Similarly the Cautious completion is not an MEU preference, although it still captures a notion of ambiguity aversion.

One objective of the current paper is to explore how Ellsberg-type phenomenon (Ellsberg (1961)) relate to complexity constraints. As in Segal (1987) and Klibanoff et al. (2005), bets on ambiguous urns are viewed here as a two stage act, where the first stage is subject to uncertainty. That ambiguity may arise from complexity in such an environment is not a new idea. Gilboa and Schmeidler (1989) states "One conceivable explanation of this phenomenon [Ellsberg-type preferences] which we adopt here is as follows: ... the subject has too little information to form a prior. Hence (s)he considers a set of priors as possible." One novelty of the current paper is that it derives ambiguity averse preferences by explicitly characterizing subjective complexity. Bewley (2002) and Gilboa and Schmeidler (1989) relax completeness and independence respectively. I do both, but in a way that is driven by explicit assumptions about perceived complexity.

The paper is organized as follows. Section 1 introduces the setting. Section 2 presents the Generalized Simple Bounds preferences derived from a weak characterization of the well-understood set. Section 3 presents the stronger Simple Bounds characterization. Section 4 discusses complete preferences derived from Simple Bounds. Section 5 presents the applications. The key axiom of Partitional Complexity is further decomposed in Appendix A, and Appendix B explores the procedural learning model which motivates this decomposition. Omitted proofs are in Appendix C, and general comparative statics results in Appendix D.

1 The model

The framework is that of Anscombe and Aumann (1963). The decision maker is characterized by a binary relation $\succsim$ over acts (the preference relation). The preference relation is assumed to be a preorder (i.e. reflexive and transitive), but need not be complete. The strict preference relation and indifference relation are defined as usual.[4] Further notation is as follows:

- Z: the set of *outcomes*.
- L: the set of vN-M *lotteries* (finite support distributions) over Z.
- Ω: the *state space*, endowed with an algebra Σ of *events*.
- F_c: the set of constant acts.
- F: the set of finite valued acts; Σ-measurable $f : \Omega \to L$ such that $|f(\Omega)| < \infty$.

The preference relation is assumed to satisfy the following basic conditions (see e.g. Gilboa et al. (2010)).

Basic Conditions:

PREORDER: $\succsim$ is reflexive and transitive.

MONOTONICITY: For every $f, g \in F$, $f(\omega) \succsim g(\omega)\ \forall\ \omega \in \Omega$ implies $f \succsim g$.

ARCHIMEDEAN CONTINUITY: For all $f, g, h \in F$, the sets $\{\lambda \in [0,1] : \lambda f + (1-\lambda)g \succsim h\}$, $\{\lambda \in [0,1] : h \succsim \lambda f + (1-\lambda)g\}$ are closed in $[0,1]$.

NONTRIVIALITY: There exist $f, g \in F$ such that $f \succ g$.

C-COMPLETENESS: $\succsim$ is complete on F_c

DOUBLE C-INDEPENDENCE: Let c_1, c_2 be constant acts. Then for any act f and $\alpha \in (0,1)$, $f \succsim (\precsim) c_1$ if and only if $\alpha f + (1-\alpha)c_2 \succsim (\precsim) \alpha c_1 + (1-\alpha)c_2$.

The key difference between the Basic Conditions and the corresponding subset of the standard SEU axioms is of course the lack of completeness. Transitivity is preserved as a basic tenet of rational preferences. For a detailed discussion of the relationship between transitive but incomplete preferences and intransitive choice see Mandler (2005). Double C-Independence is a weakening of the common C-Independence assumption, as it imposes that two of the acts involved be constant.

The Basic Conditions imply that there is a (complete) expected utility representation of preferences on constant acts, which are identified with the corresponding lottery. This is

[4] I take as primitive the reflexive and transitive relation, interpreted as weak preferences. An alternative approach, as in Galaabaatar and Karni (2013), would be to take as primitive a transitive and irreflexive strict partial order. I find my approach more convenient because I make use directly of the existence of certainty equivalents and the transitivity of the weak preference relation.

in keeping with the literature on ambiguity, stemming from Gilboa and Schmeidler (1989), which distinguishes between risky lotteries, which the agent has no difficulty evaluating, and uncertain acts, over which choice may be non-EU.[5]

Let v be a representation of preferences over L. I refer to the coarsest partition of Ω with respect to which $v \circ f$ is measurable as *f's partition*. Intuitively, the fact that what matters is the range of $v \circ f$, rather than that of f, reflects the idea that the DM does not find it difficult to evaluate individual consequences, but rather to understand the mapping from states to consequences. Moreover, distinguishing between acts with the same utility image is incompatible with Monotonicity; if $v \circ f = v \circ g$ then Monotonicity implies $f \sim g$. Monotonicity, i.e. respecting statewise dominance, I take as a basic tenet of rationality.

2 Generalized Simple Bounds

I first present a general representation, following the above discussion. The characterization here is straightforward. It becomes more challenging to further refine the representation, as I do in Section 3.

The set of well-understood acts is characterized by Partitional Complexity, plus a richness condition (recall the definitions of "well-understood" and "Partitional Complexity" from pages 2 and 3).

C0. All binary acts are well understood.

Intuitively, C0 says that there are no events which the DM does not understand. This allows us to identify a subjective probability measure on the state space that characterizes the DM's preferences over binary acts. This is in contrast to models such as Epstein and Zhang (2001) and Gul and Pesendorfer (2014), in which there are events that are inherently difficult to understand. In Epstein and Zhang (2001) it is possible for probabilistic sophistication, as defined in Machina and Schmeidler (1992), to be violated whenever an act is not measurable

[5]There is of course empirical evidence that choice over risky lotteries may also be influenced by complexity constraints (see for example Neilson (1992)). More recently, models have been developed which incorporate both ambiguity aversion and non-EU preferences over risky lotteries (see for example Dean and Ortoleva (2017)). The techniques developed in the current paper for studying choice under uncertainty can be applied to risky choice as well. Studying choice behavior when both constraints are present, along the lines of Dean and Ortoleva (2017), is an intriguing direction for future work.

with respect to a set of subjectively unambiguous events. In contrast, C0 embodies the idea that complexity is a property of comparisons between acts, rather than an inherent difficulty with understanding certain events.

Say that an event $E \subset \Omega$ is *non-null* if there exist $c^1, c^2, c^3 \in F_c$ such that $c^2_E c^1 \succ c^3_E c^1$.[6] The set E is *null* otherwise (the empty set is null, by reflexivity of $\succsim$).[7] I write $f \geq_E g$ if $f(\omega) \succsim g(\omega)$ for all $\omega \in E$. If there exists a null set E such that $f \geq_{\Omega \setminus E} g$, write $f \geq^0 g$. For simplicity, I write $\geq$ rather than $\geq_\Omega$ to represent statewise dominance.

Endow the space of partitions of Ω with the refinement partial order: write $\tau'' R \tau'$ if τ'' is a refinement of τ'. A set $\mathcal{T}$ of partitions of Ω is *downward closed* if $\tau'' \in \mathcal{T}$ and $\tau'' R \tau'$ implies $\tau' \in \mathcal{T}$. By Partitional Complexity, the set of well-understood acts is defined by a downward closed set of partitions.

Let $F_{\mathcal{T}}$ be the set of acts measurable with respect to some partition in $\mathcal{T}$. For any $f \in F$, let $\text{gsiminf}_{\mathcal{T},f} := \{h \in F_{\mathcal{T}} : f \geq^0 h, \nexists g \in F_{\mathcal{T}} \text{ s.t. } g \succ h \text{ and } f \geq^0 g\}$ and $\text{gsimsup}_{\mathcal{T},f} := \{h \in F_{\mathcal{T}} : h \geq^0 f, \nexists g \in F_{\mathcal{T}} \text{ s.t. } h \succ g \text{ and } g \geq^0 f\}$. In other words, $\text{gsiminf}_{\mathcal{T},f}$ be the set of most preferred acts among those statewise dominated (up to null sets) by f and measurable with respect to a partition in $\mathcal{T}$. For a lottery l, I write $E_l u = E_l[u(z)]$, i.e. the expected utility given distribution l over outcomes. Although $\text{gsiminf}_{N,f}$ may not be single valued, I abuse notation and write $\int_\Omega E_{\text{gsiminf}_{\mathcal{T},f}(\omega)} u dP(\omega)$.

Definition. *Preference $\succsim$ has a* ***Generalized Simple Bounds*** *representation if there exists a downward directed set $\mathcal{T}$, finitely additive probability P on Σ, and a non-constant function $u : Z \to \mathbb{R}$; such that $gsiminf_{\mathcal{T},f}$ and $gsimsup_{\mathcal{T},f}$ are non-empty for every $f \in F$; and such that for every $f, g \in F$, $f \succsim g$ if and only if at least one of the following holds:*

i. $f \geq g$.

ii. $\int_\Omega E_{gsiminf_{\mathcal{T},f}(\omega)} u \; dP(\omega) \geq \int_\Omega E_{gsimsup_{\mathcal{T},g}(\omega)} u \; dP(\omega)$

Remark. Before characterizing such preferences, it is worth clarifying the condition that $\text{gsiminf}_{\mathcal{T},f}$ and $\text{gsimsup}_{\mathcal{T},f}$ are non-empty for every $f \in F$. For arbitrary $\mathcal{T}, P$, and u, we cannot guarantee that this holds, and verifying this condition may not be easy. Fortunately, in the primary case of interest – the Simple Bounds representation in which $\mathcal{T}$ is the set

[6] Here $c^2_E c^1$ is the binary act which takes value $c^1 \in L$ on E and $c^2 \in L$ on $\Omega \setminus E$.

[7] As shown below, we are be able to identify a probability P which the DM uses to evaluate well-understood acts, and null events will be zero measure events under P.

of partitions defined by an upper bound $N \in \mathcal{Z}$ on the partitions' cardinality – I show that $\text{gsiminf}_{\mathcal{T},f}$ and $\text{gsimsup}_{\mathcal{T},f}$ (which for Simple Bounds preferences we write as $\text{siminf}_{N,f}$ and $\text{simsup}_{N,f}$) exist for any P, u. Thus non-emptiness of these sets imposes not additional restrictions on parameters. Moreover, non-emptiness of $\text{gsiminf}_{\mathcal{T},f}$ and $\text{gsimsup}_{\mathcal{T},f}$ does not appear to impose substantive restrictions on preferences, in the sense that the two additional axioms needed to characterize the representation, S-Independence and Uniform Comparability, have clear behavioral interpretations and are not directly related to the non-emptiness question.

To characterize Generalized Simple Bounds preferences, some modifications to the standard SEU axioms must be made in order to maintain the spirit of the characterization of complexity via the well-understood set of acts. In particular, the usual independence assumption must be modified. This is because the mixture of two acts with partitions in $\mathcal{T}$ need not itself have a partition in $\mathcal{T}$. Thus the standard independence axiom would artificially expand the set of well-understood acts. This is to be avoided, as part of our goal is to identify the well understood set based on axioms which are explicitly about complexity.

The modified axiom below, S-Independence, eliminates this concern. Moreover, it addresses some of the usual critiques of the independence assumption. For one, it applies only when all acts involved are well-understood. Moreover, it only applies to mixtures that do not result in a more complex act.

S-INDEPENDENCE: Let f, g, h be well-understood acts. For any $\alpha \in (0, 1)$, if $\alpha f + (1 - \alpha)h$ and $\alpha g + (1 - \alpha)h$ are both well understood, then $f \succsim g$ if and only if $\alpha f + (1 - \alpha)h \succsim \alpha g + (1 - \alpha)h$.

Finally I make an assumption limiting how comparisons involving complex acts can be made.

UNIFORM COMPARABILITY: For any $f, g \in F$ such that $\neg(f \geq g)$, if $f \succsim g$ then there exist well-understood acts h, k such that $f \geq h \succsim k \geq g$.

Uniform Comparability says that statewise dominance combined with transitivity is the *only* way to compare complex acts. Thus Uniform Comparability implies the minimal extension of preferences beyond well-understood acts.[8]

[8]That is, minimal in the set inclusion sense in the space of monotone preorders.

It is straightforward to show that a Generalized Simple Bounds representation obtains under the above assumptions.

Theorem 1. The following statements are equivalent:

i. $\succsim$ satisfies the Basic Conditions, Partitional Complexity, C0, S-Independence and Uniform Comparability.

ii. $\succsim$ has a Generalized Simple Bounds representation, with parameters $P, u, \mathcal{T}$. Moreover $\mathcal{T}$ is unique, P is unique, u is unique up to positive affine transformations.

Further discussion of the representation is deferred to Section 3, as it is convenient to discuss the general properties of the "bracketing" procedure in the context of the simpler cardinality-based refinement.

3 Simple Bounds

While the Generalized Simple Bounds representation is intuitively appealing, it tells us relatively little about the nature of complexity. We would like to know how additional assumptions on preferences refine the set of well-understood acts. Moreover, for tractability in applications it us useful to have more structure on this set. In this section I explore one such refinement.

3.1 Characterizing complexity

I first provide a tighter characterization of the set of well-understood acts, under two additional assumptions.

Cardinally Conditions. Let τ be f's partition, with typical elements T, T'.

C1. If f is well-understood then for any non-null $T \neq T'$ in τ, and any 2-element partition τ' of $T \cup T'$, acts with partition $\tau' \cup (\tau \setminus \{T, T'\})$ are well understood.

C2. If f is well-understood and E is null then $g_E f \sim f$ for all $g \in F$.

As demonstrated above, the previous imply that there is an expected utility representation of preferences over well-understood acts. In light of this, C2 simply says that the decision maker is not artificially confused by changes to a well-understood act that occur with zero probability. This is primarily a technical condition.

The key axiom is C1. This assumption says that if a partition is well-understood then so are partitions that are close to it, in the sense that they differ only on two cells, and have the same number of elements. Thus it can be viewed as saying that the well-understood property of partitions is robust to local perturbations.

Together, the axioms discussed thus far imply that there is an integer $N \geq 2$, determined endogenously by preferences, such that all and only N-simple acts are well-understood.

Theorem 2. Assume $\succsim$ satisfies the Basic Conditions, Partitional Complexity, C0, C1, and C2. Then there exists an $N \in (\mathbb{N}\backslash\{1\}) \cup \{\infty\}$ such that an act is well-understood iff its partition has at most N non-null elements.

While the restriction that the well-understood property depends only on the number of elements is strong, it is supported by a large body of experimental and observational evidence relating partition size to perceived complexity. Moffatt et al. (2015) and Sonsino et al. (2002) find experimental evidence that subjects undervalue lotteries with larger supports. Bernheim and Sprenger (2019) argue that aversion to large supports helps explain experimental data that is otherwise inconsistent with both expected utility and cumulative prospect theories. Additional experimental evidence, relating support size to ambiguity aversion, is discussed in Section 4.2

While C1 explicitly ties together the "well-understood" property across partitions of the same size, it is not immediate to conclude that only partition cardinality matters. The key step in the proof is a result about moving through the space of partitions of Ω via binary modifications of the form defined in C1. The difficulty comes from the fact that no structure is imposed on the state space. This result may be of independent interest.

Lemma 1. For any two N-element partitions $\tau = \{T_j\}_{j=1}^N, \tau' = \{T_j'\}_{j=1}^N$ of Ω, there is a finite sequence of N-element partitions $\{\tau^i\}_{i=1}^K$, starting with τ and ending with τ', such that between τ^i and τ^{i+1}, $N-2$ of the partition cells remain unchanged.

3.2 Simple Bounds representation

Given Theorem 2, it is straightforward to refine Theorem 1 to incorporate the cardinality-based notion of complexity.

Definition. *For any $f \in F$ and $N \in \mathbb{N}$, denote by* $\boldsymbol{simsup_{N,f}}$ *the set of acts h satisfying: 1) $h \in F_N$, 2) $h \geq^0 f$, and 3) there is no k satisfying 1 and 2 such that $h \succ k$.*[9]

Define $\text{siminf}_{N,f}$ analogously. Note that for any N such that all acts in F_N are well-understood, the DM is indifferent between all acts in $\text{simsup}_{N,f}$ (similarly for $\text{siminf}_{N,f}$). When there is no risk of confusion, I will therefore abuse notation and write as if siminf and simsup are single valued (for example, $f \succsim \text{siminf}_{N,f}$ even when $\text{siminf}_{N,f}$ may contain multiple acts). It is not obvious that such bounds exist, i.e. that $\text{siminf}_{N,f}$ and $\text{simsup}_{N,g}$ are non-empty. In fact, the axioms stated will imply that both are non-empty. Conversely, Proposition 2 shows that existence of these bounds places no additional restrictions on the model parameters.

Definition. *Preference $\succsim$ has a* ***Simple Bounds*** *representation if there exists an integer N, finitely additive probability P on Σ, and a non-constant function $u : Z \to \mathbb{R}$; such that $siminf_{N,f}$ and $simsup_{N,f}$ are non-empty for every $f \in F$; and such that for every $f, g \in F$, $f \succsim g$ if and only if at least one of the following holds:*

i. $f \geq g$.

ii. $\int_\Omega E_{siminf_{N,f}(\omega)} u \; dP(\omega) \geq \int_\Omega E_{simsup_{N,g}(\omega)} u \; dP(\omega)$

I refer to such preferences as Simple Bounds preferences.

Theorem 3. The following statements are equivalent:

i. $\succsim$ satisfies the Basic Conditions, Partitional Complexity, C0, C1, C2, S-Independence and Uniform Comparability.

ii. $\succsim$ has a Simple Bounds representation, with parameters P, u, N. Moreover, N is unique, P is unique, u is unique up to positive affine transformations.

The behavioral implications of Theorem 3 are discussed in Section 3.3, but it is useful to first clarify some technical points.

Condition *i.* in the definition of a Simple Bounds representation requires $f \geq g$, rather than $f \geq^0 g$. This is natural if we think that it is harder to identify null events when

[9] Requiring only $h \geq^0 f$, rather than $h \geq f$, in condition 2 is a technicality. For most applications, we can replace this condition with $h \geq f$, as Proposition 1 shows.

comparing complex acts. Alternatively, we could assume that preferences obey $\geq^0$ dominance, and make the obvious modification to Uniform Comparability, to replace *i.* with $f \geq^0 g$.

For applications, it is generally without loss to assume that $\text{simsup}_{N,f} \geq f \geq \text{siminf}_{N,f}$. Roughly speaking, violations of statewise dominance only occur if f has isolated outliers in the support of P. No continuity is required. Given $f \in F$ and $u : Z \to \mathbb{R}$, define $f_u : \Omega \to \mathbb{R}$ by $f_u := E_{f(\omega)}[u]$.

Proposition 1. $f \geq \mathit{siminf}_{N,f}$ iff $f_u^{-1}(A)$ is non-null for every open neighborhood, A, of $\inf\{f_u(\omega) : \omega \in \Omega\}$. Similarly, $\text{simsup}_{N,f} \geq f$ iff $f_u^{-1}(A)$ is non-null for every open neighborhood of $\sup\{f_u(\omega) : \omega \in \Omega\}$.

The conditions of Proposition 1 are met, for example, if P has full support on $\Omega \subseteq \mathbb{R}$ and $E_f u$ is continuous, or if its partition has no null elements, or it is a convex combination of any acts with these properties. Even if the conditions for statewise dominance fail, it is easy to see where violations will occur. The "Lebesgue approach", discussed below, helps clarify this point.

Theorem 3 includes the conclusion that $\text{siminf}_{N,f}$ and $\text{simsup}_{N,f}$ are non-empty for all f. A natural concern from a modeling perspective is that this may impose constraints on the other parameters of the representation, namely u, P, N. This would be the case if existence failed for some specification of u, P or N. The following result states that this is not the case. Let $B(\Omega)$ be the set of bounded measurable functions on Ω, and $B_N(\Omega)$ be the set of N-simple measurable functions on Ω. Abusing notation, define

$$\text{siminf}_{N,w,P} := \underset{\{b \in B_N(\Omega): b \leq^0 w\}}{\arg\max} \int_\Omega b(\omega) dP(\omega)$$

Proposition 2. For any $w \in B(\Omega)$, $\text{siminf}_{N,w,P}$ and $\text{simsup}_{N,w,P}$ are non-empty for all N, P.

A full discussion of Proposition 2 can be found in Appendix C.4. The proof is instructive, as it makes use of a dual approach to the problem of finding simple upper and lower bounds; rather than look for functions defined by partitions on Ω, I define a dual problem in terms of partitions of $w(\Omega)$. Using this dual "Lebesgue approach" finding the $\text{siminf}_{N,w,p}$ and $\text{simsup}_{N,w,P}$ for an arbitrary act f can be reduced to the problem of finding a the siminf and simsup for the identity function on the unit interval, with a suitably defined distribution $\hat{P}$.

This dual approach is useful for comparative statics (Appendix D). Moreover, it is easy to use this approach to show that finding $\mathrm{simsup}_{N,w,P}$ and $\mathrm{siminf}_{N,w,P}$ functions can be reduced to the computationally easy problem of finding the maximum-weight path in a suitably defined directed acyclic graph. It also implies a convenient structure of the sets $\mathrm{siminf}_{N,w,P}$ and $\mathrm{simsup}_{N,w,p}$.

Lemma 2. $\mathrm{siminf}_{N,w,P}$ and $\mathrm{simsup}_{N,w,P}$ are lattices.

The partial order with respect to which Lemma 2 holds is discussed in Appendix D. For $\Omega = [0,1]$, the result follows from Proposition 17. The extension to more general spaces follows from the dual approach of Appendix C.4.1.

3.3 Theorem 3 discussion

Simple Bounds preferences characterize a decision maker who 1) understands some set of acts well, and 2) uses the acts they understand well to bracket those they do not, allowing them to make some comparisons involving poorly understood acts. The characterization provided by Theorem 2 of which acts are well understood is sharp: the well-understood property is defined by a cut-off in the number of elements of an act's partition. This does not mean, however, that perturbing an act slightly to increase the size of its partition will cause a discontinuous change in choice behavior. If act f' takes values very close to those of f (uniformly over Ω), where the latter is well understood while the former is not, the simple lower and upper bounds of f' will be almost identical to f. Thus the sharp characterization of what is well understood has smoother implications for choice. This is in contrast to models in which complexity enters the representation as an additive cost, such as Puri (2020).

A natural question is how the incomplete preferences characterized above relate to those of Bewley (2002).[10] So called "Bewley preferences" have the following representation: there exists a non-empty, closed, and convex set C^* of probabilities on Σ and a non-constant function $u : Z \to \mathbb{R}$ such that $f \succsim g$ if and only if

$$\int_\Omega E_{f(\omega)} u dp(\omega) \geq \int_\Omega E_{g(\omega)} u dp(\omega) \quad \forall\, p \in C^*.$$

[10]Bewley (2002) in fact gives a representation of the strict preference. The exact unanimity representation discussed here is due to Gilboa et al. (2010).

There is no simple relationship between the two representations. The Bewley preferences satisfy the usual independence axiom. However it is easy to see that the representation in Theorem 3 does not satisfy independence. If incompleteness à la Bewley is interpreted as reflecting complexity then it must be that mixtures do not increase the complexity of a comparison. For example, let the state space be $[0, 1]$ and for any act k identify $u(k(\omega))$ with $k(\omega)$. Let f and g be binary acts, with $f = 0, g = 1$ on $[0, 1/2)$ and $f = 1, g = 0$ on $(1/2, 1]$. In the Bewley model, if f and g are comparable then so are $1/2f + 1/2h$ and $1/2g + 1/2h$ for $h(\omega) = 10\omega^2$. Decision makers may find the later comparison, which involves acts with a greater range of values and larger partitions, more difficult. Relaxing Independence and explicitly modeling complexity allows for a model in which mixtures can increase complexity and lead to incomparability.

4 Extension: Completing preferences

In many settings the decision maker is forced to make a choice. In such cases we would like to be able to make predictions about behavior even when the environment contains pairs of alternatives that are not ranked according to the incomplete preferences above. Following Gilboa, Maccheroni, Marinacci, and Schmeidler (2010), henceforth GMMS, I assume that the decision maker is characterized by a pair of binary relations $(\succsim, \succsim')$, interpreted as *objective rationality* and *subjective rationality* relations respectively.[11] I consider decision makers with objectively rational preferences that can be represented as in Theorem 3, and consider various assumptions on the subjective relation that lead to distinct complete preference relations and representations. This approach allows me to separate the decision makers ability to compare acts from their attitude towards choices between acts that they do not know how to compare. The first point is addressed by Theorem 3. Attitudes towards the unknown can be captured by intuitive axioms. Following GMMS, I first make the natural assumption that the subjective relation never reverses the objective.

CONSISTENCY: $f \succsim g$ implies $f \succsim' g$.

This notion of consistency allows for indifference according to the subjective relation

[11] According to GMMS, a choice is objectively rational if the DM can convince others that they, the DM, is right in making it. It is subjectively rational if the DM cannot be convinced that they are wrong in making it.

between acts that are strictly ranked according to the objective relation. A stronger notion of consistency rules out such differences when the acts in question are well-understood. Below I discuss the reason for imposing Strong Consistency only on well-understood acts (Section 4.4).

STRONG CONSISTENCY FOR SIMPLE ACTS: For any two well-understood acts f, g, $f \succsim g \Leftrightarrow f \succsim' g$.

Finally, I assume a cautious approach to incomparable alternatives.

CAUTION: For all $f \in F$ and $h \in F_c$, if $f \not\succsim h$ then $h \succsim' f$.

Caution is exactly the axiom used by GMMS to derive max-min expected utility (MEU) as the subjective relation given Bewley objective preferences. When the objective relation is Simple Bounds, Caution yields a representation in which acts are evaluated according to their simple lower bounds.

Definition. *Preference $\succsim'$ has a **Cautious** representation if there exists an integer N, probability P on Σ, and a non-constant function $u : Z \to \mathbb{R}$ such that, for every $f, g \in F$ we have $f \succsim' g$ iff $\int_\Omega E_{siminf_{N,f}(\omega)} u \; dP(\omega) \geq \int_\Omega E_{siminf_{N,g}(\omega)} u \; dP(\omega)$.*

Call preferences that admit a Cautious representation Cautious preferences.

Theorem 4. The following statements are equivalent:

i. $\succsim$ satisfies the Basic Conditions, Partitional Complexity, C0-C2, S-Independence and Uniform Comparability; $\succsim'$ satisfies Archimedean Continuity; and $\succsim$ and $\succsim'$ jointly satisfy Consistency, Strong Consistency for Simple Acts and Caution.

ii. $\succsim$ has a Simple bounds representation and $\succsim'$ has a Cautious representation, with common parameters P, u, N. Moreover, P is unique, u is unique up to positive affine transformations, and for all $f \in F$, $\text{siminf}_{N,f}$ and $\text{simsup}_{N,f}$ are non-empty.

Cautious preferences capture a particular attitude towards the unknown. Faced with a difficult choice, the decision maker takes a worst case view of the set of payoffs that they consider reasonable. The properties of Cautious preferences, and their relationship to models of ambiguity aversion, are discussed in detail in Section 4.1. In brief, the Cautious completion is the most ambiguity averse completion of Simple Bounds preferences, in the sense of Ghirardato and Marinacci (2002). Importantly, Cautious preferences do not satisfy the Uncertainty Aversion axiom of Gilboa and Schmeidler (1989), but do satisfy a modified version, N-Ambiguity Aversion, discussed in Hartmann and Kauffeldt (2019). Under some conditions, Cautious preferences are a special case of the Revealed Reasoning model of Saponara (2020). This connection is discussed in detail in the introduction.

4.1 Properties of Caution

Cautious preferences can also be understood through the violations of Independence that they exhibit. A decision maker with Cautious preferences may violate independence when the mixture of two acts leads to an act that is harder to approximate from below by simple acts. In particular, the mixture of two N-simple acts between which the decision maker is indifferent will in general not be N-simple, and may thus be considered inferior to the original acts. Cautious preferences are ambiguity averse in the sense of Ghirardato and Marinacci (2002).[12] In particular, for fixed u and P, decision makers with a higher capacity are less ambiguity averse than those with lower capacity.

Given the discussion of Bewley (2002) above and the formal parallel between this paper and GMMS, it is natural to consider the relationship between Cautious preferences and Gilboa and Schmeidler (1989) MEU preferences. Preferences have an MEU representation if there exists a convex set C of probabilities on Σ and a non-constant function $u : Z \to \mathbb{R}$ such that f is preferred to g if and only if

$$\min_{p \in C} \int_{\Omega} E_{f(\omega)} dp(\omega) \geq \min_{p \in C} \int_{\Omega} E_{g(\omega)} dp(\omega).$$

As GMMS show, such preferences are derived from Bewley preferences through the Consistency and Caution axioms, just as Cautious preferences are derived from Simple Bounds. These representations do not coincide however. This can be seen most easily by considering the Uncertainty Aversion axiom of Gilboa and Schmeidler (1989).

UNCERTAINTY AVERSION: For every $f, g \in F$, if $f \sim g$ then $(1/2)f + (1/2)g \succsim g$.

Uncertainty Aversion captures the notion that mixing acts smooths payoffs, and thus reduces exposure to uncertainty. It is easy to see that this axiom is violated by Cautious

[12] Simple Bounds preferences are also ambiguity averse, if we extend the notion to incomplete preferences.

preferences. As noted above, mixing any two N-simple acts between which the decision maker is indifferent leads to an act which is no better than either of the original acts. This is because the mixed act may not be N-simple, and will thus be approximated from below, whereas the original acts were perfectly understood. More generally, whenever the two acts considered are comonotonic, Cautious preferences will satisfy the reverse Uncertainty Aversion: if $f \sim g$ then $g \succsim (1/2)f + (1/2)g$. Two acts f, g are said to be comonotonic if there is no $\omega, \omega' \in \Omega$ such that $f(\omega) \succ f(\omega')$ and $g(\omega') \succ g(\omega)$. [13] The preference against the mixture will be strict unless there are elements of $\text{siminf}_{N,f}$ and $\text{siminf}_{N,g}$ which share the same partition. The intuition for this reversal is that mixing between comonotonic acts does not smooth payoffs across states in the same way that mixing between non-comonotonic acts can. On the other hand, since the two acts may have had a lot of variations on different regions of the state space, the mixed act will be more difficult to approximate from below than any of the two acts individually. This contrast illustrates that MEU and Cautious preferences capture very different notions of aversion to uncertainty.

Uncertainty Aversion is satisfied by Cautious preferences whenever the mixture of f and g is N-simple, and indeed the inequality may be strict. A modification of Uncertainty Aversion along these lines is studied by Hartmann and Kauffeldt (2019), who state the following axiom

N-AMBIGUITY AVERSION: $f_1, \ldots, f_n \in F$, $\alpha_1, \ldots, \alpha_n \geq 0$, $\sum_{i=1}^n \alpha_i = 1$, $\sum_{i=1}^n \alpha_i f_i = f \in F_N$ such that $f_1 \sim f_2 \cdots \sim f_n$ implies $f \succsim f_1$.

The preferences studied by Hartmann and Kauffeldt (2019) differ from Cautious preferences, primarily in that they satisfy the Comonotonic Independence axiom of Schmeidler (1989).

COMONOTONIC INDEPENDENCE: For all pairwise comonotonic acts f, g, h, and all $\alpha \in (0,1)$, $f \succ g$ implies $\alpha f + (1-\alpha)h \succ \alpha g + (1-\alpha)h$.

[13] This can be seen as follows. For a set $I \subseteq \Omega$ and act f define $\underline{f}(I) = \inf\{u \circ f(\omega) : \omega \in I\}$. Then the value of $\text{siminf}_{N,f}$ is given by

$$\max_{\{I_i\}_{i=1}^N \in T^N(\Omega)} \sum_{i=1}^N P(I_i)\underline{f}(I).$$

When f and g are comonotonic $\underline{f\alpha g}(I) = \alpha \underline{f}(I) + (1-\alpha)\underline{g}(I)$ for all I, where $f\alpha g = \alpha f + (1-\alpha)g$. The claim follows.

Schmeidler's motivation for Comonotonic Independence, as opposed to the usual Independence axiom, is similar to my motivation of S-Independence. Schmeidler (1989) notes that arbitrary mixtures may lead to acts which define a much finer (larger) algebra than the original acts, and thus violations of Independence may occur. However, if "f, g and h are pairwise comonotonic, then the comparison of f to g is not very different from the comparison of $[\alpha f + (1-\alpha)h$ to $\alpha g + (1-\alpha)h]$". In some circumstances however, comonotonicity may be too weak a notion to guarantee that mixtures do not alter the the acts under consideration in ways that lead to violations of independence. Consider the following simple example. Let the state space be the interval $[0,1]$ and identify $u(f(\omega))$ with $f(\omega)$. Let $f = 2$ on $[0, 1/2)$ and $f = 4$ on $(1/2, 1]$. Let $g = 1$ on $[0, 2/3)$ and $g = 6$ on $(2/3, 1]$. Suppose $f \succ g$. The mixture $1/2f + 1/2g$ has both a greater spread between its highest and lowest outcomes and a larger partition than f. Both these changes may lead the DM to favor g. The Simplicity Conditions, along with S-Independence, can be seen as making precise the notion of "not very different". Cautious preferences capture aversion to uncertainty in the sense that the decision maker, faced with an act that she does not fully understand, assigns to the act the minimum value consistent with monotonicity and her preferences over the acts that she understands well. Uncertainty here stems directly from the complexity of the acts under consideration. However, cautious preferences satisfy the following condition, capturing a notion of aversion to comonotonic mixtures.[14]

COMONOTONIC MIXTURE AVERSION: For all comonotonic acts f, g with $f \sim g$ and $\alpha \in (0,1)$, $f \succsim \alpha f + (1-\alpha)g$.

4.2 Empirical evidence

Cautious preferences display features of ambiguity aversion. One of the useful features of this model is that it explicitly separates the decision makers understanding of the environment from her attitude towards the unknown. This novel model of ambiguity aversion can explain some experimental findings at odds with existing models.

[14] In ongoing work, I characterize a broad class of preferences satisfying Comonotonic Mixture Aversion. I show that within this class, cautious preferences are simultaneously *i*) maximally ambiguity averse, in the sense of Ghirardato and Marinacci (2002), *ii*) maximally N-Ambiguity averse, and *iii*) maximally averse to mixtures of N-simple acts. Details available upon request.

Chew et al. (2017) find that subjects are averse to increases in the number of possible compositions (the size of the state space in my framework) of a deck of cards on which bets are made. For example, consider an individual betting on red in a deck of red and black cards, where the number of red cards is known to be between n and $100 - n$. The authors find that subjects aversion to the ambiguous deck, measured by the difference between their certainty equivalents for an ambiguous act and the corresponding compounded lottery, is decreasing in n. The data contradict the predictions of the maxmin expected utility (Gilboa and Schmeidler, 1989) and recursive expected utility (Klibanoff et al., 2005; Seo, 2009), which predict an aversion to increasing the number of possible compositions only in certain cases. Similarly, Viscusi and Magat (1992) document increasing ambiguity aversion as the range of uncertain outcomes increases.

4.3 Alternative approaches

There may be situations in which decision makers take the opposite approach to what they recognize as their limited understanding of the objects of choices.

ABANDON: For all $f \in F$ and $h \in F_c$, if $h \not\succsim f$ then $f \succsim' h$.

Replacing Caution with Abandon yields the obvious converse representation, which I call Reckless preferences, whereby acts are evaluated according to their simsup rather than siminf. Such preferences seem to contradict common assumptions about risk and ambiguity aversion. They have a flavor of, but are distinct from, optimism or overconfidence. The decision maker does not behave exactly as if she thought high payoff states are more likely to occur, but rather as if the underlying act is as good as possible without violating her limited understanding of the situation. Such a model may be useful for understanding the behavior of decision makers who seem to favor nebulous prospects with high potential over those that are well-understood. I will sometimes use the notation $U(f)$ or $U(f, P)$ to denote perceived utility. If the agent is cautious $U(f, P) = E_P[E_{\text{siminf}_{N,f}(\omega)} u]$, and if the agent is reckless $U(f, P) = E_P[E_{\text{simsup}_{N,f}(\omega)} u]$.

An alternative approach to incompleteness would be to proceed as in Bewley (2002) and assume that there is a default option and the DM satisfies an inertia condition, whereby an alternative act is chosen over the default if and only if it is preferred, according to the underlying incomplete preference. While a full coverage of inertia is beyond the scope of this paper, the results presented here are helpful for understanding such a model. What matters for choice under inertia is the simsup of the default act and the siminf of the alternative. The behavior of these objects is studied in the comparative statics and applications sections.[15]

4.4 A note on Consistency

Given the interpretation of the subjective preference as an extension of the incomplete objective preference, it would seem natural to impose Strong Consistency everywhere: for all $f, g \in F$, $f \succsim g$ implies $f \succsim' g$, and $f \succ g$ implies $f \succ' g$. Consistency allows the decision maker to be indifferent under $\succsim'$ between some acts that are strictly ranked under $\succsim$.[16] It turns out that Strong Consistency everywhere is incompatible with Archimendean Continuity of $\succsim'$. This conflict between Caution, Strong Consistency, and Continuity is not unique to this setting. GMMS face the same trade-off obtaining MEU preferences as the Cautious completion of Bewley preference (see Section 4.1), and the incompatibility holds for a broad class of incomplete preference models.[17] Rather than relaxing consistency to retain continuity, as in Theorem 4, one can drop the continuity requirement for $\succsim'$ and impose Strong Continuity everywhere. This yields "lexicographic" preferences: acts are first ranked by statewise dominance, and then according to their siminf if they are not ranked according to statewise dominance. The difference between the lexicographic model and that of Theorem 4 is one of continuity; the two models have nearly identical properties and predictions.

5 Applications

The subsequent applications focus on Cautious and Reckless preferences. However the results are relevant for Simple Bounds preferences more generally, as these are also characterized by the siminf and simsup. For example in the insurance choice application of Section 5.2 under inertia, an individual chooses some plan f over the default act f if and only if $\text{siminf}_{N,f}$ is preferred to $\text{simsup}_{N,g}$. Thus the results on insurance valuation for a Cautious individual in Section 5.2 also answer the question of how an individual with inertia will value new

[15]The earlier working paper version of this paper includes further discussion of inertia.

[16]In fact, under Strong Consistency for Simple Acts, this occurs for acts f, g only if a) either $f \geq g$ or $g \geq f$, and b) $\text{siminf}_{N,f} \cap \text{siminf}_{N,g} \neq \varnothing$.

[17]As a technical note, Strong Consistency for Simple Acts renders imposition of further basic conditions on $\succsim'$ redundant, as these are inherited from $\succsim$ on the set of well-understood acts.

offerings relative to their default plan.

The applications of this section build heavily on comparative statics results from Valenzuela-Stookey (2022). Further discussion of how to solve for $\text{siminf}_{N,f}$ and $\text{simsup}_{N,f}$ in applications, both analytically and computationally, can also be found in Valenzuela-Stookey (2022).

5.1 Consumption-Savings

In this section I study consumption-savings decisions in a two period model. The following observation is useful for understanding the role of complexity in this setting. Let f be an increasing and real-valued act on $\Omega = \mathbb{R}$ (recall that using the Lebesgue approach discussed in Section 3.2 and Appendix C.4.1, it is essentially without loss of generality to consider such acts). Let P be the cdf of the agent's belief and u the DM's utility function. Fix a selection s from $\text{siminf}_{N,f}$, and recall that since f is monotone, s is defined by a vector of cut-offs $(t_i)_{i=0}^N$. Define the *lower-perceived distribution* P_N^s to be the distribution which places mass at the cut-off points so as to satisfy $E_{P_N^s}[f] = E_P[\text{siminf}_{N,f}]$. Similarly for $s \in \text{simsup}_{N,f}$ defined by cut-offs $(t_i)_{i=0}^N$, the *upper-perceived distribution* $\tilde{P}_N^s$ is defined as the distribution which places mass at the cut-off points so as to satisfy $E_{\tilde{P}_N^s}[f] = E_P[\text{simsup}_{N,f}]$.[18]

Observation 1. *Let $f \circ u$ be increasing. Then $\tilde{P}_N^{s'} \succsim_{FOSD} P \succsim_{FOSD} P_N$ for any $s \in \text{siminf}_{N,f}$ and $s' \in \text{simsup}_{N,f}$. If, moreover, $u \circ f$ is integrable then $E_P[u \circ f] = \lim_{M\to\infty} E_{P_M}[u \circ f] = E_{\tilde{P}_M}[u \circ f]$.*

Observation 1 can be extended to more general environments using the Lebesgue approach of section C.4.1. This observation facilitates comparisons between constrained and unconstrained agents.

5.1.1 Portfolio choice with a safe and risky asset

Suppose the DM is cautious and faces the problem of allocating wealth between consumption, purchase of a risk-free asset and purchase of a risky asset. I compare choices under a capacity constraint to those of a fully rational agent ($N = \infty$). Let R_b be the certain return on the

[18] So if P has no mass points, P_N^s places mass of $P(t_{i+1}) - P(t_i)$ on t_i, for $i = 0, \dots, N-1$ and is zero elsewhere, and $\tilde{P}_N^s$ places mass of $P(t_i') - P(t_{i-1}')$ on $i = 1, \dots, N$. In other words, Observation 1 is simply the observation that for an increasing act, the siminf (simsup) is a step function, where the steps touch $u \circ f$ at their leftmost (rightmost) endpoints.

risk free asset, and R_s the uncertain return on the risky asset with cdf P and pdf p on the interval $[\underline{R}_s, \bar{R}_s]$. The DM with capacity N solves $\max_{b,s} u(w-b-s) + \beta V^N(R_b b + R_s s)$, where

$$V^N(R_b b + R_s s) = \max_{\{\{t_i\}_{i=0}^N \in I\}} \sum_{i=1}^{N} [P(t_i) - P(t_{i-1})] u(R_b b + t_{i-1} s), \tag{1}$$

and I is the set of interval partitions of $[\underline{R}_s, \bar{R}_s]$. Note that $\lim_{N\to\infty} V^N(R_b b + R_s s) = \int u(R_b b + rs) dP(r)$. As Observation 1 makes clear, for any *fixed* portfolio choice the unconstrained DM perceives a FOSD shift of the returns perceived by the constrained DM. Of course, as the constrained DM changes their portfolio allocation the perceived distribution of R_s induced by siminf will also change, so this is not the same problem as comparing portfolio choice under two different fixed yield distributions.

Proposition 3. Let u be CRRA with a coefficient of relative risk aversion greater than or equal to 1.[19] Then

i. For a given level of aggregate savings $x > 0$, a Cautious DM allocates a greater portion of savings to the safe asset than an unconstrained DM.

ii. A Cautious DM saves more overall than an unconstrained DM.

Part i of Proposition 3 is perhaps unsurprising; the Cautious DM takes a pessimistic view of uncertainty, and so invests less in the risky asset. Part ii is more striking; despite perceiving savings as less profitable, the Cautious DM saves more. Loosely, the reason is that what matters for the savings decision is not the level of returns, but the marginal return on an additional dollar saved, and the assumptions on u imply that the Cautious DM, who anticipates being poorer in the future than the unconstrained DM, wants to save more. The proof is slightly more nuanced, as one must also account for the fact that the perceived distribution (or equivalently, the siminf) depends on the level of savings.

5.1.2 Equilibrium Asset Prices

In the previous section, we studied the savings problem of a single agent. Consider now an economy with a representative agent who chooses between a safe and risky asset. In order

[19] Part (i) holds under the weaker condition that $u'''(z)u'(z)/u''(z)^2 \geq 2$ for all z. I state it for CRRA for simplicity. The assumption that the risk aversion coefficient is greater than 1 is standard. Part (ii) also holds under weaker conditions, CRRA just makes it easy to show that (9) in the proof holds. As the proof illustrates, neither of these conditions is necessary for the result

for the market to clear the price of the risky asset and the return on the safe asset must adjust so that the representative agent demands zero of both assets. Returns are as in the previous section. Normalize the price of the safe asset to 1 and let the price of the risky asset be given by ρ. I assume that the agent has an endowment of w in each period. Let $P^{b,s}$ be the perceived distribution given investment b in the safe asset and s in the risky asset. The constrained DM solves

$$\max_{b,s} u(w - b - \rho s) + \beta \int u(w + bR_b + sr) dP^{b,s}(r).$$

For the unconstrained DM the problem is identical, except that $P^{b,s}$ is replaced with P. It is intuitive that a cautious (reckless) constrained DM should be biased towards the safe (risky) asset, relative to the unconstrained DM. The constraint simply coarsens the DM's understanding of the stochastic payoff of the risky asset, which leads the cautious agent to undervalue it and reckless agent to overvalue. In fact, using the condition that the assets that must be in zero net supply in equilibrium, we can extend this intuition to make comparisons between intermediate capacity levels.

Proposition 4. The equilibrium risky asset price with a Cautious (Reckless) representative agent is increasing (decreasing) in the agent's capacity.

When the agent is cautious, Proposition 4 gives an explanation for the so called "equity premium puzzle".

5.2 Insurance valuation

I now compare the salience of various features of insurance plans for a complexity-constrained versus a fully rational ($N = \infty$) individual. I focus on a Cautious individual, but the analysis is equally relevant for an individual who faces inertia with respect to some fixed default plan, as discussed above.

I take as given a class of basic insurance contracts that are characterized by a premium p, i.e. price for the plan; a deductible d, below which the individual bears all losses; a coverage rate c specifying the fraction of losses above the deductible covered by the plan; an out-of-pocket expenditure cap m, which is the maximum amount that an individual will have to pay, excluding the premium. I do not provide a foundation for the use of these piece-wise linear contracts, but they are by far the most common form of insurance contract. These contracts have the feature that the individual's ex-post wealth is decreasing in the realized loss in $[0, \bar{\omega}]$. Recall that in this case siminf and simsup will be measurable with respect to N-element interval partitions, which I will describe by the cut-off states $t_0 = 0, t_1 \ldots, t_N = \bar{\omega}$.

5.2.1 The setting

Consider an individual facing a bounded loss distribution on $\Omega = [0, \bar{\omega}]$ with absolutely continuous CDF P. The loss distribution will be fixed throughout. Let w be the individual's endowment wealth. In autarky the ex-post wealth is $w - \omega$, where ω is the realized loss.

I do not consider here the choice of the contract by the insurer, only the valuation by the individual. Say that a capacity-constrained individual *over-reacts* (*under-reacts*) to a change from one plan to another if the magnitude of the difference between their values for the two plans is greater (less) than that of a fully rational individual.

(a) No cap

(b) Cap

Figure 1: siminf with and without out-of-pocket maximum

5.2.2 Salient plan features

I will consider changes in the deductible and in the coverage rate, meaning the percentage c of losses above the deductible covered by the contract ($c = 1$ in the full insurance case). Proposition 5 states that a cautious agent over-reacts to changes in both parameters when there is no out-of-pocket expenditure cap. On the other hand, when there is a binding cap the agent's valuation is unaffected by changes in the coverage rate, provided the coverage rate is sufficiently high.

Proposition 5. Consider marginal changes to a contract with no out-of-pocket expenditure cap. A **cautious** agent:

1. over-reacts to changes in the deductible.
2. over-reacts to changes in the coverage rate.

Proposition 5 is silent on which of the distortions is relatively larger. If we focus on marginal changes to a baseline plan with full insurance above the deductible, then the over-reaction to the coverage rate will be relatively larger than that to the deductible if the decision maker believes that high losses are sufficiently likely. In this case the highest cut-off defining the siminf will be close to d (but always strictly below d, by Lemma 4), and so the response to a deductible change will be close to that of a fully rational DM. However caution will still cause the DM to drastically over-react to reductions in coverage rate. Similarly, when N is sufficiently high the highest cut-off defining the siminf will be close to d, so again the over-reaction to the coverage rate will be the larger of the two when N is high enough.

Restricting attention to contracts with near full insurance above the deductible, it is easy to show a stronger comparative statics result.

Proposition 6. Consider a contract with deductible d and $c = 1$. For both the deductible and the coverage rate the magnitude of the response of a cautious agent to marginal changes is decreasing in N.

Many of the settings in which biases towards low deductibles are observed involve full coverage above the deductible. In this case cautious individuals are always biased towards low deductible plans, and this bias is decreasing in their capacity.

Proposition 7. With $c = 1$, the amount a cautious individual is willing to pay to lower the deductible by a given amount is decreasing in their capacity.

Finally, I consider how a capacity constraint affects the individual's willingness to pay to decrease the out of pocket maximum.

Proposition 8. The amount a cautious agent would be willing to pay to decrease the out-of-pocket maximum is decreasing in their capacity.

The finding that lower capacity individuals place a greater value on the out of pocket maximum may help explain the widely documented bias towards full insurance, for example by Shapira and Venezia (2008). Full insurance plans are easy to understand, since the loss is independent of the state. In general, constrained individuals will overvalue full insurance plans relative to those for which the realized losses are a more complicated function of the state.

5.2.3 Dominated choices

Bhargava et al. (2017) find that many individuals choose dominated plans, and that the propensity to do so is positively correlated with both high expected losses and earnings levels.[20] The cautious model can provide an explanation for these observations.[21] The result depends on the nature of plan dominance. These results also demonstrate the usefulness of the comparative statics properties discussed in the appendix. Bhargava et al. (2017) observe individuals choosing low deductible plans even when the increase in the premium relative to a high deductible plan (holding other plan features constant) is greater than the maximum possible savings from the lower deductible. Figure 2 illustrates a situation in which the maximum possible savings from the low deductible, high premium plan is equal to the increase in the premium. Call this as a *weakly dominated low-deductible plan.*

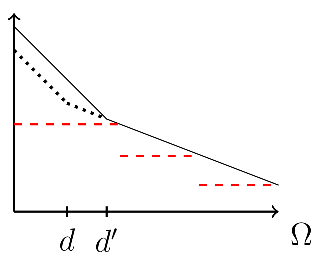

Figure 2: Indifference with a weakly dominated low-deductible plan

In Figure 2 the solid black line is the high deductible, low premium plan, and the black

[20]Interestingly, the authors find that the number of parameters needed to describe a plan, an alternative measure of complexity encountered in the literature, does not predict dominated choices. This suggest that the partitional notion of complexity may be more relevant in this setting.

[21]My model does not predict strictly dominated choices, but allows for indifference between pairs of plans ordered by weak dominance, even for full support beliefs. If dominated choices are related to act complexity then the empirical evidence is informative about what types of acts are perceived to be complex. Additional factors, such as difficulty understanding how plans map states to payments, may combine with complexity considerations to produce strictly dominated choices.

dashed line is a low premium, high deductible plan. Both plans have the same coverage rate. The red dotted line is the siminf, which in this case is the same for both plans. This occurs whenever the lowest cut-off for the high deductible plan is above its deductible. Let $l(\omega|d,c)$ be the amount paid by the consumer when the loss is ω given a contract with deductible d and coverage rate c (this is stated formally in Lemma 5 in the Online Appendix). More interesting than the fact that dominated plans can be chosen is are the conditions which are conducive to such mistakes. Bhargava et al. (2017) observe that dominated choices are correlated with both expected losses and earnings levels. Proposition 9 predicts the former. To the extent that earnings are correlated with the capacity to evaluate acts, Proposition 10 predicts that low earners will be more prone to mistaken indifference. The following corollary establishes a single-crossing property of dominated choice which implies that individuals who are more pessimistic about their losses are more prone to make dominated choices (in both cases, all other model parameters are held constant).

Proposition 9. If an individual with belief P chooses a weakly dominated low-deductible plan then so does one with belief P' if $P' \succeq_{MLR} P$.

Proposition 10. If an individual with capacity N' chooses a weakly dominated low-deductible plan then so is one with capacity N if $N' \geq N$.

5.2.4 Discussion

This section relates to the large empirical literature documenting behavioral phenomena in choices of complicated contracts. The model is able to explain many observed choice patterns which differ from the predictions of standard EU theory. Abaluck and Gruber (2011) find that consumers underweight out-of-pocket spending relative to premiums. Moreover, cautious agents will respond more than the fully rational to changes in the coverage rate when there is an out-of-pocket expenditure cap and the coverage rate is high (see Proposition 5). Cutler and Zeckhauser (2004) document a bias towards low deductible plans. This is consistent with a cautious decision maker (see Proposition 5). The results of this section highlight the dependence of the qualitative nature of comparisons to the rational model on details of the insurance plans under consideration. Such variation forms an interesting basis for further empirical work, and provides a more nuanced perspective on "behavioral biases".

5.2.5 MEU in insurance

The insurance plans considered here all induce payoffs that are monotone in the the state. If the set C of beliefs over which the DM with MEU preferences minimizes contains a first-order stochastically dominant belief then this will be the minimizing belief regardless of plan characteristics. In this sense the individual will behave exactly like an expected utility maximizer.

In some cases MEU and cautious preferences predict similar behavior. For example, in the MEU framework we can consider individuals who minimize over smaller sets of beliefs, which is analogous to the idea of increasing capacity. Consider comparing the relative value of two plans which induce ex-post wealth functions f_1, f_2 for two individuals, one of whom has a larger set of beliefs. If the larger set of beliefs contains an upper bound in the first-order stochastic dominance (FOSD) order, and the payoff difference $f_1 - f_2$ is increasing (decreasing) in the state then the individual with the larger belief set will value f_1 more (less) relative to f_2 than the individual with the smaller set. For example, the individual with a larger set of beliefs will value reductions in the out-of-pocket maximum more than an individual with a smaller set of beliefs.

Appendix

A Axiomatizing Partitional Complexity

In this section I discuss an axiomatization of Partitional Complexity. A number of the axioms state that "there exists" an act with certain properties. While axioms with existential qualifiers are not testable, they are intended here to illustrate more precisely the behavioral content of Partitional Complexity. Moreover, these axioms relate directly to the procedural learning model of Appendix B, which further motivates the axioms.

A bet on an event E is a binary act that has a strictly better consequence on E than on the complement of E. Given any acts f, g, let $f_E g$ be the binary act which is equal to f on E and g elsewhere. The following axioms characterize the set of acts that are well-understood by the decision maker.

Simplicity Conditions. Let τ be f's partition, with typical elements T, T'.

A0. For any event $E \in \Sigma$ there exists a bet on E that is well-understood.

A1. If f is well-understood then for any partition τ' coarser than τ, there exists a well-understood act with partition τ'.

A2. Let f be a well-understood act. Let g measurable with respect to f's partition, and for which there exists a constant act $a \in f(\Omega)$ such that for all $\omega \in \Omega$ either $g(\omega) = f(\omega)$ or $g(\omega) = a$. Then if g is well-understood, $\alpha f + (1-\alpha)g$ also is well-understood for all $\alpha \in (0, 1)$.

The Simplicity Conditions (and A3 below) are further motivated by a procedural model of learning, presented in Appendix B.

Roughly, A1 says that acts with coarser partitions are easier to understand. For example, if we merge two cells in the partition of f and replace the outcome on the merged cell with the conditional expectation of f over the two cells, we might expect that the act has become easier to evaluate. (In the procedural model of choice of Appendix B I show that it is precisely this type of conditional-mean modification which makes an act easier to compare.)

A2 is similar to the Certainty Independence assumption of Gilboa and Schmeidler (1989). Certainty Independence would imply that if f is well-understood and c is a constant act, then $\alpha f + (1-\alpha)c$ would also be well-understood. A2 generalizes this conclusion to mixtures with non-constant acts g that are well-understood, provided g has the same partition and range as f, and is constant on the set of states on which it differs from f. Standard arguments in favor of Certainty Independence, that mixing with a constant act does not help to hedge against uncertainty, therefore apply.

Under the Basic Conditions, the Simplicity Conditions characterize the set of well-understood acts. This characterization can be separated into two parts. First, A0-A2 imply that i) if any act with a given partition is well-understood then so are all such acts, and ii) with the refinement order on the space of partitions, the set of partitions for which measurable acts are well-understood is a lower-set.[22]

Proposition 11. Under the Basic Conditions, the following are equivalent.

1. Preferences satisfy A0-A2.

2. Preferences satisfy Partitional Complexity.

Given Proposition 11, we can also relax the Cardinality Conditions, again with an eye towards the procedural model of Appendix B.

A3. If f is well-understood then for any non-null $T \neq T'$ in τ, and any 2-element partition τ' of $T \cup T'$, there exists a well-understood act with partition $\tau' \cup (\tau \setminus \{T, T'\})$.

By Proposition 11, replacing C1 with A3 does not affect the characterization of the well-understood set.

B Learning motivation for the simplicity conditions

I present here a simple procedural model of a decision maker using data to inform their choice between uncertain acts. The purpose of this section is twofold. First, it demonstrates that the Simplicity Conditions are satisfied under natural assumptions on the learning model. Second, it formalizes the intuition that acts with coarser partitions are easier to understand.

The setting is that of a standard frequentist inference problem. A decision maker is endowed with a dataset χ of K i.i.d. draws from an unknown distribution P.[23] They must

[22]Recall the distinction between an act "with partition τ" and a τ-measurable act. The former means τ is the *coarsest* partition w.r.t which the act is measurable.

[23]This sampling procedure can be interpreted literally as sampling from an unknown distribution. Alternatively, we can think of it as a reduced form model of contemplation in which the DM accesses a latent belief P, similar in spirit to drift-diffusion models of cognition.

compare an arbitrary simple real valued act f and constant act c. These can be thought of as the utility images of acts with outcomes in an arbitrary space. The DM uses their data to estimate the expected value of f, as well as the risk arising from sampling uncertainty. Let the empirical distribution of the sample be $\hat{P}_\chi$. The empirical expectation of f is

$$\hat{\mathrm{E}}_\chi[f] = \sum_{i=1}^{N} f(T_i) \frac{1}{K} \sum_{x \in \chi} \mathbb{1}\{x \in T_i\}. \tag{2}$$

The error due to sampling uncertainty is $\varepsilon_f(\chi) := \hat{\mathrm{E}}_\chi[f] - \mathrm{E}[f]$. Denote the true distribution of $\varepsilon_f(\chi)$ across different samples by G_f. G_f is unknown to the DM, as it depends on the unknown distribution P. Instead, the DM uses an estimate $\hat{G}_f$ of G_f when making decisions. I assume that the DM uses the bootstrap to estimate $\hat{G}_f$.[24] To complete the description of the environment, assume that the DM uses a decision rule that maps $\hat{\mathrm{E}}_\chi(f)$ and $\hat{G}$ to a ranking between f and c. I also allow the DM to declare that f and c are un-rankable. This is a standard problem of statistical inference. I will show that any protocol that is consistent, in a weak sense, with second order stochastic dominance rankings of the error distributions will satisfy A1. I will show further that reasonable decision rules in this environment satisfy all of the Simplicity Conditions.

First, I will discuss how A1 is related to second order stochastic dominance rankings of the error distributions. Let $\tau = \{T_i\}_{i=1}^N$ be a act f's partition, and let $\tau' = \{T', T_3, \ldots, T_N\}$ where $T' = T_1 \cup T_2$. It seems intuitive that learning about acts measurable with respect to τ' will be easier than learning about those measurable with respect to τ, since the set of events to which the DM must assign probabilities is strictly smaller (in the inclusion order). Proposition 12 formalizes the sense in which this is true. Of course, the difficulty of a comparison depends not only on the partitions of the acts involved, but also on their values. Let $\tilde{f}$ be an act such that $\tilde{f}(x) = \mathrm{E}[f|T']$ for $x \in T'$, and $\tilde{f}(x) = f(x)$ otherwise.

[24]The bootstrap procedure for estimating G_f is as follows. Draw a sample $\bar{\chi}$ of size K from the empirical distribution (i.e. sample the data with replacement) and calculate $\bar{\varepsilon}_f(\bar{\chi}) := \hat{E}_{\bar{\chi}}[f] - \hat{E}_\chi[f]$. Do this repeatedly and use the resulting empirical distribution of $\bar{\varepsilon}_f(\bar{\chi})$ as the estimate of G_f. The idea is to treat the empirical distribution $\hat{P}_\chi$ as if it were the true distribution, for the purposes of estimating G_f. Since $n^{-1/2}\varepsilon_f$ is an asymptotically pivotal statistic, the bootstrap estimator provides an asymptotic refinement of the normal approximation, and therefore performs better in finite samples (see Horowitz (2001)). Thus it is reasonable for the DM to use the bootstrap to calculate $\hat{G}_f$.

For any act h, let $\varepsilon_h(\chi) = \hat{\mathrm{E}}_\chi[h] - \mathrm{E}[h]$, and let G_h be the distribution of ε_h. Say that distribution F *strictly second-order stochastically dominates* distribution H ($F >_{SOSD} H$) if $\int u(x)dF(x) \geq \int u(x)dH(x)$ for all concave u, with strict inequality if u is strictly concave.

Proposition 12. $G_{\tilde{f}}$ strictly second-order stochastically dominates G_f.

The intuition for this result is straightforward. Consider the errors made in the estimation of $\mathrm{E}[f]$ versus $\mathrm{E}[\tilde{f}]$. Notice that for all datasets χ, $\hat{\mathrm{E}}_\chi[\tilde{f}|T'] = \mathrm{E}[\tilde{f}|T']$. The randomness of $\hat{\mathrm{E}}_\chi[f|T']$ simply adds noise to the distribution of errors from the estimation of $\mathrm{E}[f]$, relative to those of $\mathrm{E}[\tilde{f}]$. A similar conclusion holds for $\bar{f}$, where $\bar{f} = \hat{\mathrm{E}}_\chi[f|T']$ on T', and $\bar{f} = f$ elsewhere.

Corollary 1. $\hat{G}_{\bar{f}} >_{SOSD} \hat{G}_f$

This follows immediately from Proposition 12, since the bootstrap estimator treats the empirical distribution as if it were the true distribution. We return now to the DM's problem of choosing between the act f and a constant act c. A natural assumption about decisions in this framework is that greater uncertainty about $\mathrm{E}[f]$ makes it harder for the DM to compare f to constant acts (comparable in this setting meaning that the DM is willing to state a preference for one of the two acts).

CONFOUNDING SAMPLING UNCERTAINTY: For any acts f, f' with $\hat{\mathrm{E}}_\chi[f] = \hat{\mathrm{E}}_\chi[f']$ and $\hat{G}_{f'} >_{SOSD} \hat{G}_f$ and any constant act c, if f is comparable to c then so is f'.

Proposition 13. Any decision rule satisfying Confounding Sampling Uncertainty will satisfy A1

Proposition 13 is an immediate implication of Corollary 1. As a concrete example of a protocol satisfying Confounding Sampling Uncertainty, consider the following. Given data set χ the DM concludes that $f \succsim c$ if and only if $\mathrm{E}_{\hat{G}_f}\left[\phi\left(\hat{\mathrm{E}}[f]\right)\right] \geq \phi(c) - k$ for some increasing, strictly concave function ϕ and constant $k > 0$. Moreover $c \succsim f$ if $c \geq \hat{\mathrm{E}}_\chi[f]$. I will refer to this as the *smooth sampling uncertainty* model. It is similar in spirit to the smooth ambiguity model of Klibanoff et al. (2005), where here the higher order uncertainty derives from sampling uncertainty, rather than subjective ambiguity. Additionally, the smooth sampling uncertainty model allows the DM to express incomplete preferences.

Clearly, f will have a certainty equivalent only if $\hat{G}_f$ is not too dispersed. In particular, f will have a certainty equivalent c if and only if i) $\hat{\mathrm{E}}_{\mathcal{X}}[f] = c$, and ii) $\mathrm{E}_{\hat{G}_f}\left[\phi\left(\hat{\mathrm{E}}[f]\right)\right] \geq \phi(c) - k$. Corollary 1 implies that the smooth sampling uncertainty model satisfies Confounding Sampling Uncertainty since ϕ is concave. In fact, the model satisfies all of the Simplicity Conditions, plus A3 and C4.

Proposition 14. Smooth sampling uncertainty satisfies the Simplicity Conditions, A3, and C4.

Smooth sampling uncertainty is by no means the only model for which the Simplicity Conditions will be satisfied. As the proof of Proposition 14 shows, A1 and A3 are very close from a learning perspective. Given Proposition 13, A3 is satisfied so long as the decision rule is suitably continuous in the estimated error distribution $\hat{G}$.

C Appendix: Omitted Proofs

C.1 Proof of Theorem 3 and Theorem 1

I prove Theorem 1, Theorem 3 is an immediate corollary.

By Partitional Complexity, we know that $F_{\mathcal{T}} \subseteq F_{CE}$ for some downward direct set of partitions $\mathcal{T}$ ($F_{\mathcal{T}}$ be the set of acts measurable with respect to some partition in $\mathcal{T}$). By C0, we can choose such $\mathcal{T}$ which contains all binary partitions. The first part of the proof shows that preferences on $F_{\mathcal{T}}$ have an expected utility representation under the Basic Conditions and S-Independence. Then Uniform Comparability implies the desired representation.

Part 1. The restriction of $\succsim$ on L satisfies the von Neumann-Morgernstern axioms, and so is represented by $v(l) = E_l[u]$. Without loss of generality assume that $v(L) \supset [-1, 1]$. Fix a partition $\tau = \{T_i\}_{i=1}^N \in \mathcal{T}$. The restriction of $\succsim$ to acts measurable with respect to τ satisfies the standard SEU axioms (see for example Gilboa and Schmeidler (1989)). Therefore there exists a probability measure P^τ on τ such that for all τ-measurable acts f, g, $f \succsim g$ iff $E_{P^\tau}[v \circ f] \geq E_{P^\tau}[v \circ g]$.

It remains to show that there exists a probability P on Σ such that $P^\tau(A) = P(A)$ for all $\tau \in \mathcal{T}$ and $A \in \Sigma$. For any $A \in \Sigma$ and any $\tau', \tau'' \in F_{\mathcal{T}}$ such that $A \in \tau'$, $A \in \tau''$, it must be that $P^{\tau'} = P^{\tau''}$, since the certainty equivalent for the act $\mathbb{1}_A$ must be the same regardless of which of $P^{\tau'}, P^{\tau''}$ is used to represent preferences. Thus $P(A) = \{P^{\tau}(A) : \tau \in \mathcal{T}\}$ is a well defined function. P is non-negative since each P^{τ} is.

To show additivity of P, it suffices to consider acts in F_2. I show now how additivity of P is implied by S-Independence. To see this, it is sufficient to consider 2-element partitions and $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Let $\tau^1 = \{\{\omega_1\}, \{\omega_2, \omega_3\}\}$, $\tau^2 = \{\{\omega_1, \omega_3\}, \{\omega_2\}\}$, and $\tau^3 = \{\{\omega_1, \omega_2\}, \{\omega_3\}\}$. Let $\mathbb{1}_A$ be an act such that $v \circ f(\omega) = 1$ for $\omega \in A$, and 0 otherwise. Then $P^{\tau^i}(\omega_i) = E_{P^{\tau^i}}[\mathbb{1}_{\{\omega_i\}}]$ for $i \in \{1,2,3\}$. We wish to show that P defined as $P(\omega_i) = P^{\tau^i}(\omega_i)$ for $i \in \{1,2,3\}$ is a well-defined probability on Ω that is consistent with P^{τ^1}, P^{τ^2}, and P^{τ^3}. To see this, let $f_1, f_3 \in F_2$ be acts such that $v \circ f_1(\{\omega_2, \omega_3\}) = 1$ and $v \circ f_1(\omega_1) = -1$; and $v \circ f_3(\{\omega_1, \omega_2\}) = 1$ and $v \circ f_3(\omega_3) = -1$. Then $\frac{1}{2}f_1 + \frac{1}{2}f_3 = \mathbb{1}_{\omega_2}$, so $E_{P^{\tau^2}}[\frac{1}{2}f_1 + \frac{1}{2}f_3] = P^{\tau^2}(\omega_2)$. But by S-Independence we also know that $\frac{1}{2}f_1 + \frac{1}{2}f_3 \sim \frac{1}{2}U(f_1) + \frac{1}{2}U(f_3)$, where $U(f_i)$ is a constant act such that $v(U(f_i)) = E_{P^{\tau^i}}[f_i]$. Then, we also have, with the usual abuse of notation, that $\frac{1}{2}f_1 + \frac{1}{2}f_3 \sim \frac{1}{2}(P^{\tau^1}(\{\omega_2, \omega_3\}) - P^{\tau^1}(\omega_1) + P^{\tau^3}(\{\omega_1, \omega_2\}) - P^{\tau^3}(\omega_3))$. Thus $\frac{1}{2}(P^{\tau^1}(\{\omega_2, \omega_3\}) - P^{\tau^1}(\omega_1) + P^{\tau^3}(\{\omega_1, \omega_2\}) - P^{\tau^3}(\omega_3)) = P^{\tau^2}(\omega_2)$. Using this equality it is easy to see that P is well defined and consistent; for example $P(\omega_2) = P^{\tau^2}(\omega_2) = 1 - P^{\tau^1}(\omega_1) - P^{\tau^3}(\omega_3) = 1 - P(\{\omega_1, \omega_3\})$.

By exactly the same argument, we can show that P as defined above is consistent with $P^{\tau'}$ and $P^{\tau''}$ for any 2-element partitions τ'', τ'. This gives the desired representation of $\succsim$ for N-simple acts.

Part 2. That Uniform Comparability implies the representation in Theorem 1 is immediate; when $\text{gsiminf}_{\mathcal{T},f}$ and $\text{gsimsup}_{\mathcal{T},g}$ both exist with $\text{gsiminf}_{\mathcal{T},f} \succsim \text{gsimsup}_{\mathcal{T},g}$, they play the role of h and k in the axiom respectively.

I now show that $\text{gsiminf}_{\mathcal{T},f}$ must be non-empty for all f. The proof for $\text{gsimsup}_{\mathcal{T},f}$ is analogous. Suppose $\text{gsiminf}_{\mathcal{T},f}$ is empty. Since f is bounded, there exist constant acts $\bar{f}, \underline{f} \in F_c$ such that $\bar{f} \succsim f \succsim \underline{f}$. Let

$$s = \sup_{h \in \{h \in F_{\mathcal{T}} :\ f \geq^0 h\}} \int_\Omega E_h u \; dP(\omega)$$

By Uniform Comparability, there exists an $h \in \{h \in F_{\mathcal{T}} :\ f \geq^0 h\}$ such that $h \succsim \underline{f}$. Choose λ^* such that $\lambda^* \int_\Omega E_{\bar{f}} u dP(\omega) + (1 - \lambda^*) \int_\Omega E_{\underline{f}} u dP(\omega) = s$. Since, by hypothesis, there are no $h \in F_{\mathcal{T}}$ with $f \geq^0 h$ that have expected utility equal to s, Uniform Comparability implies

$\neg(f \succsim \lambda^* \bar{f} + (1-\lambda^*)\underline{f})$. But this violates Archimedian Continuity.

C.2 Theorem 2

Theorem 2 is immediate from Lemma 1, which I prove here.

Proof. of Lemma 1 Let $\tau^i = \{T^i_j\}_{j=1}^N$ be the last partition in the sequence constructed thus far (we begin with $\tau^1 = \tau$). I first show how to construct τ^{i+1}. Assume $\tau^i \neq \tau'$, otherwise we are done.

Step 1.

Check if there exists $T^i_1 \in \tau^i$ and a collection $\{T'_j\}_{j=1}^J \subset \tau'$ such that $\cup_{j=1}^J T'_j = T^i_1$, for $J \geq 2$. If not, proceed to Step 2. Otherwise, find $T^i_2 \in \tau^i$ such that $T^i_2 \notin \tau'$. Such a T^i_2 must exists, since both τ^i and τ' have N elements. Define $\tau^{i+1} := \{T'_1, (T^i_1 \cup T^i_2) \setminus T'_1, T^i_3, \ldots, T^i_N\}$.

Step 2.

In this case, there exists $T' \in \tau'$ and $T^i_1, T^i_2 \in \tau^i$ such that *a*) $T' \cap T^i_1$ and $T' \cap T^i_2$ are both non-empty, and *b*) $T^i_1 \setminus T' \neq \varnothing$. That *a*) holds simply follows from the fact that $\tau^i \neq \tau'$. If *b*) fails, then there must exist $T^i_1 \in \tau^i$ and a collection $\{T'_j\}_{j=1}^J \subset \tau'$ satisfying the conditions for Step 1, so we would not have arrived at Step 2.

To generate τ^{i+1}, fix all elements of τ^i other than T^i_1, T^i_2. Let $r_1 = T^i_1 \cap T'$ and let $r_2 = T^i_2 \cap T'$. By condition *a*) above, r_1 and r_2 are non-empty. By condition *b*), $r_1 \cup r_2 \neq T_1 \cup T_2$. Then define $\tau^{i+1} := \{r_1 \cup r_2, (T^i_1 \cup T^i_2) \setminus (r_1 \cup r_2), T^i_3, \ldots, T^i_N\}$.

As long as $\tau^i \neq \tau'$ the algorithm above delivers $\tau^{i+1} \neq \tau^i$. Moreover, τ^k is a coarsening of $\tau \vee \tau'$ for all k. Since there are at most N^2 elements of $\tau \vee \tau'$, the set of possible coarsenings is finite. Thus the algorithm eventually delivers τ' as long as there are no cycles. To see that cycles cannot occur, first notice that if $T \in \tau^i$ and $T \in \tau'$ then $T \in \tau^{i+k}$ for all $k \geq 1$. Therefore the algorithm can only execute Step 2 finitely many times along any sequence, since each time it does so it delivers τ^{i+1} which has one more cells in common with τ' than did τ^i. Between the steps at which Step 1 is executed there can be no cycles; once r and r' are merged into the same cell in a Step 2 execution, they are never separated into different cells. $\square$

C.3 Proof of Theorem 4

Take $f \notin F_{CE}$. I first show that $f \sim' \text{siminf}_{N,f}$. Suppose $\exists\, c \in F_c$ such that $c \succ f$ (otherwise the claim is trivial). By consistency $f \succsim' \text{siminf}_{N,f}$. Suppose $\text{siminf}_{N,f} \not\succsim' f$. By continuity there exist constant acts $c_1 \succ c_2$ and $\lambda^* \in [0,1)$ such that $\lambda^* c_1 + (1-\lambda^*)c_2 \sim \text{siminf}_{N,f}$. Then for all $\lambda > \lambda^*$ it follows from Theorem 3 that $f \not\succsim \lambda c_1 + (1-\lambda)c_2$. Caution then implies that $\lambda c_1 + (1-\lambda)c_2 \succsim' f$. By continuity of $\succsim'$, $\lambda^* c_1 + (1-\lambda^*)c_2 \succsim' f$. The claim follows. Strong consistency for simple acts implies that the restrictions of $\succsim$ and $\succsim'$ to F_{CE} have the same representation. The theorem follows.

C.4 Existence of simsup and siminf

Rather than look for functions defined by partitions on Ω, I define a dual problem in terms of partitions of $w(\Omega)$. I show that $\text{simsup}_{N,w,P}$ and $\text{siminf}_{N,w,P}$ can be mapped to increasing functions on $w(\Omega)$, and then exploit this monotonicity and fact that $w(\Omega)$ is a bounded interval of $\mathbb{R}$. Put another way, Ω inherits both an order and a topology from $\mathbb{R}$ and the measurable function w, which greatly simplifies the problem of finding simple bounds.

C.4.1 "Lebesgue Approach"

Let $\tau(\Omega)$ be the set of all partitions of Ω, and $\tau^N(\Omega)$ the set of N-element partitions. When Ω is a partially ordered set, say that $\tau = \{T_i\}_{i=1}^N$ is an *interval partition* if T_i is an interval for all i.[25]

Fix $w \in B(\Omega)$. For any $h \in B_N(\Omega)$ and let $\tau_h = \{T_i\}_{i=1}^N$ be h's partition. Define $T_i' := \{r \in w(\Omega) : w(\omega) = r \text{ for some } \omega \in T_i\}$. Define $\tau_h' := \{T_i'\}_{i=1}^N$ as the cover of $w(\Omega)$ *induced* by h. Say that h *induces an interval partition* of $w(\Omega)$ if τ_h' is a partition of $w(\Omega)$ and T_i' is an interval for all i.

Let Q be the law of w, defined by $Q(A) = P(w^{-1}(A))$ for any Borel set A. Let S be Q's support. Since w is measurable, Q is a Borel measure on $\mathbb{R}$, and hence a Radon measure. Thus $Q(A) = 0$ for any $A \in \Omega \setminus S$ (see Parthasarathy (2005), ch 2). The idea behind the proof is to look at N-simple functions on S, rather than on Ω. To do this, I need to show that it is possible to move between $B_N(\Omega)$ and $B_N(S)$. I will focus on the existence of

[25] By an interval in a partially ordered space $(J, \geq)$ I mean a set $I \subseteq J$ such that for all $x, y \in I$ and all $z \in J$ such that $x \geq z \geq y$, $z \in I$. I do not define an interval to be closed, as is sometimes done.

$\text{siminf}_{N,w,P}$, as the argument for simsup is exactly symmetric.

The following lemma shows that $\text{siminf}_{N,w,P}$, if it exists, will live in the subset of $B_N(\Omega)$ that induce an interval partitions of S. The lemma has a symmetric counterpart for $\text{simsup}_{N,w,P}$.

Lemma 3. For any $w \in B(\Omega)$, let $h \in B_N(\Omega)$ be such that $w \geq^0 h$. If h does not induce an interval partition of S then there exists an N-simple function $\hat{h}$ that does, and such that $w \geq^0 \hat{h}$ and $E_P[\hat{h}] \geq E_P[h]$.

Proof. Let $A \subseteq \Omega$ be the set of states ω such that $h(\omega) > w(\omega)$. We can restrict attention to functions h such that $P(A) = 0$, as this is a requirement for $\text{siminf}_{N,w,P}$. To begin, assume that $A \cap w^{-1}(S) = \varnothing$, so $w \geq_{w^{-1}(S)} h$ (i.e. $w(\omega) \geq h(\omega)$ for all ω such that $w(\omega) \in S$).

Suppose that τ'_h is not a partition of S. For any $r \in S$ and any T_i, $T_j \in \tau_h$ such that $T_i \cap w^{-1}(r) \neq \varnothing$ and $T_j \cap w^{-1}(r) \neq \varnothing$, define $\hat{h}$ as $\hat{h}(\omega) = \max\{h(T_i), h(T_j)\}\ \forall\, \omega \in w^{-1}(r)$, and $\hat{h} = h$ elsewhere. Then $E_P[\hat{h}] \geq E_P[h]$, and $\tau'_{\hat{h}}$ will be a partition of $w(\Omega)$. For any such r, $w \geq_{w^{-1}(S)} h$ implies $r \geq \max\{h(T_i), h(T_j)\}$. Therefore $w \geq_{w^{-1}(S)} \hat{h}$, and since $P(A) = 0$, $w \geq^0 \hat{h}$

If there are elements of τ'_h that are not intervals then there exist states $\omega_1, \omega_2 \in \Omega$ with $w(\omega_1) < w(\omega_2)$ and $h(\omega_1) > h(\omega_2)$. Then define $\hat{h}$ such that $\hat{h}(\omega_2) = \hat{h}(\omega_1) = h(\omega_1)$, and $h = \hat{h}$ elsewhere. Clearly $E_P[\hat{h}] \geq E_P[h]$ and $w \geq^0 \hat{h}$.

Now, I want to show that it is without loss to assume $A \cap w^{-1}(S) = \varnothing$. For any $\omega \in A$, if there exists $r \in h(\Omega \setminus A)$ such that $w(\omega) \geq r$ then we can replace $h(\omega)$ with r without altering the value of h, or the fact that it is N-simple. Assume therefore that $w(\omega) < \min h(\Omega \setminus A)$ for all $\omega \in A$. Since $w \geq_{\Omega \setminus A} h$, this implies that $w(\omega) < \inf w(\Omega \setminus A)$ for all $\omega \in A$. But then $w(A) \cap S = \varnothing$, or equivalently $A \cap w^{-1}(S) = \varnothing$. $\square$

Corollary 2. When $\text{siminf}_{N,w,P}$ and $\text{simsup}_{N,w,P}$ are non-empty, they contain functions that induce interval partitions of S.

For any function $\tilde{h} \in B_N(w(\Omega))$ we can define a function $h \in B_N(\Omega)$ by $h(\omega) = \tilde{h}(w(\omega))$. Moreover, $E_Q[\tilde{h}] = E_P[h]$ by the definition of Q. Lemma 3 and Corollary 2 are useful because they allow us to do the converse: given a function $h \in B_N(\Omega)$ that induces a partition of $w(\Omega)$ we can define a function $\tilde{h} \in B(w(\Omega))$ by $\tilde{h}(r) = h(w^{-1}(r))$. By $h(w^{-1}(r))$ I mean the value taken by h for all $\omega \in w^{-1}(r)$. For $\tilde{h}$ to be well defined it is therefore necessary that h

induce a partition of $w(\Omega)$. Interval partitions are easy to work with, as we will see, since they can be described by a vector of N cut-offs.

C.4.2 Existence proof

Proposition 15. Suppose Ω is a closed interval in $\mathbb{R}$, P has full support, and $w \in B(\Omega)$ is continuous, and increasing. Then $\text{siminf}_{N,w,P}$ is non-empty.

Proof. Since P has full support and w is continuous, it is without loss to assume that $w \geq \text{siminf}_{N,w,P}$. To find $\text{siminf}_{N,w,P}$, it will be sufficient to restrict attention to increasing functions with interval partitions. This follows since for any non-interval partition T there exist states $\omega_1, \omega_2 \in \Omega$ satisfying $w(\omega_1) \leq w(\omega_2)$ and $h_T(\omega_1) > h_T(\omega_2)$. Then adding ω_2 to the partition cell that contains ω_1 leads to a new N-simple function h' with $h' \leq w$ and $E_P[h'] \geq E_P[h_\tau]$. When working with interval partitions label the cells so that $T_i > T_{i-1}$ (in the obvious order). Moreover, since w is increasing we can assume that each interval in τ contains its left endpoint (since it is always optimal to assign a state to the higher interval). Any such partition can by fully described by a vector $z_\tau = (t_1, \ldots, t_{N-1})$ of $N-1$ cutoffs. Given a cutoff vector z, let h_z be the function constructed as above using the partition described by z. The set of such cutoff vectors is compact. Moreover $z \mapsto E_P[h_z]$ is upper semi-continuous; this follows from right-continuity of the CDF and continuity of w. □

Proof. (Proposition 2). By Corollary 2, it is without loss to look $\text{siminf}_{N,w,P}$ in the subset of $B_N(\Omega)$ that induce interval partitions of S. It is therefore without loss to look for functions $\tilde{h} \in B_N(S)$ that have interval partitions, i.e. we solve $\max_{\hat{h} \in B_N(S): w \geq_S \hat{h}} E_Q[\hat{h}]$. This problem has a solution by Proposition 15. Let $\tilde{h}$ be the solution, and define $h = \tilde{h} \circ w$ on S. Let $h = c$ on $\Omega \setminus S$, for some $c \in h(S)$. Then h will be N-simple and satisfy $w \geq^0 h$. Moreover, Corollary 2 implies that $h \in \text{siminf}_{N,w,P}$. □

C.5 Proof of Proposition 1

Proof. I give the proof for siminf; simsup is symmetric. If $f_u^{-1}(A)$ is non-null for every open neighborhood of $\inf E_f u(\Omega)$ then there must be some ω such that $\inf E_f u(\Omega) \geq E_{\text{siminf}_{N,f}} u(\omega)$. If there was a violation of statewise dominance for some ω' then it could be removed by specifying $E_{\text{siminf}_{N,f}} u(\omega') = E_{\text{siminf}_{N,f}} u(\omega)$. If the condition doesn't hold then $\exists\, \omega$ such that it is strictly sub-optimal to have $E_f u(\omega) \geq E_{\text{siminf}_{N,f}} u(\omega)$. □

C.6 Proposition 11

Proof. Suppose an act f, with partition $\{T_i\}_{i=1}^N$, has a certainty equivalent. I wish to show that any other act g with the same partition also has a certainty equivalent. Throughout the proof, label the partition so that $g(T_{i+1}) \succ g(T_i)$ for all i. For notational simplicity, I will identify each act f with its utility image $v \circ f$. Finally, assume that there exists a constant act $\bar{c}, \underline{c}$ such that $\bar{c} \succ g(\omega) \succ \underline{c}$ for all ω (in the end we will establish existence of a certainty equivalent when such a $\bar{c}, \underline{c}$ do not exist). The proof will proceed by induction on N. The induction hypothesis for each $K < N$ is that all acts measurable with respect to a K-element coarsening of $\{T_i\}_{i=1}^N$ have certainty equivalents. Note that, by the Basic Conditions, preferences on F_c have an expected utility representation.

Before proceeding to the induction proof, I show the following claim.

Claim 1: if $\lambda g + (1-\lambda)c_1 \sim c_2$ for $c_1, c_2 \in F_c$ with $\bar{c} \succsim c_1 \succsim \underline{c}$ and $\lambda \in (0,1)$, then g has a certainty equivalent. The proof of Claim 1 is as follows. By Double C-Independence, it suffices to show that there exists $c_3 \in F_c$ such that $\lambda c_3 + (1-\lambda)c_1 \sim c_2$. If $c_1 \sim c_2$ then we are done. Suppose $c_1 \succ c_2$. By Monotonicity and $c_1 \succsim \underline{c}$, $c_2 \succ \underline{c}$. Then there exists $\alpha \in (0,1)$ such that $\alpha \underline{c} + (1-\alpha)c_1 \sim c_2$. Let $c_3 = \frac{\alpha}{\lambda}\underline{c} + \frac{\lambda-\alpha}{\lambda}c_1$. Since $g(\omega) \succ \underline{c}$, and given the expected utility representation on F_c, Monotonicity implies that $\lambda > \alpha$, so c_3 is well defined. If $c_2 \succ c_1$ replace $\underline{c}$ with $\bar{c}$.

Now for the induction proof. I first show that all binary acts have certainty equivalents. For any event E, let f be a bet on E that has a certainty equivalent (which exists by A0), and g be another arbitrary bet on E. Let $E^c = \Omega \setminus E$. There are a few cases to consider. Suppose $f(E) \succ g(E) \succ g(E^c) \succ f(E^c)$. Then, using the expected utility representation on F_c, $\exists\, \lambda \in (0,1)$ and $\hat{c} \in F_c$ such that $\lambda f(E^c) + (1-\lambda)\hat{c} \sim g(E^c)$ and $\lambda f(E) + (1-\lambda)\hat{c} \sim g(E)$ ($\lambda = (u(g(E)) - u(g(E^c)))/(u(f(E)) - u(f(E^c)))$). By Double C-Independence $\lambda f + (1-\lambda)\hat{c}$ is well-understood, and since this act is payoff equivalent to g, g is as well. Suppose instead that $g(E) \succ f(E) \succ f(E^c) \succ g(E^c)$. Then, as before, there exist $\lambda, \hat{c}$ such that $\lambda g(E) + (1-\lambda)\hat{c} \sim f(E)$ and $\lambda g(E^c) + (1-\lambda)\hat{c} \sim f(E^c)$, so $\lambda g + (1-\lambda)\hat{c}$ has a certainty equivalent. Then g has a certainty equivalent as well, by Double C-Independence.

Suppose $g(E) \succsim f(E) \succsim g(E^c) \succsim f(E^c)$. Then $\exists\, \lambda_1, \lambda_2 \in [0,1)$ such that $\lambda_1 \bar{c} + (1-\lambda_1)f(E) \sim g(E)$, and $\lambda_2 \bar{c} + (1-\lambda_2)f(E^c) \sim g(E^c)$. Let $\lambda = \min\{\lambda_1, \lambda_2\}$, and suppose WLOG that this is equal to λ_1. By Double C-Independence, $f' := \lambda\bar{c} + (1-\lambda)f$ has a certainty equivalent. Note $f'(E) \sim g(E)$. By continuity $\exists\ \alpha \in [0,1]$ such that $\alpha f'(E) + (1-\alpha)f'(E^c) \sim g(E^c)$. Then $\alpha f'(E) + (1-\alpha)f'$ is payoff equivalent to g and has a certainty equivalent by Double C-Independence, so g has a certainty equivalent. The remaining cases are analogous. Since by A0 every event has a well understood bet, we can conclude that any binary act has a certainty equivalent. This is the first step in the induction.

Now suppose that f's partition has $N > 2$ elements, and f has a certainty equivalent. Then by A1 and the induction hypothesis, all acts g' measurable with respect to coarser partitions than f, and satisfying $\bar{c} \succ g'(\omega) \succ \underline{c}$ for all ω, have certainty equivalents. I now show that it is without loss to assume that f is increasing with respect to the same order as g, i.e. $f(T_{i+1}) \succ f(T_i)$ for all i. To see this, it suffices to show that if $f(T_a) \succ f(T_b) \succ f(T_c)$ then there is an act f' with the same partition that has a certainty equivalent and such that $f'(T_a) \succ f'(T_c) \succ f'(T_b)$ (a symmetric argument shows that there is f'' such that $f''(T_b) \succ f''(T_c) \succ f''(T_a)$). Define h by $h(T_a) = h(T_c) = f(T_a)$ and $h = f$ on $\Omega \setminus (T_a \cup T_c)$. Then h has a certainty equivalent by the induction hypothesis. Moreover, by continuity there exists $\alpha \in (0,1)$ such that $\alpha h(T_c) + (1-\alpha)f(T_c) \succ f(T_b)$. Moreover, by A2, $f' = \alpha h + (1-\alpha)f$ has a certainty equivalent, as desired. So from now on, assume WLOG that f and g are comonotone.

By Claim 1, it is without loss to consider g such that $f(T_N) \succ g(T_N)$ and $g(T_1) \succ f(T_1)$. Then there exists $\lambda \in (0,1)$ such that $\lambda f(T_N) + (1-\lambda)f(T_1) \sim g(T_N)$. Define h_N by $h_N = f(T_1)$ on T_N and $h_N = f$ elsewhere. Then h_N has a certainty equivalent by the induction hypothesis. Let $f'_N = \lambda f + (1-\lambda)h_N$. Then f'_N has a certainty equivalent by A2. Moreover, by Monotonicity, f_N defined as $f_N = g(T_N)$ on T_N and $f_N = f'_N$ elsewhere also has a certainty equivalent. Then there also exists $c_{N-1} \in \{f_N(T_N), f_N(T_1)\}$ and $\lambda' \in (0,1)$ such that $\lambda' f_N(T_{N-1}) + (1-\lambda')c_{N-1} \sim g(T_{N-1})$. Define h_{N-1} by $h_{N-1} = c_{N-1}$ on T_{N-1} and $h_{N-1} = f_N$ elsewhere. Then by the induction hypothesis h_{N-1} has a certainty equivalent, and by A2 $f'_{N-1} = \lambda' f_N + (1-\lambda')h_{N-1}$ also has a certainty equivalent. As before, define f_{N-1} by $f_{N-1} = g(T_{N-1})$ on T_{N-1} and $f_{N-1} = f'_{N-1}$ elsewhere. Then f_{N-1} also has a certainty equivalent. Proceeding in this way, we arrive at an act $f_1 = g$ which has a certainty equivalent, as desired.

I now need to address the assumption that there exist constant acts $\underline{c}, \bar{c}$ such that $\bar{c} \succ g(\omega) \succ \underline{c}$ for all ω. Suppose that we have established existence of a certainty equivalent for all acts f, measurable with respect to a given partition, that satisfy this interiority assumption. Suppose g is such that $g(T_N) \succsim c \succsim g(T_1)\ \forall\ c \in F_c$ (recall the ordering of T_i). Let $g' = \frac{1}{2}g + \frac{1}{2}\left(\frac{1}{2}g(T_N) + \frac{1}{2}g(T_1)\right)$. Then g' satisfies the interiority assumption, and so has a certainty equivalent by hypothesis, which can be written as $\lambda g(T_1) + (1-\lambda)g(T_N)$, where $\lambda \in \left[\frac{1}{4}, \frac{3}{4}\right]$ by Monotonicity. The existence of a certainty equivalent for g will follow from Double C-Independence if we can show that there exists a $\kappa \in [0,1]$ such that $\frac{1}{2}(\kappa g(T_N) + (1-\kappa)g(T_1)) + \frac{1}{2}\left(\frac{1}{2}g(T_N) + \frac{1}{2}g(T_1)\right) \sim \lambda g(T_1) + (1-\lambda)g(T_N)$. Equating coefficients, this holds for $\kappa = 2\lambda - \frac{1}{2}$, which is well defined since $\lambda \in \left[\frac{1}{4}, \frac{3}{4}\right]$. □

C.7 Proof of Proposition 12

Proof. For any act g with partition τ and any set A define the random variable $\hat{\mathrm{E}}_\chi[g|A]$ as follows:

$$\hat{\mathrm{E}}_\chi[g|A] = \begin{cases} \frac{1}{\hat{P}_\chi(A)} \sum_{i=1}^N f(T_i) \frac{1}{K} \sum_{x\in\chi} \mathbb{1}\{x \in T_i \cap A\} & \text{if } \hat{P}_\chi(A) > 0 \\ \mathrm{E}[g|A] & \text{otherwise} \end{cases}$$

This definition is necessary since we will be dealing with finite samples, so $\hat{P}_\chi(A) = 0$ with positive probability even when $P(a) > 0$. Let f_b be the act equal to b on T' and f elsewhere. In what follows $T'^c = \Omega - T'$. Notice that $\varepsilon_f(\chi)$ can be written as

$$\begin{aligned} \varepsilon_f(\chi) &= \hat{\mathrm{E}}_\chi[f|T'](1 - \hat{P}_\chi(T'^c)) + \hat{\mathrm{E}}_\chi[f|T'^c]\hat{P}_\chi(T'^c) - E[f] \\ &= \hat{\mathrm{E}}_\chi[f|T'] - \mathrm{E}[f|T'] + \left(\hat{\mathrm{E}}_\chi[f|T'^c] - \hat{\mathrm{E}}_\chi[f|T']\right)\hat{P}_\chi(T'^c) \\ &\quad - \left(\mathrm{E}[f|T'^c] - \mathrm{E}[f|T']\right)P(T'^c) \end{aligned}$$

Similarly $\varepsilon_{f_b}(\chi) = \left(\hat{\mathrm{E}}_\chi[f|T'^c] - b\right)\hat{P}_\chi(T'^c) - \left(\mathrm{E}[f|T'^c] - b\right)P(T'^c)$. Thus

$$\varepsilon_f(\chi) = \varepsilon_{f_b}(\chi) + \underbrace{\hat{\mathrm{E}}_\chi[f|T'] - \mathrm{E}[f|T'] + \left(b - \hat{\mathrm{E}}_\chi[f|T']\right)\hat{P}_\chi(T'^c) + \left(\mathrm{E}[f|T'] - b\right)P(T'^c)}_{\xi(\chi)}$$

The weak inequality in the definition of SOSD will follow by Jensen's inequality if I can show that $\mathrm{E}[\xi|\{\chi : \varepsilon_{f_b}(\chi) = m\}] = 0$ for all m in the range of ε_{f_b}, where the expectation is taken with respect to the measure on datasets induced by P and the i.i.d. sampling procedure. The strict inequality will follow since the distribution of ξ is non-degenerate.

Notice that, because sampling is i.i.d., $\mathrm{E}\left[\hat{\mathrm{E}}_\chi[f|T'] \;\middle|\; \hat{\mathrm{E}}_\chi[f|T'^c], \hat{P}_\chi(T'^c)\right] = \mathrm{E}[f|T']$, where we use here the specification of $\hat{\mathrm{E}}_\chi[f|T'] = \mathrm{E}[f|T']$ when $\hat{P}_\chi(T'^c) = 1$. Independence between $\hat{\mathrm{E}}_\chi[f|T'], \hat{\mathrm{E}}_\chi[f|T'^c]$, and $\hat{P}_\chi(T'^c)$ does not hold for finite samples due to integer restrictions. Fortunately all that we need is the stated conditional mean independence condition. Given this condition

$$\begin{aligned}
\mathrm{E}\left[\xi|\varepsilon_{f_b}\right] &= \mathrm{E}\left[\mathrm{E}\left[\xi|\hat{P}_\chi(T'^c)\right]|\varepsilon_{f_b}\right] \\
&= \mathrm{E}\left[(b - \mathrm{E}[f|T'])(\hat{P}_\chi(T'^c) - P(T'^c))\middle|\varepsilon_{f_b}\right]
\end{aligned}$$

where the first equality is just the law of iterated expectations and the second follows from conditional mean independence. Clearly if $b = E[f|T']$ the expectation is zero, so we are done. □

C.8 Proof of Proposition 14

Proof. A1 has already been shown. A2 follows immediately from Jensen's inequality. A3 follows from the fact that ϕ is strictly concave and $\hat{G}_{\bar{f}} >_{SOSD} \hat{G}_f$; for any 2-element partition τ'' of $T_1 \cup T_2$ we can choose a binary act b on τ'' that is arbitrarily close to a constant and such that $\hat{E}_\chi[b|\tau''] = \hat{E}_\chi[f|\tau'']$. So b can be chosen such that the act f'' that is equal to b on τ'' and f elsewhere will satisfy $E_{\hat{G}_{f''}}\left[\phi\left(\hat{E}[f'']\right)\right] > E_{\hat{G}_f}\left[\phi\left(\hat{E}[f]\right)\right]$. A0 is satisfied because for any event A we can choose a bet on A that is arbitrarily close to constant. C2 holds since null events are those for which there is not data, and when using the bootstrap the value on such sets is irrelevant. □

C.9 Proofs for the applications

C.10 Proof of Proposition 3

Proof. Proposition 3(i). In what follows N and x, the level of savings, will be fixed. Assume without loss of generality that the siminf is unique for any α (the same proof applies for any selection). Let P^α be the perceived distribution (induced by the cutoffs in (1)) when the proportion of savings allocated to the risky asset is α. For a fixed level of savings x the allocation problem becomes.

$$\max_{\alpha\in[0,1]} \int u((1-\alpha)xR_b + \alpha xr)dP^\alpha(r). \tag{3}$$

Notice that for the unconstrained DM, for whom P^α is replaced with P, the objective is concave in α. This means that the derivative of the objective crosses zero (at most once) from above. The result will follow if I can show that at any α the derivative of the objective in (3) with respect to α is greater for the unconstrained DM than for the constrained one. Recall that $\text{siminf}_{N,f}$ (equivalently P^α) is defined by a maximization problem, which satisfies the conditions of the envelope theorem (Milgrom and Segal, 2002). Then, by the envelope theorem, the derivative with respect to α of the objective function in (3) is given by

$$\int u'((1-\alpha)xR_b + \alpha xr)(xr - xR_b)dP^\alpha(r).$$

Recall that $P \succsim_{FOSD} P^\alpha$ for all α. It is therefore sufficient to show that $u'((1-\alpha)xR_b + \alpha xr)(xr - xR_b)$ is increasing in r. To do this I show that for any $b, s \geq 0$

$$\frac{d}{dr}[u'(R_b b + rs)R_b] \leq \frac{d}{dr}[u'(R_b b + rs)r],$$

or equivalently

$$(r - R_b)s \leq -\frac{u'(R_b b + rs)}{u''(R_b b + rs)} \tag{4}$$

Notice that by assumption $R_b \in (\underline{R}_s, \bar{R}_s)$, so that the left hand side of (4) is negative for r low enough, while the the right hand side is always strictly positive. So it is sufficient to show that the derivative with respect to r of the LHS of (4) is less than that of the RHS for all r, i.e.

$$\frac{d}{dr}[(r - R_b)s] \leq \frac{d}{dr}\left[-\frac{u'(R_b b + rs)}{u''(R_b b + rs)}\right] \iff \frac{u'''(R_b b + rs)u'(R_b b + rs)}{u''(R_b b + rs)^2} \geq 2$$

This holds for all r for CRRA utility $u(x) = x^{1-\gamma}/(1-\gamma)$ when $\gamma \geq 1$. □

Proof. Proposition 3(ii) To prove the proposition define

$$V(x,N) = u(w-x) + \beta \max_{\alpha} \int u((1-\alpha)xR_B + \alpha xr)dP^{\alpha,x}(r).$$

where $P^{\alpha,x}$ is the lower-perceived distribution corresponding to the siminf for the act induced by x, α. Define

$$V(x,\infty) = u(w-x) + \beta \max_{\alpha} \int u((1-\alpha)xR_B + \alpha xr)dP(r).$$

I will show that V has decreasing differences, in the sense that for $x'' > x'$

$$V(x',\infty) - V(x'',\infty) \geq V(x',N) - V(x'',N),$$

A sufficient condition for the above inequality is that for all $x \in [x', x'']$

$$\begin{aligned}\frac{d}{dx}\left[\max_{\alpha}\int u((1-\alpha)xR_B + \alpha xr)dP^{\alpha,x}(r)\right] \\ \geq \frac{d}{dx}\left[\max_{\alpha}\int u((1-\alpha)xR_B + \alpha xr)dP(r)\right]\end{aligned} \tag{5}$$

Let $\alpha(x,N)$, $\alpha(x,\infty)$ be the optimal allocation proportions. Let $\xi(\alpha,r) = (1-\alpha)xR_B + \alpha xr$. I prove that the inequality in (5) holds in two parts. First

$$\frac{d}{dx}\left[\max_{\alpha}\int u(\xi(\alpha,r))dP^{\alpha,x}(r)\right] \tag{6}$$

$$= \int u'(\xi(\alpha(x,N),r))((1-\alpha(x,N))R_b + \alpha(x,N)r)dP^{\alpha,x}(r) \tag{7}$$

$$\geq \int u'(\xi(\alpha(x,N),r))((1-\alpha(x,N))R_b + \alpha(x,N)r)dP(r). \tag{8}$$

where the equality in (7) follows from the envelope theorem (which implies that the marginal effect on $P^{\alpha,x}$ is second order because $P^{\alpha,x}$ is the solution to the maximization problem defining the siminf). The inequality in (8) will follow by $P \geq_{FOSD} P^{\alpha,x}$ if the integrand in (8) is decreasing in r. Taking the derivative and rearranging we can see that this is the case if and only if the coefficient of relative risk aversion $-u''(z)z/u'(z)$ is greater than or equal to 1.

To complete the proof that (5) holds I show that

$$\int u'(\xi(\alpha(x,N),r))((1-\alpha(x,N))R_b+\alpha(x,N)r)dP(r)$$
$$\geq \frac{d}{dx}\left[\max_{\alpha}\int u((1-\alpha)xR_B+\alpha xr)dP(r)\right].$$

This will follow by the envelope theorem and Proposition 3(i), which says that $\alpha(x,\infty)\geq\alpha(x,N)$. We need only show that for $\alpha\in[\alpha(x,N),\alpha(x,\infty)]$ the expression on the LHS of the above inequality is decreasing in α. To be precise, we need that $\int u'((1-\alpha)xR_B+\alpha xr)((1-\alpha)R_b+\alpha r)dP(r)$ is decreasing in α. Taking the derivative, we need

$$\int[u''(\xi(\alpha,r))((1-\alpha)xR_b+\alpha xr)+u'(\xi(\alpha,r))](r-R_b)dP(r)\leq 0. \tag{9}$$

Notice that the term in brackets in the integrand, $u''((1-\alpha)xR_B+\alpha xr)((1-\alpha)xR_b+\alpha xr)+u'((1-\alpha)xR_B+\alpha xr)$, is less than or equal to zero by the assumption of $-u''(z)z/u'(z)\geq 1$, but $(r-R_b)$ is negative for low values of r and positive for high values. Assuming CRRA utility, (9) reduces to

$$\int u'((1-\alpha)xR_B+\alpha xr)(xr-xR_b)dP(r)\geq 0 \tag{10}$$

Notice that the LHS of (10) is exactly the derivative with respect to α of expected utility. Thus for all $\alpha\leq\alpha(x,\infty)$, (10) holds by concavity of the objective function. Then (5) holds, as desired. □

C.11 Proof of Proposition 4

Proof. Let $P_N^{b,s}$ be the induced distribution for the agent with capacity N (either Cautious or Reckless). The first order conditions for the constrained agent's problem are given by

$$[b]\quad u'(w-b-\rho s)=\frac{d}{db}\left[\beta\int u(w+bR_b-sr)dP_N^{b,s}(r)\right]$$
$$[s]\quad u'(w-b-\rho s)p=\frac{d}{ds}\left[\beta\int u(w+bR_b-sr)dP_N^{b,s}(r)\right].$$

In equilibrium (with $b=s=0$) the FONC for b implies that $R_B=1/\beta$, regardless of capacity or attitude. Let ρ^N be the equilibrium risky asset price for the capacity N agent.

Zero net supply requires that

$$u'(w)\rho^N = \frac{d}{ds}\left[\beta \int u(w+sr)dP_N^{0,s}(r)\right]_{s=0}. \tag{11}$$

Clearly when $s=0$ we have $\beta \int u(w+sr)dP_{N''}^{0,s}(r) = \beta \int u(w+sr)dP_{N'}^{0,s}(r)$ for all N', N''. For any $s \neq 0$ for a cautious (reckless) DM, $\beta \int u(w+sr)dP_{N'}^{0,s}(r) < (>)\ \beta \int u(w+sr)dP_{N''}^{0,s}((r)$ for $N'' > N'$, since $P_N^{0,s}$ is the solution to a maximization (minimization) problem. Therefore the derivative on the RHS of (11) is increasing (decreasing) in N when the DM is cautious (reckless). The result follows. $\square$

C.12 Proofs for Section 5.2

Let $f'_-(\omega)$ and $f'_+(\omega)$ be left and right derivatives of f at ω respectively. The proof of Proposition 5 makes use of the following lemma.

Lemma 4. Let f be a decreasing and continuous function. Suppose f has a kink at d ($f'_+(d) > f'_-(d)$). Then siminf(f) is constant in a neighborhood of d.

Proof. Suppose that there is a cut-off at d, so $t_n = d$ and siminf(f) is discontinuous at t_n. Fix t_{n-1}, t_{n+1}. The value generated by a cutoff in (t_{n-1}, t_{n+1}) is

$$V(t) = [P(t) - P(t_{n-1})]u(f(t)) + [P(t_{n+1}) - P(t)]u(f(t_{n+1}))$$

Then optimality of $t_n = d$ implies that the left derivative of V at d must be positive and the right derivative must be negative:

$$V'_-(d) = [P(d) - P(t_{n-1})]u'(f(d))f'_-(d) + p(d)[u(f(d)) - u(f(t_{n+1}))] \geq 0 \tag{12}$$

$$V'_+(d) = [P(d) - P(t_{n-1})]u'(f(d))f'_+(d) + p(d)[u(f(d)) - u(f(t_{n+1}))] \leq 0 \tag{13}$$

Equations (12) and (13) imply

$$f'_+(d) \leq \frac{-p(d)}{u'(f(d))[P(d) - P(t_{n-1})]}[u(f(d)) - u(f(t_{n+1}))] \leq f'_-(d)$$

which contradicts $f'_+(d) > f'_-(d)$. $\square$

Let t_N be the highest cut-off defining siminf$(y|d)$. Clearly $t_N \leq d$, since y is flat above d for a full insurance contract. Moreover Lemma 4 implies that $t_N < d$.

Proof. Proposition 5. Let $l(\omega|d,c)$ be the amount paid by the consumer when the loss is ω given a contract with deductible d and coverage rate c. Denote the perceived value to a cautious agent of an insurance contract characterized by d, c as

$$U^N(d,c) = \max_{\hat{t}_1,\ldots,\hat{t}_N} \sum_{n=1}^{N+1} [P(\hat{t}_n) - P(\hat{t}_{n-1})] u(w - l(\hat{t}_n|d,c)) \tag{14}$$

A fully rational agent would value the contract at

$$U^\infty(d,c) = \int_0^d u(w-\omega)dP(\omega) + \int_d^{\bar{\omega}} u(w - d - (1-c)(\omega - d))dP(\omega)$$

Let $\{t_1, \ldots, t_N\}$ be the solution to the maximization problem in 14, and let $n^* = \max\{n \in \{1, \ldots, N\} : t_n \leq d\}$ be the index of the highest cut-off below d. By the envelope theorem

$$U_d^N(d,c) = -c \sum_{n=n^*+1}^{N+1} [P(t_n) - P(t_{n-1})] u'(w - d - (1-c)(t_n - d)).$$

Moreover $U_d^\infty(d,c) = -c \int_d^{\bar{\omega}} u'(w - d - (1-c)(\omega - d))dP(\omega)$. Notice that

$$\begin{aligned} U_d^N(d,c) &< c[P(d) - P(t_{n^*})] u'(w - d - (1-c)(t_{n^*+1} - d)) \\ &\quad - c \sum_{n=n^*+1}^{N+1} [P(t_n) - P(t_{n-1})] u'(w - d - (1-c)(t_n - d)) \\ &\leq -c \int_d^{\bar{\omega}} u'(w - d - (1-c)(\omega - d))dP(\omega) \quad = \quad U_d^\infty(d,c) \end{aligned}$$

where the first inequality follows since Lemma 4 implies $t_{n^*} < d$, and the second from concavity of u. Notice that the second inequality holds with equality if and only if $c = 1$ for u strictly concave. This proves part 1 of Proposition 5.

The proof of part 2 is similar. In this case

$$
\begin{aligned}
U_c^N(d,c) &= \sum_{n=n^*+1}^{N+1}[P(t_n)-P(t_{n-1})]u'(w-d-(1-c)(t_n-d))(t_n-d) \\
&> -[P(d)-P(t_{n^*})]u'(w-d-(1-c)(t_n-d))(t_n-d) \\
&\quad + \sum_{n=n^*+1}^{N+1}[P(t_n)-P(t_{n-1})]u'(w-d-(1-c)(t_n-d))(t_n-d) \\
&\geq \int_d^{\bar{\omega}} u'(w-d-(1-c)(\omega-d))(\omega-d)dP(\omega) \\
&= U_c^{\infty}(d,c)
\end{aligned}
$$

where the first inequality follows from Lemma 4 and second from concavity of u. □

C.12.1 Proof of Proposition 6

This result follows from a similar argument as Proposition 5. The key to the proof of Proposition 5 was the result of Lemma 4 that $t_{n^*} < d$. When the baseline contract is full insurance above d the following result allows us to draw the analogous conclusion that the value of the greatest cut-off below d is increasing in N.

Let $t^N = t_0^N, \ldots, t_{N+1}^N$ be the cutoffs defining siminf$(y|d)$ when the agent has capacity N and $t^{N+1} = t_0^{N+1}, \ldots, t_{N+2}^{N+1}$ be the cutoffs for capacity $N+1$. When $c = 1$ we have the immediate corollary ofLemma 7 and Valenzuela-Stookey (2022) Lemma 1: the largest cut-off below d is increasing in N. Notice that for $c = 1$, $n^* = N$ where N is the capacity of the agent. This follows since ex-post wealth is constant above d, and so it would not be optimal to have a cut-off above d.

Corollary 3. For a full insurance contract above a deductible, $t_N^N < t_{N+1}^{N+1}$.

We can now prove Proposition 6.

Proof. (Proposition 6) For $c = 1$ an envelope theorem implies that

$$
\begin{aligned}
U_d^N(d,c) &= -[1-P(t_N^N)]u'(w-d) \\
&\leq [P(t_{N+1}^{N+1})-P(t_N^N)]u'(w-d) - [1-P(t_N^N)]u'(w-d) \\
&= U_d^{N+1}(d,c)
\end{aligned}
$$

The proof for decreasing c (recall that c is bounded above by 1), is similar. $\square$

C.12.2 Propositions 9 and 10

Propositions 9 and 10 follow immediately from Lemma 7 and Valenzuela-Stookey (2022) Lemma 12 and Theorem 1 respectively.

C.12.3 Proof of Propositions 8 and 7

I give the proof Proposition 8 here. The argument for Proposition 7 is essentially identical.

Proof. Let m be the out of pocket maximum, and $U^N(m)$ be the perceived value, holding c and d fixed. Let $\underline{l}(m) = \min\{\omega \in \Omega : l(\omega|d,c,m) = m\}$. Let $C^*(N,[0,\underline{l}(m)])$ be the cut-offs corresponding to the elements of the plan siminf. Since P is absolutely continuous and $\underline{l}(\cdot)$ is continuous, $C^*(N,[0,\cdot])$ is upper-hemicontinuous by Berge's maximum theorem.

Recall that for a decreasing and continuous function f on an interval of the reals the cell function corresponding to simple lower bounds is given by

$$v([a,b]) = f(b)P(\{\omega \in [a,b]\}). \tag{15}$$

Since utility is strictly decreasing on $[0,\underline{l}(m)]$ and P is full support the cell function generating $U^N(m)$ is strictly submodular by Lemma 7. Moreover, it satisfies the conditions for regularity in Tian (2015), so by Tian (2015) Theorem 3, C'' and C' are sandwiched for all $C'' \in C^*(N+1,[0,\underline{l}(m)])$ and $C' \in C^*(N,[0,\underline{l}(m)])$, so upper-hemicontinuity implies that for ε small enough there exist sandwiched selections from $C^*(N,m)$ and $C^*(N+1,m-\varepsilon)$.

For an interval $I \subseteq \Omega$ and $C \in C^N(I)$ let $V(I,C)$ be the coarse value corresponding to the cell function in (15). Since utility is constant above $\underline{l}(m)$, we can write $U^N(m) = V([0,\underline{l}(m)],C^*(N,[0,\underline{l}(m)])) + (1-P(\underline{l}(m)))u(w-m)$ (with an abuse of notation when $C^*(N,[0,\underline{l}(m)])$ is non-singleton). The result follows from Valenzuela-Stookey (2022), Theorem 4. $\square$

C.12.4 Lemma 5

Lemma 5. Consider two plans with no out-of-pocket maximum, coverage rate c, deductibles d and d', with $d < d'$, and premiums p, p' such that $p' - p = (1-c)(d'-d)$. Then a cautious agent is indifferent between the two plans if the lowest cut-off defining siminf of the high deductible plan is (weakly) greater than d'. [26]

Proof. Let t be the lowest cut-off for the siminf of the high deductible plan. Since the high deductible dominates the low deductible plan, any set of cut-offs defines a less preferred lower bound for the latter than for the former. Since the siminf of the high deductible plan is also dominated statwise by the low deductible plan it is an element of the siminf for the low deductible plan as well. □

D Comparative Statics

Given the centrality of the siminf and simsup in the representations of incomplete and complete preferences discussed above, it helpful in applications to understand how these functions vary with the parameters of the problem. I discuss real valued acts, with the understanding that all conclusions apply to the utility images of any acts. Moreover, I make statements about *every* element of siminf and simsup, with the understanding that these apply "up to sets of measure zero under P". I first focus on properties of $\text{siminf}_{N,w,P}$ when $\Omega = [\underline{\omega}, \bar{\omega}] \subset \mathbb{R}$ and w is increasing. Results in this setting can be extended in two ways. First, since $\text{siminf}_{N,w,P} = -\text{simsup}_{N,-w,P}$ all results regarding siminf can be translated directly to simsup. Second, using the dual "Lebesgue approach" of Appendix C.4.1, results for increasing functions on an interval can be translated to results for arbitrary bounded functions on an arbitrary state space. Finally, for simplicity I will assume throughout that $w \geq \text{siminf}_{N,w,P}$. Using Proposition 1 we know when this will hold. Moreover, Proposition 1 tells us where violations of statewise dominance can occur. The results presented below will apply with only minor modifications when $\neg(w \geq \text{siminf}_{N,w,P})$, as we can just ignore the zero measure set on which violations of statewise dominance occur.

[26]The siminf for the high deductible plan need not be unique. I state the result in this way for simplicity, but it holds as long as the condition is satisfied for a selection from the siminf. In fact, if siminf for the high deductible plan is single valued then the converse holds as well.

D.1 Cell functions

Let $\mathcal{P}^N$ be the space of N-element partitions of Ω, with typical element $\tau = (T_i)_{i=1}^N$. We can write the problem of finding $\text{siminf}_{N,w,P}$ as

$$\max_{\tau \mathcal{P}^N} V(\tau) := \sum_{i=1}^{N} v(T_i)$$

where $v(T_i) := \sup_{\omega \in T_i} w(\omega)P(T_i)$. Problems of this form are studied in Tian (2015), Tian (2016), and Valenzuela-Stookey (2022). Following Tian (2015), I refer to v as the siminf *cell function*, and V as the *coarse value*. Comparative statics for siminf follow from the results of these three papers, given the following observations. First, it is without loss of generality to restrict attention to interval partitions. The following is implied by the proof of Lemma 3.

Lemma 6. If w is increasing then every element of $\text{siminf}_{N,w,P}$ and $\text{simsup}_{N,w,P}$ has partition consisting only of intervals.

Suppose $\tau \in T^N(X)$ is an interval partition of the sub-interval $X \subset \Omega$. Then τ can be described by a set of cut-offs $C = \{t_i\}_{i=1}^{N-1} \in \mathbb{R}^{N-1}$ giving the interior endpoints of the partition intervals. I will sometimes use the convention $t_0 = \inf\{X\}$ and $t_N = \sup\{X\}$ when referring to a partition of the interval X. Cut-off states can be assigned to partition cells in any way.[27] Since for finding the siminf it will always be optimal to assign cut-off states to the higher interval I will assume this assignment from now on. Let $C^N(\Omega)$ be the set of cut-offs defining interval partitions of Ω. Abusing notation, I will write $V(\Omega, C)$ to denote the coarse value of the partition induced on Ω by cut-off vector $C \in C^N(\Omega)$.

Endow $C^N(\Omega)$ with the pointwise partial order. Denote the least upper bound and greatest lower bound of two cut-off vectors $C, C' \in C^N(\Omega)$ by $C \vee C'$ and $C \wedge C'$. Define the union of two sets of cut-offs in the obvious way, as the ordered union of the two sets. [28]

[27] Technically, a full description of the partition should include, in addition to the cut-offs, a vector in $\{0,1\}^{N+1}$ identifying whether each cut-off state is assigned to the interval immediately above or below it. A cut-off assigned the the lower interval is considered to be lower than the same cut-off assigned to the higher interval, otherwise the usual order on $\mathbb{R}^N$ is used to order cut-off vectors.

[28] By this definition the union of two cut-off vectors defines a partition which is the join, in the refinement sense, of the partitions defined by each of the individual cut-off vectors. This is different than the join of the cut-off vectors, which I define as the coordinate-wise maximum. The later notion is restricted to cut-off vectors of the same length, whereas the union can be taken of any two cut-off vectors.

For an interval $I = [a, b]$ and c in I, let $\Delta v(c, I) = [v([a, c]) + v([c, b])] - v(I)$.

Definition. *A cell function v is (strictly)* ***submodular*** *if for all intervals I, I' with $I \cap I' \neq \varnothing$, we have $v(I \cap I') + v(I \cup I')\ (<) \leq\ v(I) + v(I')$.*

Notice that for submodular cell functions $v(c, I)$ is always non-negative. Moreover, we have the the following observation

Proposition 16. (Tian (2015), Observation 1) A cell function is (strictly) submodular if and only if for all intervals I, I' with $(I' \subsetneq I\ )\ I' \subseteq I$ and any $c \in I'$ we have $\Delta v(c, I)\ (>\ ) \geq\ \Delta v(c, I')$.

Definition. *A coarse value is (strictly)* ***supermodular*** *if it is (strictly) supermodular as a function of cut-off vectors.*

Proposition 17. (Tian (2015), Observation 3) The coarse value is (strictly) supermodular if and only if the cell function is (strictly) submodular.

It is easy to verify that the siminf cell function satisfies the condition of Proposition 16, and is thus submodular, so that the associated coarse value is supermodular by Proposition 17.

Lemma 7. The siminf cell function $v(\cdot)$ is submodular. It is strictly submodular if P has full support and w_a is strictly increasing.

Proof. For any interval $[a, b]$ and $c \in [a, b]$ we have

$$\begin{aligned} \Delta v(I, c) &= \underline{w}_a([a, c])P([a, c]) + \underline{w}_a(c, b)P([c, b]) - \underline{w}_a([a, b])P([a, b]) \\ &= (\underline{w}_a([c, b]) - \underline{w}_a([a, c]))P([c, b]) \end{aligned}$$

where the last equality follows since $\underline{w}_a([a, c]) = \underline{w}_a([a, b])$ when w_a is increasing, as it is here by assumption. Suppose $[a, b] \subseteq [l, m]$. Then $\underline{w}_a([c, b]) = \underline{w}_a([c, m])$, $\underline{w}_a([l, c]) \leq \underline{w}_a([a, c])$ and $P([c, b]) \leq P([c, m])$, so $\Delta v([a, b], c) \leq \Delta v([l, m], c)$. The claim follows from Proposition 16. □

Corollary 4. The coarse value for the siminf cell function is supermodular. It is strictly supermodular if w_a is strictly increasing and P has full support.

Lemma 7 and Corollary 4 allow us to apply the comparative statics results of Tian (2015), Tian (2016), and Valenzuela-Stookey (2022).

References

J. Abaluck and J. Gruber. Choice inconsistencies among the elderly. *American Economic Review*, 101(4):1180–1210, 2011.

D. S. Ahn and H. Ergin. Framing contingencies. *Econometrica*, 78(2):655–695, 2010.

F. J. Anscombe and R. J. Aumann. A definition of subjective probability. *Annals of mathematical statistics*, 34(1):199–205, 1963.

R. J. Aumann. Utility theory without the completeness axiom. *Econometrica: Journal of the Econometric Society*, pages 445–462, 1962.

B. D. Bernheim and C. Sprenger. Direct tests of cumulative prospect theory. *Available at SSRN 3350196*, 2019.

T. F. Bewley. Knightian decision theory. part i. *Decisions in economics and finance*, 25(2): 79–110, 2002.

S. Bhargava, G. Loewenstein, and J. Sydnor. Choose to lose: Health plan choices from a menu with dominated option. *The Quarterly Journal of Economics*, 132(3):1319–1372, 2017.

S. H. Chew, B. Miao, and S. Zhong. Partial ambiguity. *Econometrica*, 85(4):1239–1260, 2017.

D. M. Cutler and R. Zeckhauser. Extending the theory to meet the practice of insurance. *Brookings-Wharton Papers on Financial Services*, 2004(1):1–53, 2004.

M. Dean and P. Ortoleva. Allais, ellsberg, and preferences for hedging. *Theoretical Economics*, 12(1):377–424, 2017.

F. Echenique, L. Pomatto, and J. Vinson. Twofold conservatism in choice under uncertainty. *arXiv preprint arXiv:2012.14557*, 2020.

F. Echenique, M. Miyashita, Y. Nakamura, L. Pomatto, and J. Vinson. Twofold multiprior preferences and failures of contingent reasoning. *Journal of Economic Theory*, 202:105448, 2022.

D. Ellsberg. Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, pages 643–669, 1961.

L. G. Epstein and J. Zhang. Subjective probabilities on subjectively unambiguous events. *Econometrica*, 69(2):265–306, 2001.

T. Galaabaatar and E. Karni. Subjective expected utility with incomplete preferences. *Econometrica*, 81(1):255–284, 2013.

P. Ghirardato and M. Marinacci. Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102(2):251–289, 2002.

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

I. Gilboa, F. Maccheroni, M. Marinacci, and D. Schmeidler. Objective and subjective rationality in a multiple prior model. *Econometrica*, 78(2):755–770, 2010.

F. Gul and W. Pesendorfer. Expected uncertain utility theory. *Econometrica*, 82(1):1–39, 2014.

L. Hartmann and T. F. Kauffeldt. A hierarchy of ambiguity aversion. Working Paper, 2019.

J. L. Horowitz. The bootstrap. In *Handbook of econometrics*, volume 5, pages 3159–3228. Elsevier, 2001.

P. Klibanoff, M. Marinacci, and S. Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892, 2005.

E. Lehrer and R. Teper. Extension rules or what would the sage do? *American Economic Journal: Microeconomics*, 6(1):5–22, 2014.

M. J. Machina and D. Schmeidler. A more robust definition of subjective probability. *Econometrica: Journal of the Econometric Society*, pages 745–780, 1992.

M. Mandler. Incomplete preferences and rational intransitivity of choice. *Games and Economic Behavior*, 50(2):255–277, 2005.

R. Mehra and E. C. Prescott. The equity premium: A puzzle. *Journal of monetary Economics*, 15(2):145–161, 1985.

P. Milgrom and I. Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2): 583–601, 2002.

P. G. Moffatt, S. Sitzia, and D. J. Zizzo. Heterogeneity in preferences towards complexity. *Journal of Risk and Uncertainty*, 51(2):147–170, 2015.

W. S. Neilson. Some mixed results on boundary effects. *Economics Letters*, 39(3):275–278, 1992.

P. Ortoleva. The price of flexibility: Towards a theory of thinking aversion. *Journal of Economic Theory*, 148(3):903–934, 2013.

K. R. Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Soc., 2005.

I. Puri. Simplicity theory of choice under risk. Working Paper, 2020.

N. Saponara. Revealed reasoning. *Journal of Economic Theory*, page 105096, 2020.

D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, pages 571–587, 1989.

U. Segal. The ellsberg paradox and risk aversion: An anticipated utility approach. *International Economic Review*, pages 175–202, 1987.

K. Seo. Ambiguity and second-order belief. *Econometrica*, 77(5):1575–1605, 2009.

Z. Shapira and I. Venezia. On the preference for full-coverage policies: Why do people buy too much insurance? *Journal of Economic Psychology*, 29(5):747–761, 2008.

D. Sonsino, U. Benzion, and G. Mador. The complexity effects on choice with uncertainty–experimental evidence. *The Economic Journal*, 112(482):936–965, 2002.

J. Tian. Optimal interval division. Working Paper, 2015.

J. Tian. Monotone comparative statics for cut-offs. Working Paper, 2016.

Q. Valenzuela-Stookey. Monotone portfolio problems. Working Paper, 2022.

W. K. Viscusi and W. A. Magat. Bayesian decisions with ambiguous belief aversion. *Journal of Risk and Uncertainty*, 5(4):371–387, 1992.

J. Von Neumann and O. Morgenstern. Game theory and economic behavior, 1944.