Quantile Maximization in Decision Theory*

Marzena J. Rostek†

This version: May 2007

Abstract

This paper introduces a model of preferences in which, given beliefs about uncertain outcomes, an individual evaluates an action using a quantile of the induced distribution. The choice rule of Quantile Maximization unifies maxmin and maxmax as maximizing the lowest and the highest quantiles of beliefs distributions, respectively, and generalizes them to any intermediate quantile.

Taking preferences over acts as a primitive, we axiomatize Quantile Maximization in a Savage setting. We derive probability measure(s) representing subjective beliefs, and a unique quantile that is maximized by implied preferences over probability distributions. The probability measure is unique for all quantiles other than the extreme, for which a set of capacities is derived. The result also provides a novel characterization of probabilistic sophistication that demonstrates that neither monotonicity nor the continuity axioms assumed in the literature are essential to probabilistic sophistication. We further axiomatize all quantiles with respect to capacities. We characterize risk preferences of quantile maximizers and discuss how the distinct from the Expected Utility properties of the model, robustness and ordinality, can be useful in studying choice behavior and problems of resource allocation, treatment effects, and robust economic policy design.

KEYWORDS: Quantile, Uncertainty, Robustness, Ordinal Utility, Value-at-Risk, Scenario-Based Analysis, Subjective Probability, Probabilistic Sophistication

JEL CLASSIFICATION: C72, D11, D80, D81

*I am grateful to Itzhak Gilboa, Stephen Morris and Ben Polak for their numerous helpful suggestions throughout the project. I also thank Dino Gerardi for his advice; Scott A. Boorman, Don Brown, Markus Brunnermeier, Larry Epstein, Simon Grant, Philippe Jehiel, Barry O'Neill, Jacob Sagi, Bernard Salanie, Barry Schachter, Peter Wakker, Jan Werner, John Weymark, and especially Marek Weretka for helpful discussions. I have also benefited from comments from seminar participants at Caltech, Columbia, Duke FUQUA, Minnesota, Northwestern, Oxford, Rochester, Wisconsin-Madison, Yale; and conference participants at the RUD 2006, the FUR XII 2006, the Midwest Economic Theory Meeting 2005, the XVI International Conference on Game Theory in Stony Brook 2005, and the X Spring Meeting of Young Economists 2005. All errors are mine.

†University of Wisconsin-Madison, Department of Economics, 1180 Observatory Drive, Madison, WI 53703-1393; E-mail: mrostek@ssc.wisc.edu.

1 Introduction

This paper examines choice behavior of an individual choosing among uncertain alternatives who selects the one with the highest quantile of the utility distribution. For example, she might be maximizing median utility, as opposed to mean utility, as she would if she were an Expected Utility maximizer. More generally, she might be comparing alternatives through some other quantile that corresponds to any given number between 0 and 1.

Although largely ignored in decision theory, quantiles have been widely used in many applied areas of economics:

• One of the most popular tools of risk management in finance, insurance, and banking is *Value-at-Risk* (*VaR*), defined as a quantile of the distribution of losses. VaR has become an industry standard, and its use is encouraged by the Federal Reserve Bank and the Securities and Exchange Commission, and the Bank for International Settlements.[1]

• In econometrics, quantiles have been used in techniques of robust estimation (LAD), providing an alternative to classical mean-based estimators, which are sensitive to large errors. Quantile regression (Koenker and Bassett [1978a]) is also becoming increasingly popular in optimal allocation design and evaluations of the effects of social policies or treatment effects by allowing distributional targeting and explicit analysis of distributional consequences. (Chamberlain [1994], Buchinsky [1995], Koenker and Hallock [2001] provide comprehensive surveys. See also Bhattacharya [2007], Chernozhukov, Imbens and Newey [2007], Chesher [2003].)

• In public economics, quantiles have been applied in measurement (e.g., population-based poverty lines), scenario-based analysis, used as order statistics, etc.

To which properties of quantiles then can one attribute their comparative advantage over moments, such as mean and variance? Two key characteristics of quantiles, robustness and ordinality, have proven to be attractive among practitioners: Quantile-based techniques are robust to fat tails, which are often met in practice, and offer predictions not driven by outliers. More pragmatically, quantiles enable policy making that is precautionary and at the same time does not require assumptions about the extreme tails of distributions, about which often little is known in practice (e.g., income distributions). Unlike tools based on mean or standard deviation, quantiles do not require that there exist moments of any order, which is a problem for example in non-life insurance, finance and income studies. In risk analysis, quantiles allow focusing on downside risk instead of the requirement to treat positive and negative deviations symmetrically, as forced by standard deviation. In

[1] Although used since the beginning of the 19th century, VaR became popular among trading institutions during the 1990s with the influential report on derivatives practices of the Group of 30 in 1993, the RiskMetrics service launched by JP Morgan in 1994 to promote the use of VaR, and the market risk capital requirements set for banks by the Basel Committee on Banking Supervision in 1995. For more on the history of Value at Risk, its applications, and ongoing theoretical as well as econometric research that it has spurred, see VaR's Web site: `http://www.GloriaMundi.org`.

terms of ordinality, quantiles are well suited to the econometric analysis of nonlinear models, unlike techniques based on expectations. Quantiles have the advantage of not requiring any parametric assumptions about utilities, so they can deal with the behavior of agents whose preferences are not known or a group of agents with heterogenous preferences (e.g., in curve-based grading schemes, agency problems, expert recommendations). In such environments, cardinal representations of choice behavior may not be appropriate and thus should not affect policy recommendations. *Quantile Maximization* does complement the existing models of choice under uncertainty, virtually all of which imply the use of cardinal properties of utility functions over outcomes.

To our knowledge, apart from the work of Manski (1988) and a recent contribution by Chambers (2007), on which we elaborate below, quantiles have not been studied in choice theory. There are two famous exceptions: *maxmin* and *maxmax.* Decision makers selecting an alternative with the highest minimal or maximal payoff can be viewed as maximizing the lowest or the highest quantile, respectively. It is worth emphasizing that Quantile Maximization nests these two choice rules and it does so in a non-trivial way. Maxmax and especially maxmin have been applied in game theory, robust control, individual and social choice, bargaining, voting, and other areas in economics. These criteria have, however, been commonly criticized for basing the choice on what may be extreme and unlikely outcomes. Indeed, maxmin agents would not invest, would not drive, and so on.

Admittedly, many choice rules, including the Expected Utility itself, can model less extreme choice behavior. Surprisingly, there is no model that captures more moderate prefernces while preserving the qualitative properties of maxmin and maxmax that Expected Utility does not exhibit, such as ordinality and robustness. Perhaps the closest related concept is the α*-maxmin* rule, defined as a convex combination of the minimal and the maximal payoffs with the fixed weights α and $1-\alpha$ (Hurwicz [1951], Arrow and Hurwicz [1972]).[2] The α-maxmin decision rule is, however, not ordinal. Furthermore, like maxmin and maxmax, α-maxmin loses the entire information contained in the prior except for its support. Incidentally, Quantile Maximization offers a family of ordinal and robust criteria intermediate between maxmin and maxmax. Quantiles do preserve ordinality, and compared to the extreme maxmin and maxmax, quantiles incorporate richer (outcome and probability) information from uncertain alternatives faced by an agent, additionally gaining the attribute of robustness. Further, as this paper establishes, there is a sense in which Quantile Maximization is a unique ordinal rule.

[2] α-maxmin was introduced in a context of "complete ignorance" and subsequently applied also to decision problems under uncertainty. Maxmin and maxmax are useful decision criteria in settings where an analyst has no probabilistic information about events. Nonetheless, such information is often available, especially about part of the domain, which is all that quantiles require. Our Theorem 3 demonstrates how quantiles admit imprecise information, represented by the multiplicity of beliefs, ensuring local accuracy around a unique quantile.

This paper formalizes the concept of Quantile Maximization in choice-theoretic language to provide a foundation for both its practice and the application in economic theory, and to study its implications for decision making.

Results. We model an individual choosing between uncertain alternatives who evaluates each alternative by the τ^{th}-quantile of the implied distributions, and selects the alternative with the highest quantile payoff. Thus, under Quantile Maximization, a decision maker is characterized by a scalar $\tau \in [0, 1]$; subjective beliefs over events π; and a rank order over outcomes. The central theoretical contributions of the paper are the axiomatization of Quantile Maximization (with respect to probabilities and capacities) and a novel characterization of probabilistic sophistication. As a by-product, our results axiomatize maxmin and maxmax under uncertainty. To the best of our knowledge, we are the first to derive and characterize the implied beliefs of agents who choose according to maxmin or -max.[3] We next describe the main results.

We provide an exact characterization of the model by jointly axiomatizing Quantile Maximization and subjective probabilities in a Savage setting. That is, taking preferences over Savage-style acts (maps from states to outcomes) as a primitive, we find conditions that are necessary and sufficient for those preferences to admit a quantile representation. We derive probability measure(s) representing subjective beliefs, an ordinally unique utility index and a unique quantile that is maximized by induced preferences over probability distributions. The axiomatization uncovers an important difference in how beliefs enter the preferences of the extreme and the intermediate quantiles: For all values of τ strictly between 0 and 1, the probability measure derived to represent subjective beliefs is unique (and also convex-ranged and finitely additive). For the extreme values of τ, equal to 0 or 1, we derive a set of nonatomic measures that are monotone, but not necessarily finitely additive (capacities). This is intuitive: Choices of 0- or 1-maximizers do not depend on beliefs, just on their support; hence, these are consistent with any measure that assigns strictly positive (and less than one) values to the same outcomes. That beliefs of the extreme-τ maximizers satisfy additional properties is behaviorally interesting: While it might seem that agents choosing according to maxmin or maxmax can only distinguish between null and non-null events, our results reveal that the agents can "see" more. In particular, the property of monotonicity with respect to events related by strict set inclusion implies that maxmin and -max agents can also compare events nested in that sense.

We characterize Quantile Maximization with five conditions. Our central axiom that leads to existence and uniqueness of τ is a new monotonicity condition. The key implication

[3] The maxmin in uncertain settings discussed here should be distinguished from the maxmin over multiple priors, axiomatized by Gilboa and Schmeidler (1989) to study choice behavior under ambiguity.

of this axiom is that for any act, there exists an event, called a *pivotal event*, such that exchanging outcomes outside of this event in a way that preserves their rank with respect to the outcome on the pivotal event does not affect preferences over acts. Intuitively, after we derive beliefs, the axiom will imply that the induced lottery preferences remain unaffected by changing parts of the cumulative distributions below and above some quantile. We dub this axiom *Pivotal Monotonicity*, $P3^Q$. Compared to Savage (1954), our set of conditions retains P1 (Ordering) and P5 (Nondegeneracy). We drop all his remaining axioms, including P2 (the Sure-Thing Principle) and P3 (Eventwise Monotonicity). Savage's P4 (Weak Comparative Probability) is in fact implied by our axioms, but dispensing with P2 requires adding an additional condition that ensures additivity of derived probability measures. This is achieved by the axiom that provides the likelihood judgment, derived from preferences over acts, with a weak-order structure. We call the condition *Comparative Probability* ($P4^Q$). Finally, due to ordinality, the Archimedean axiom typically employed in a Savage setting, P6 (Small Event Continuity), is too strong for our model, as it implies mixture continuity.[4] We weaken P6 just enough to retain its implications for the nonatomicity of probability measures and to ensure that the quantile is left continuous (*Event Continuity*, $P6^Q$). In summary, our three new conditions characterize the monotonicity-substitution ($P3^Q$) and the continuity ($P6^Q$) properties of quantile maximizer's preferences and the additivity ($P4^Q$) of their beliefs.

While Pivotal Monotonicity seems suggestive of how the quantile may be obtained by comparing probability distributions (when probabilities are derived), the axiom does not say that the pivotal event is unique in a given act, or how pivotal events relate across acts. Nonetheless, the main challenge in axiomatizing Quantile Maximization was to derive a probability-measure representation for subjective beliefs. We could not directly use either Savage's (1954) or other arguments in the literature: Derivation of probabilities typically involves defining a likelihood relation – a binary relation on events – induced from the preference relation on acts ("separation of probabilities from preferences"). According to the commonly used definition (employed by Savage, and earlier by Ramsey [1931] and de Finetti [1937]), event E is assessed as *more likely than* event F if, for any pair of outcomes x and y, where x is strictly preferred to y, an individual strictly prefers betting on x when E occurs than when F occurs. In the quantile model, however, the likelihood relation generates only two equivalence classes on the entire collection of events: All events are judged either equally likely to the null set or to the whole state space. (For example, think of a median maximizer comparing events with probabilities 0.7 and 0.9.) Even if there is a probability measure that represents the beliefs of τ-maximizers, the likelihood relation will not allow one to recover it from a data set as rich as recording all choices in all decision problems.

Nonetheless, we show that the information contained in preferences over acts suffices for recovering probability measures. Our approach is to construct a sub-collection of "small"

[4]See, e.g., Grant (1995, Step 3 in the proof of Theorem II).

events and define a new likelihood relation, which although incomplete on the whole collection, is still complete on the sub-collection of "small" events. We use the new relation to construct a unique, finitely additive and convex-ranged measure that represents beliefs of the agent about events in that sub-collection. We then uniquely extend the derived measure to the remaining events which, as we show, can be partitioned into "small" events. As hinted above, the preferences of 0- and 1-maximizers[5] depend on, and hence reveal, less structure about the primitive on acts than do those of individuals with $\tau \in (0,1)$. In constructing the measure, the difference arises in the ability to strictly rank disjoint non-null events. While we show that this is possible for the class of preferences leading to $\tau \in (0,1)$, for $\tau = 0$ or $\tau = 1$ the rankings of acts are invariant to swapping outcomes on any disjoint non-null events. For $\tau = 0$ and $\tau = 1$, we derive a set of capacities.

Finally, we show that when our new Comparative Probability axiom is dropped, the remaining four conditions provide a new axiomatization of Quantile Maximization with respect to capacities – for all quantiles, not just extreme ones. Thus, curiously, the role of the new axiom is only to ensure additivity of the probability measure (cf. P4 in Savage [1954], P4* in Machina and Schmeidler [1992], and P4^{CE} in Grant [1995]).

Axiomatization contributes in two ways to the growing body of literature on probabilistic sophistication that was initiated by Machina and Schmeidler (1992). The goal of this line of research is to understand when choices of a decision maker are consistent with that person's having beliefs that conform to a probability measure, without restricting the actual decision rule to be an Expected-Utility or other functional form. Nonetheless, with the exception of the recent result by Chew and Sagi (2006), the existing derivations impose continuity or monotonicity conditions on the representation functional, thus narrowing down the class of preferences.

In fact, the continuity condition prevalent in the literature (Savage's P6) cannot hold in the model presented in this paper. The reason is that P6 implies mixture continuity of the representation functional, which cannot be satisfied by an ordinal representation, such as Quantile Maximization. Our result thus identifies a class of preferences disjoint from those studied in the literature for which the characterizations of probabilistic sophistications based on P6 cannot be used.

As an illustrative example, consider a median maximizer. Her choices violate all axioms in Machina and Schmeidler, except P1 (Ordering), P4 (Weak Comparative Likelihood) and P5 (Nondegeneracy) and all axioms in Grant (1995) except for P1 and P5'. Thus the median maximizer would not be probabilistically sophisticated according to these two characterizations.

[5] Clearly, uniquely pinning down a number $\tau \in [0,1]$ can be accomplished only after the measure representation of beliefs is derived. The maxmin and -max agents are identified by a condition on preferences, formally defined in Section 4.

Chew and Sagi (2006) established probabilistic sophistication without continuity or monotonicity. We do use the weakest, in the sense made precise in Section 7, notion of monotonicity of risk preferences (weak stochastic dominance). However, as with Savage's P3, our monotonicity axiom only enters into the derivation of a probability measure to show that non-null events are judged more likely than is the empty set. This is exactly how Chew and Sagi (2006) used their weakening of P3.[6]

Another related concern about the developments in probabilistic sophistication, which has not been emphasized thus far, is that (through mixture continuity) these results impose restrictions on the set of outcomes from which acts are defined. Admittedly, the existence of subjective beliefs about events should not depend on the properties of the set of outcomes. Our proof neither assumes nor implies any conditions placed on the outcome set. Consequently, unlike Machina and Schmeidler's (1992) and Grant's (1995), our results can be used to show that individuals without well defined utility functions (e.g., lexicographic agents) can also be probabilistically sophisticated.

Applications. From an empirical perspective, being free from parametric assumptions and moment restrictions, Quantile Maximization can be a useful tool in applications. We show that despite its ordinality property, the model admits an elegant characterization of risk attitudes: τ itself provides a comparative measure and a complete ranking of attitudes toward risk, ranging from extreme risk aversion ($\tau = 0$) to extreme risk loving ($\tau = 1$). The model allows an analyst to study risk attitudes without having first to characterize the concavity of utilities from the data. To formulate policy recommendations, it suffices to recover a unique parameter (τ) to pin down the entire preference ordering over acts. The number τ is a counterpart of the convexity of the Bernoulli utility function in expectation-based models.

From a modeling perspective, Quantile Maximization provides a tractable framework to study choice behavior formally in environments where robustness to the assumptions about utility functions is a concern, or where cardinal representations will be unduly restrictive or inappropriate; for example, when choice variables are categorical, when a policy maker has little knowledge about agents' preferences or seeks a rule that can apply to a population with heterogenous preferences.[7] Furthermore, quantiles can address designing optimal policies in resource allocation problems where no redistribution of outcomes – such as survival rates after a surgery, or test-scores – is possible. In Section 7.5, we discuss specific settings in

[6]Furthermore, how to establish probabilistic sophistication (without monotonicity and continuity) by directly defining a likelihood relation, rather than inducing it indirectly via an exchangeability relation, remains an open question.

[7]Since the policy maker has no information about the utility function over outcomes (as in the former example), or the rule will apply to potentially all utilities over outcomes (as in the latter one), an ordinal framework is suitable. As mentioned above, we show that within the class of such (ordinal and probabilistically sophisticated) rules, Quantile Maximization is a unique criterion that respects any monotonicity.

which the distinct theoretical properties of the quantile model – robustness to appropriate changes in distributions and ordinality – are desirable.

Related Literature. The maxmin choice rule has been formally studied by Roy (1952, *safety first* rule), Milnor (1954), Rawls (1971, *justice as fairness* theory), Maskin (1979), Barbera and Jackson (1988), Cohen (1992, *security level*), Segal and Sobel (2002), and others. Studies that formalize maxmax include those of Cohen (1992, *potential level*), Segal and Sobel (2002), and Yildiz (2007, *wishful thinking*). This paper complements these results by characterizing exactly what maxmin- and -max entail regarding an agent's beliefs about the likelihood of events.

Ordinal representations of preferences have been advocated by Börgers (1993, *pure-strategy dominance*), Chambers (2007) and earlier by Manski (1988, *quantile utility* model, *utility mass* model). Manski was the first to draw attention to the decision-theoretic attributes of Quantile Maximization and examine risk preferences of quantile maximizers, and one of the first to illustrate the appeal of quantile-based techinques in applied economics. Interested in the class of functions restricted only by appropriate monotonicity and ordinal covariance, Chambers (2007) demonstrated that these properties characterize generalized quantiles. Thereby, he axiomatized quantile-based functionals, taking as a primitive the real-valued bounded measurable functions, and provided several characterization results that illuminated the relation among the functionals, commonly used in applications, that satisfy the two defining conditions. We should also mention that in the field of Artificial Intelligence, the ordinal approach to modeling choice has been increasingly popular over the past decade. The goal is to develop decision rules that can be implemented by information systems, such as recommender systems that require less information about utilities and beliefs (*qualitative decision theory*). Some of the decision criteria already proposed can be interpreted as modeling preferences that are intermediate between maxmin and maxmax. (See for example, Boutilier [1994], Dubois *et al.* [2000], Dubois *et al.* [2002] and references therein.)[8]

Structure of the Paper. This paper proceeds as follows: Section 2 presents the model of Quantile Maximization and characterizes its properties. Section 3 states our axioms, and Section 4 provides the main results, namely the representation theorem and a characteriza-

[8] The idea of modeling preferences intermediate between maxmin and maxmax is not new. Apart from α-maxmin, Cohen and Jaffray (1980) characterized decision criteria under complete ignorance, demonstrating that only extreme outcomes can have first-order effects. Other related models involve a combination of Expected Utility and maxmin (Gilboa [1988], Jaffray [1988], Cohen [1992]) or, like *neo-additive capacities* by Chateauneuf, Eichberger and Grant (2006), place a fixed weight on the extreme outcomes and apply Expected Utility on the outcomes in between. These concepts were intended to explain deviations from Expected Utility, and they still rely on the expected-utility operator. As a result, the rules behave very differently than the original maxmin and maxmax.

tion of probabilistic sophistication. Section 5 outlines the proofs of the two central results. Section 6 relates the axiomatization and our characterization of probabilistic sophistication to the literature. Section 7 examines properties of risk preferences and discusses applications of the model. Finally, Section 8 offers concluding remarks. All proofs, unless otherwise noted in the text, appear in the Appendices.

2 The Quantile Maximization Model

Let $\mathcal{S} = \{..., s, ...\}$ denote a set of states of the world, and let $\mathcal{X} = \{..., x, y, ...\}$ be an arbitrary set of *outcomes*. An individual chooses among finite-outcome *acts*,[9] maps from states to outcomes. $\mathcal{F} = \{..., f, g, ...\}$ is the set of all such acts. The set of events $\mathcal{E} = 2^{\mathcal{S}} = \{..., E, F, ...\}$ is the set of all subsets of $\mathcal{S}$. A collection $\{\mathcal{S},\ \mathcal{X},\ \mathcal{E},\ \mathcal{F}\}$ defines the Savagean model of purely subjective uncertainty. An individual is characterized by a binary relation over acts in $\mathcal{F}$, which will be defined as a preference relation and taken to be the primitive of the model. As will become clear in the sequel, it is essential to work with the strict binary relation $\succ$. Indifference and weak preference will be defined as usual (here and for all strict binary relations throughout): $f \sim g \Leftrightarrow f \not\succ g$ and $f \not\prec g$, $f \succsim g \Leftrightarrow f \succ g$ or $f \sim g$. Let $\succ_x$ denote the preference relation over certain outcomes, $\mathcal{X}$, obtained as a restriction of $\succ$ to constant acts. We say that event E is *null* if for any two acts, $f,\ g$ which differ only on E, we have $f \sim g$.

Define the set of *simple* (finite-outcome) probability distributions over the outcomes (lotteries):

$$\mathcal{P}_0(\mathcal{X}) = \left\{ P = (x_1, p_1, ..., x_N, p_N) \middle| \sum_{n=1,...,N} p_n = 1,\ x_n \in \mathcal{X},\ p_n \geq 0, N \in \mathbb{N}_{++} \right\}. \tag{1}$$

Finally, δ_x denotes the degenerate lottery $P = (x, 1)$.

Let π stand for a probability measure on $\mathcal{E}$, and let u be a utility over outcomes $u : \mathcal{X} \to \mathbb{R}$. For each act, π induces a probability distribution over payoffs, referred to as a *lottery*. For an act f, Π_f denotes the induced cumulative probability distribution of utility $\Pi_f(z) = \pi[s \in \mathcal{S} | u(f(s)) \leq z,\ z \in \mathbb{R}]$. Then, for a fixed act f and $\tau \in (0, 1]$, the τ^{th} *quantile* of the distribution of the random variable $u(x)$ is a (generalized) inverse of the cumulative distribution at τ. The generalized inverse is defined as the smallest value z, such that the probability that a random variable will be less than z is not smaller than τ:

$$Q^{\tau}(\Pi_f) = \inf\{z \in \mathbb{R} | \pi[u(f(s)) \leq z] \geq \tau\}, \tag{2}$$

[9] An act f is said to be *finite-outcome* if its outcome set $f(\mathcal{S}) = \{f(s) | s \in \mathcal{S}\}$ is finite.

while for $\tau = 0$, the quantile is defined as[10]

$$Q^0(\Pi_f) = \sup\{z \in \mathbb{R} | \pi[u(f(s)) \leq z] \leq 0\}. \tag{3}$$

Definition 1 *A decision maker is said to be a* τ-quantile maximizer *if there exists a unique* $\tau \in [0,1]$*, a probability measure* π *on* $\mathcal{E}$ *and utility* u *over outcomes in* $\mathcal{X}$*, such that for all* $f, g \in \mathcal{F}$*,*

$$f \succ g \Leftrightarrow Q^\tau(\Pi_f) > Q^\tau(\Pi_g). \tag{4}$$

By analogy with the Expected Utility, where the mean is a single statistic evaluating a distribution, when choosing among lotteries, a τ-maximizer assesses the value of each lottery by the τ^{th} quantile realization. Notice that, although in general a correspondence, generically in payoffs, the set of optimal choices is a singleton.

The quantile model nests two choice rules famous in the literature of choice under risk: maxmin and maxmax. A decision maker choosing according to *maxmin* picks the act with the highest minimal payoff:

$$f \succ g \Leftrightarrow \min_{\{x \in f(\mathcal{S}) | \pi(x) > 0\}} u(x) > \min_{\{x \in g(\mathcal{S}) | \pi(x) > 0\}} u(x). \tag{5}$$

Maxmax dictates selection of the act with the highest maximal payoff:

$$f \succ g \Leftrightarrow \max_{\{x \in f(\mathcal{S}) | \pi(x) > 0\}} u(x) > \max_{\{x \in g(\mathcal{S}) | \pi(x) > 0\}} u(x). \tag{6}$$

That the maxmin and maxmax decision makers are, respectively, the 0- and 1-quantile maximizers follows from

$$Q^0(\Pi_f) = \min_{\{x \in f(\mathcal{S}) | \pi(x) > 0\}} u(x), \; Q^1(\Pi_f) = \min_{\{x \in f(\mathcal{S}) | \pi(x) > 0\}} u(x). \tag{7}$$

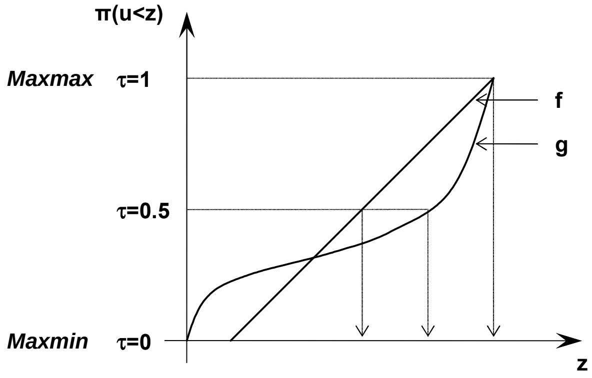
Quantile Maximization can thus be viewed as a generalization of those extreme choice rules to any intermediate quantile. While the main focus of the paper will be on finite-outcome acts, in Example 1, we illustrate the relation between maxmin, maxmax and Quantile Maximization using infinite-outcome acts.

Example 1 *Consider an individual who is facing a choice between two acts,* f *and* g*. Let* π *be the probability measure that represents the agent's beliefs. The cdf's induced by the acts,* f *and* g*, and measure* π *are plotted in Figure 1. The* 0*-quantile maximizer would choose* f*, the* 1*-quantile maximizer would be indifferent, and the median- (*$\tau = 0.5$*) maximizer would*

[10] Cf. Denneberg (1994). Clearly, the separate formulation for $\tau = 0$ merely ensures that the inverse is in the support; otherwise, the definition is conceptually the same as in the case of $\tau \in (0, 1]$ in that it picks the smallest value z (from the support of a given lottery), such that the probability that a realization will be less than z is at least 0.

prefer g.

Figure 1. Distributions induced by acts in Example 1



3 Axioms

Consider the following five axioms on $\succ$. The numbering is Savage's, the names of his conditions are adapted from Machina and Schmeidler (1992), and the superscript "^{Q}" (for "Quantile") is added to new axioms. The precise relationship of these axioms with those of Savage is presented at the end of the section.[11]

AXIOM P1 (**ORDERING**): *The relation $\succ$ is a weak order.*

This standard condition defines $\succ$ as a preference relation. To state the next axiom, for a fixed act $f \in \mathcal{F}$ and event E, such that $f^{-1}(x) = E$ for some $x \in f(\mathcal{S})$, we define the unions of events which by f are assigned outcomes strictly more and strictly less preferred to x, respectively:[12]

$$E_{fx+} = \{s \in \mathcal{S} | f(s) \succ x\}, \tag{8}$$
$$E_{fx-} = \{s \in \mathcal{S} | f(s) \prec x\}. \tag{9}$$

Note that since the acts are finite-ranged, every act induces a natural partition of the state space $\mathcal{S}$, which is the coarsest partition with respect to which it is measurable. The event E is an element of such a partition. Let the function g_{x+} be any mapping $g_{x+} : E_{fx+} \to \mathcal{X}$ with $g_{x+}(s) \succsim x$, for all $s \in E_{fx+}$ and similarly, let g_{x-} be any mapping $g_{x-} : E_{fx-} \to \mathcal{X}$ with $g_{x-}(s) \precsim x$, for all $s \in E_{fx-}$.

[11] In addition, the axioms are compared with other conditions in the literature in Section 6.

[12] For notational clarity, we assume (w.l.o.g.) that the set $\{y \in f(\mathcal{S}) | y \sim x, f(E) = x\}$ is a singleton. Alternatively, the unions (8) and (9) could be defined with respect to $f^{-1}\left(\bigcup_{x \in f(\mathcal{S})} y \middle| y \sim x\right)$ for some $x \in f(\mathcal{S})$.

AXIOM P3^Q (**PIVOTAL MONOTONICITY**): *For any act $f \in \mathcal{F}$, there exists a non-null event E, such that $f^{-1}(x) = E$ for some $x \in f(\mathcal{S})$, and for any outcome y, and subacts g_{x+}, g_{x-}, g_{y+}, and g_{y-} :*

$$\begin{bmatrix} g_{x+} \text{ if } E_{fx+} \\ x \text{ if } E \\ g_{x-} \text{ if } E_{fx-} \end{bmatrix} \succsim \begin{bmatrix} g_{y+} \text{ if } E_{fx+} \\ y \text{ if } E \\ g_{y-} \text{ if } E_{fx-} \end{bmatrix} \Leftrightarrow x \succsim y. \tag{10}$$

Before we explain the roles this axiom serves, we first interpret the following key implication: For an act $f \in \mathcal{F}$, event E, such that $f^{-1}(x) = E$ for some $x \in f(\mathcal{S})$, and all subacts g_{x+}, g_{x-}, g_{y+}, and g_{y-} define

$$f_E = \begin{bmatrix} g_{x+} \text{ if } E_{fx+} \\ x \text{ if } E \\ g_{x-} \text{ if } E_{fx-} \end{bmatrix}; \tag{11}$$

It follows from P3^Q that for any act $f \in \mathcal{F}$, there exists a non-null event E_f, such that $f^{-1}(x) = E$ for some $x \in f(\mathcal{S})$, and for all subacts g_{x+}, g_{x-}, g_{y+}, and g_{y-}:

$$f \sim f_E. \tag{12}$$

The last condition states that for a given act, there exists an event, which will be called a *pivotal event*, such that changing outcomes outside of that event in a (weakly) rank-preserving way does not affect preferences over acts – a form of separability. Crucially, what are held fixed during the transformation are the events assigned to outcomes which in the original act f are either strictly preferred or strictly less preferred to x, the outcome on the pivotal event. These events are fixed in the sense that after the transformation they will still map to outcomes preferred or less preferred, respectively, to x, with a weak preference permitted.[13] The measurability requirement that the act f be constant for the pivotal event ensures that the conditions (10) and (12) are non-trivial. (Otherwise, the state space could be taken as pivotal for any act.) This axiom will be the key to guaranteeing the existence and uniqueness of a number $\tau \in [0, 1]$. Intuitively, it implies that the induced preferences over lotteries will not change by replacing parts of the cumulative probability distributions below and above some quantile.

As its name suggests, P3^Q also provides preferences over acts with an appropriate, local notion of monotonicity. It states that replacing an outcome y on the pivotal event by a

[13] Or, for each act $f \in \mathcal{S}$, the state space can be partitioned into three event classes: one mapping to outcome(s) to which the act as a whole is indifferent, one mapping to outcomes strictly less preferred to these, and one mapping to outcomes strictly more preferred. The axiom asserts that replacing the outcomes within these three classes (i.e., preserving the ranking of outcomes across classes) leaves the agent indifferent.

(weakly) preferred outcome x always leads to a (weakly) preferred act. It is noteworthy that it suffices that the preference be monotonic on the pivotal event only; the axioms jointly allow for extending the monotonicity to the whole collection of events $\mathcal{E}$.

In addition, together with other axioms, Pivotal Monotonicity will imply that how an outcome in an act is assessed by a decision maker depends only on the likelihood of the event to which the outcome is assigned, and not on the event itself; that is, that the utility over outcomes, $u(x)$, is event independent. In a similar manner, in the presence of other axioms, P3^Q will render the property of being pivotal state-independent.[14] This will set up a relation between pivotal events across acts, on which Pivotal Monotonicity as such is silent.

AXIOM P4^Q (COMPARATIVE PROBABILITY): *For all pairs of disjoint events E and F, outcomes $x^* \succ x$, and subacts g and h,*

$$\begin{bmatrix} x^* \text{ if } s \in E \\ x \text{ if } s \in F \\ g \text{ if } s \notin E \cup F \end{bmatrix} \succ \begin{bmatrix} x^* \text{ if } s \in F \\ x \text{ if } s \in E \\ g \text{ if } s \notin E \cup F \end{bmatrix} \Rightarrow \begin{bmatrix} x^* \text{ if } s \in E \\ x \text{ if } s \in F \\ h \text{ if } s \notin E \cup F \end{bmatrix} \succsim \begin{bmatrix} x^* \text{ if } s \in F \\ x \text{ if } s \in E \\ h \text{ if } s \notin F \cup F \end{bmatrix}. \tag{13}$$

P4^Q asserts that replacing outcomes on the common subact mapping from $(E \cup F)^c$ does not strictly reverse the likelihood ranking of events E and F. It implies that adding a common complement event to either E or F will not strictly reverse the likelihood ranking between them. This, in turn, will ensure that the "more likely than" relation over events, to be induced from preferences over acts, is a weak order,[15] and it will provide its representation with a finitely additive form. Notice that the axiom has no effect in the cases leading to τ equal to 0 or 1; that it does not imply Savage's P4; and that no events are required to be non-null.

AXIOM P5 (NONDEGENERACY): *There exist acts f and g, such that $f \succ g$.*

This is the familiar non-triviality condition. By requiring that the individual not be indifferent among all outcomes, P5 assures that both the preference relation and the derived

[14] Formally, after we define a likelihood relation, we establish that exchanging outcomes between equally likely events will not alter agent's preferences (Lemma 3).

[15] We record the following interesting feature of preferences in the quantile model: Unlike in Savage, it is the monotonicity axiom, P3^Q, that together with P1, induces the standard likelihood relation, defined below in (22). P4^Q gives the weak-order structure to a new likelihood relation this paper constructs for the quantile model (Section 5), which uses more information contained in preferences $\succ$. The role of P4^Q in defining the new likelihood relation is less direct compared to the role played by P4 in Savage (1954) in the sense that is only required to hold for pairs of acts mapping events E and F to (the same) outcomes $x^* \succ x$.

likelihood relation are well-defined (in particular, non-reflexive) weak orders. It also permits establishing the uniqueness of a probability-measure representation of beliefs.

Before we state the final axiom, we introduce conditions that identify two interesting classes of preferences.

(**L**, "lowest"): For any act $f \in \mathcal{F}$, the pivotal event maps to an outcome from the least preferred equivalence class w.r.t. $\succ_x$ in the outcome set $\{x \in \mathcal{X} | x \in f(\mathcal{S})\}$. (14)

(**H**, "highest"): For any act $f \in \mathcal{F}$, the pivotal event maps to an outcome from the most preferred equivalence class w.r.t. $\succ_x$ in the outcome set $\{x \in \mathcal{X} | x \in f(\mathcal{S})\}$. (15)

Intuitively, these preferences will lead to $\tau = 0$ and $\tau = 1$, respectively.

Definition 2 *A preference relation over acts $\mathcal{F}$, $\succ$, satisfying $P3^Q$, is called* extreme *if either (**L**) or (**H**) holds. It is called* non-extreme *if neither (**L**) nor (**H**) is satisfied.*

Let us define two continuity properties that will be used in the final axiom.

(**P6**^{Q_*}) *For all events $E, F \in \mathcal{E}$, if for any pair of outcomes $x \succ y$,*

$$\begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \notin F \\ y \text{ if } s \in F \end{bmatrix}, \tag{16}$$

then there exists a finite partition $\{G_1, ..., G_N\}$ of $\mathcal{S}$, such that, for all $n = 1, ..., N$,

$$\begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \notin F \cup G_n \\ y \text{ if } s \in F \cup G_n \end{bmatrix}. \tag{17}$$

(**P6**$^{Q^*}$) *For all events $E, F \in \mathcal{E}$, if for any pair of outcomes $x \succ y$,*

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \notin F \end{bmatrix}, \tag{18}$$

then there exists a finite partition $\{H_1, ..., H_M\}$ of $\mathcal{S}$, such that, for all $m = 1, ..., M$,

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \in F \cup H_m \\ y \text{ if } s \notin F \cup H_m \end{bmatrix}. \tag{19}$$

AXIOM P6^Q (**EVENT CONTINUITY**): *The relation $\succ$ satisfies* P6^{Q_*} *for all pairs of events in $\mathcal{E}$ if $\succ$ is non-extreme or (**H**) holds; $\succ$ satisfies* P6$^{Q^*}$ *for all pairs of events in $\mathcal{E}$*

if (**L**) *holds or for a pair of a null event and any event* E *in* $\mathcal{E}$ *if* $\succ$ *is non-extreme.*

For the non-extreme preferences, the main force of this Archimedean axiom comes from the implication that the state space is infinite. Furthermore, it ensures that the quantile in the representation is left-continuous. Being formulated in terms of two-outcome acts, it has no further implications for risk preferences (i.e., the restriction of the implied lottery preferences to constant lotteries).

DISCUSSION. How "far" is Pivotal Monotonicity from the quantile representation? We note that the axiom does not require that the pivotal event is unique in a given act; and, therefore, it does not say how to relate pivotal events across acts. More importantly, the unique number τ in the unit interval $[0,1]$ can only be pinned down after the measure representation for beliefs is derived; and it is constructing the measure(s) that requires the most work. Alternatively, we now demonstrate that, given the remaining axioms, another condition could be used in place of Pivotal Monotonicity. Lemma 1 establishes a remarkable property of the preferences $\succ$: For any pair of acts, replacing the outcomes in their ranges in a weakly rank-preserving (w.r.t. $\succ_x$) way, does not affect the agent's preferences over these acts.

Consider act $f \in \mathcal{F}$, such that for some disjoint events E and F, $f^{-1}(E) = x^*$ and $f^{-1}(F) = x$. Define g_{x^*+} as a mapping $\mathcal{S} \to \mathcal{X}$, such that $g_{x^*+}(\mathcal{S}) \succsim x^*$, $g_{x^*-,x+}$ as a mapping, such that $x^* \succsim g_{x^*-,x+}(\mathcal{S}) \succsim x$, and g_{x-} as a mapping, such that $x \succsim g_{x-}(\mathcal{S})$.

Lemma 1 *Assume Weak Order (P1), Pivotal Monotonicity (P3^Q) and Nondegeneracy (P5). For all events E and F, all pairs of outcomes $x^* \succ x$ and $y^* \succ y$, and all sub-acts g_{x^*+}, $g_{x^*-,x+}$, g_{x-}, h_{y^*+}, $h_{y^*-,y+}$, and h_{y-},*

$$\begin{bmatrix} g_{x^*+} \text{ if } s \in G_1 \\ x^* \text{ if } s \in E \\ g_{x^*-,x+} \text{ if } s \in G_2 \\ x \text{ if } s \in F \\ g_{x-} \text{ if } s \in G_3 \end{bmatrix} \succ \begin{bmatrix} g_{x^*+} \text{ if } s \in G_1 \\ x^* \text{ if } s \in F \\ g_{x^*-,x+} \text{ if } s \in G_2 \\ x \text{ if } s \in E \\ g_{x-} \text{ if } s \in G_3 \end{bmatrix} \Rightarrow \begin{bmatrix} h_{y^*+} \text{ if } s \in G_1 \\ y^* \text{ if } s \in E \\ h_{y^*-,y+} \text{ if } s \in G_2 \\ y \text{ if } s \in F \\ h_{y-} \text{ if } s \in G_3 \end{bmatrix} \succ \begin{bmatrix} h_{y^*+} \text{ if } s \in G_1 \\ y^* \text{ if } s \in F \\ h_{y^*-,y+} \text{ if } s \in G_2 \\ y \text{ if } s \in E \\ h_{y-} \text{ if } s \in G_3 \end{bmatrix}. \tag{20}$$

Crucially, in the presence of Weak Order (P1) and Nondegeneracy (P5), condition (20) is equivalent to P3^Q; underlying this equivalence (established in Lemma 8 in Appendix 1A) is the ordinal nature of the model. Quantile Maximization could, thus, be axiomatized with condition (20) replacing Pivotal Monotonicity. We choose to work with Pivotal Monotonicity for expositional purposes and to facilitate comparisons with other models of choice.

One other consequence of property (20) is worth highlighting. The condition will allow inducing a likelihood relation over events, which it implicitly defines. Because of Lemma 1 the quantifiers in P4^Q can be considerably weakened compared to Savage's P4.

Relation to Savage's (1954) axioms. The only conditions that Quantile Maximization literally shares with the Subjective Expected Utility are the P1 (Ordering) and Nondegeneracy (P5). Compared to Savage, we drop the Sure-Thing Principle (P2) and weaken the continuity (P6, Small-Event Continuity) axiom of preferences. Our condition leading to the derived likelihood relation and the one driving the monotonicity of preferences are independent from their Savage's counterparts (respectively: P4, Weak Comparative Probability, and P3, Eventwise Monotonicity).

Not surprisingly, P2 (the Sure-Thing Principle) is too strong for Quantile Maximization.[16] Similarly to P3 (Eventwise Monotonicity), P2 fails when a change in the common subact affects how other outcomes rank w.r.t. $\succ_x$ in an act's outcome set. Precisely, what fails is the quantification "for all subacts," which our new axiom, P3^Q (Pivotal Monotonicity), weakens together with yet another quantifier in P3: "for all events." In the presence of other axioms, P2 still holds in the restricted class of acts, even after the second relaxation of quantifiers, but it has no independent implications (cf. Lemma 1). This weakening of P2 does not, however, preserve the structure in preferences that was used in the Subjective Expected Utility model to obtain additivity of the probability measure. To recover additivity, we strengthen P4 (Weak Comparative Probability) to P4^Q (Comparative Probability).

The Archimedean axiom of the Subjective Expected Utility theory, P6 (Small-Event Continuity), does not hold under Quantile Maximization. To see why, fix $\tau = 1$ and a pair of acts $f \succ g$. Then, taking $x \succ f(\mathcal{S})$, gives

$$f \prec \begin{bmatrix} x \text{ if } s \in E_n \\ g \text{ if } s \notin E_n \end{bmatrix} \tag{21}$$

for an arbitrary event E_n.[17] The original P6 ensures both that no consequence is infinitely desirable or undesirable, as well as that the derived probability measure is nonatomic. Crucially, what τ-maximization violates is the former but not the latter. We weaken the axiom P6 to P6^Q (Event Continuity) so that it retains only the continuity implications for probabilities.

[16] For example, consider three equally likely events E_1, E_2 and E_3. A median maximizer ($\tau = 0.5$) prefers act $\begin{bmatrix} 3 \text{ if } E_1 \\ 2 \text{ if } E_2 \\ 0 \text{ if } E_3 \end{bmatrix}$ to act $\begin{bmatrix} 4 \text{ if } E_1 \\ 1 \text{ if } E_2 \\ 0 \text{ if } E_3 \end{bmatrix}$, but she prefers $\begin{bmatrix} 5 \text{ if } E_3 \\ 4 \text{ if } E_1 \\ 1 \text{ if } E_2 \end{bmatrix}$ to $\begin{bmatrix} 5 \text{ if } E_3 \\ 3 \text{ if } E_1 \\ 2 \text{ if } E_2 \end{bmatrix}$.

[17] Analogous counter-examples can be constructed for any number $\tau \in [0, 1]$. We provide a sketch of the argument, completion of which relies on results proved in the sequel. For an intermediate value of τ, $\tau \in (0, 1)$, take an act f and let E be its pivotal event such that $f(E) = x$ and $\pi(E \cup E_{f-}) = \tau$. On a nonnull subevent of E, $\hat{E}$, which exists by P6^Q, replace x with $z \succ x$ to obtain the preference reversal $f \prec \begin{bmatrix} z \text{ if } s \in \hat{E} \\ g \text{ if } s \notin \hat{E} \end{bmatrix}$.

We close this section by remarking that the conditions (P6^{Q_*}) and (P6$^{Q^*}$) can be interpreted in terms of likelihood relations – we will use that interpretation in the sequel. Although the definition of likelihood we adopt to construct probabilities differs from the commonly used one (see Section 5), the standard definition, implicitly employed in P6^{Q}, still allows us to retrieve useful information from preferences. Formally, as proposed by Ramsey (1931) and adopted by Savage (1954), the likelihood relation $\succ^*$, a binary relation on $\mathcal{E}$, is defined through Savage's P4 (Appendix 1), implied by our P1 and P4^{Q}:

$$E \succ^* F \text{ if for all } x \succ y, \quad \begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \notin F \end{bmatrix}. \tag{22}$$

We also employ the following definition, which differs from $\succ^*$ in that it maps the events E and F, whose likelihood is being compared, to the *less* preferred outcome:

$$E \succ_* F \text{ if for all } x \succ y, \quad \begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \notin F \\ y \text{ if } s \in F \end{bmatrix}. \tag{23}$$

Given these definitions, conditions P6^{Q_*} and P6$^{Q^*}$ can be restated as follows:

(**P6**^{Q_*}, restated): *For all events $E, F \in \mathcal{E}$, if $E \succ_* F$, then there exists a finite partition $\{G_1, ..., G_N\}$ of $\mathcal{S}$ such that, for all $n = 1, ..., N$, $E \succ_* F \cup G_n$.*

(**P6**$^{Q^*}$, restated): *For all events $E, F \in \mathcal{E}$, if $E \succ^* F$, then there exists a finite partition $\{H_1, ..., H_M\}$ of $\mathcal{S}$ such that, for all $m = 1, ..., M$, $E \succ^* F \cup H_m$.*

In all cases leading to $\tau \in (0, 1]$, the relation $\succ_*$ will be central to deriving measure-representations for beliefs. The reason for altering the relation from the commonly used $\succ^*$ to $\succ_*$ is that $\succ^*$ would yield right-continuity of the quantile representation functional. We follow the convention in the literature and define (and derive) quantiles as left-continuous. The distinctive formulation of the condition in P6^{Q} for the subclass (**L**) of the extreme preferences is due to the fact that, in this case, P6^{Q_*} fails. (Section 5 clarifies this point.)

4 Axiomatic Foundations of Quantile Maximization

4.1 Probabilistic Sophistication

This section presents the first of two central theorems of the paper. The result shows that quantile maximizers' preferences $\succ$ over uncertain alternatives are consistent with them having a subjective probability distribution over the states in $\mathcal{S}$, thereby establishing probabilistic sophistication. Since the seminal paper by Machina and Schmeilder (1992), a formal definition of probabilistic sophistication, a property of preferences over acts, has been evolving. We adopt the following conceptualization:[18] Fix a probability measure π on the set of events $\mathcal{E}$. Each act $f \in \mathcal{F}$ can be mapped to a lottery in $\mathcal{P}_0(\mathcal{X})$ in a natural way, through the mapping $f \to \pi \circ f^{-1}$. We say a decision maker (or, the relation $\succ$) is *probabilistically sophisticated* if she is indifferent between two acts that induce identical probability distributions over outcomes. Formally, for all P, Q in $\mathcal{P}_0(\mathcal{X})$, and all f, g in $\mathcal{F}$,

$$\left(P = Q,\ \pi \circ f^{-1} = P,\ \pi \circ g^{-1} = Q\right) \Rightarrow f \sim g. \tag{24}$$

In passing, we define a mapping from acts in $\mathcal{F}$ to lotteries in $\mathcal{P}_0(\mathcal{X})$ using a fixed (possibly non-additive) measure λ. For an act $f \in \mathcal{F}$, rank the outcomes in f's outcome set $f(\mathcal{S})$ w.r.t. $\succ$: $x_1 \precsim x_2 \precsim ... \precsim x_N$ for some $N \in \mathbb{N}_{++}$; next, map the corresponding events $E_1, E_2, ..., E_N$ to numbers $p_1, p_2, ..., p_N$ in $[0, 1]$ according to: $p_1 = \lambda(E_1)$ and $p_n = \lambda\left(\bigcup_{m \leq n} E_m\right) - \lambda\left(\bigcup_{m' \leq n-1} E_{m'}\right)$ for $n \in \{2, ..., N\}$. As $\sum_{n \leq N} p_n = 1$, the mapping $f \to \lambda \circ f^{-1}$ uniquely yields a lottery $P \in \mathcal{P}_0(\mathcal{X})$. Further, given the mapping $f \to \lambda \circ f^{-1}$, a non-additive measure λ (uniquely) implies a cumulative probability distribution for a given act $f \in \mathcal{F}$, denoted by Λ_f. Theorem 1 characterizes subjective beliefs about the likelihood of events for individuals whose preferences over acts satisfy P1, P3^Q, P4^Q, P5, P6^Q.

Theorem 1 *Suppose a preference relation $\succ$ over $\mathcal{F}$ satisfies P1 (Ordering), P3^Q (Pivotal Monotonicity), P4^Q (Comparative Probability), P5 (Nondegeneracy), and P6^Q (Event Continuity). Then,*

A. There exists a unique, finitely additive, convex-ranged probability measure π with respect to which the relation $\succ$ is probabilistically sophisticated if and only if it is not extreme.

B. If the relation $\succ$ is extreme, there exists a set of nonatomic capacities $\Lambda(\mathcal{E})$ on $\mathcal{E}$, such that the condition (24) holds for any capacity $\lambda \in \Lambda(\mathcal{E})$.

The result reveals two interesting behavioral characteristics of beliefs underlying the choice of quantile maximizers. Theorem 1 first unveils that beliefs enter differently into the decision-making of agents with extreme versus non-extreme preferences. The theorem identifies a condition on preferences that satisfy axioms P1, P3^Q, P4^Q, P5, P6^Q under which quantile maximizers, like expected utility maximizers, behave as if they based their choice on a unique probability measure. This is the case as long as their preferences are not extreme. The result further asserts that the beliefs of maxmin and -max agents although not additive are nonetheless monotone with respect to event inclusion, and the agents can hence distinguish among such events whenever the event differences are nonnull. Thus, the choices of maxmin and -max decision makers reflect more from their beliefs than merely whether events (and their complements in $\mathcal{S}$) are null or not. In Section 5, we explain at

[18] It was first formalized by Grant (1995). Chew and Sagi (2006) have employed it to provide the most general characterization of probabilistic sophistication to date.

preference level what engenders the differences in the properties of the probability measures representing beliefs of individuals with the non-extreme compared to the extreme preferences.

Discussion. Theorem 1 has two more general implications for modeling probabilistic sophistication as such. First, the qualitatively different properties of beliefs of extreme and non-extreme quantile maximizers (additivity, uniqueness, and convex-rangedness) invite a question about whether the definition of probabilistic sophistication should not be strengthened. Note carefully that for agents with the the extreme preferences, condition (24), increasingly used in the literature to define probabilistically sophisticated agents, is satisfied even if each act being compared is evaluated through a different (and non-additive) measure in $\Lambda(\mathcal{E})$.[19] Furthermore, we establish that the preferences of the extreme-quantile maximizers remain intact whenever outcomes are exchanged between arbitrary disjoint non-null events, and not only equally likely events (Section 5). Should probabilistic sophistication permit that? Perhaps, its definition ought to require both additivity and uniqueness of the measure representing beliefs?[20]

Second, as we discuss in the next section, nothing in the conditions imposed on preferences $\succ$ over acts in $\mathcal{F}$ in Theorem 1 implies that there is a real-valued utility index providing them with a numerical representation. This is in stark contrast with the expected utility model, and indeed, with any cardinal one. Formally, the presence of P6-like axiom imposes a restriction on the set of outcomes $\mathcal{X}$ ($\succ$-denseness of a countable subset) which,

[19] In fact, the preferences of extreme (as well as not) quantile maximizers also satisfy a stronger version of probabilistic sophistication: Define a preference relation over lotteries, $\succsim_P$, induced from the underlying preferences over acts $\succsim$:

$$\text{If } \pi \circ f^{-1} = P,\ \pi \circ g^{-1} = Q \text{ for some } f, g \in \mathcal{F}, \text{ then } (f \succsim g \Rightarrow P \succsim_P Q). \tag{25}$$

Then, $\succsim$ is probabilistically sophisticated if there exists a measure μ on the set of events, inducing a relation $\succsim_P$ over lotteries, such that for all P, Q in $\mathcal{P}_0(\mathcal{X})$, and all f, g in $\mathcal{F}$

$$\left(P \succsim_P Q,\ \mu \circ f^{-1} = P,\ \mu \circ g^{-1} = Q\right) \Rightarrow f \succsim g. \tag{26}$$

This characteristic of preferences over acts $\succsim$ entails that they can be recovered from the knowledge of μ and lottery preferences $\succsim_P$ alone. It is straightforward to show that the stronger definition is equivalent with (24) if the relation $\succ$ is a weak order and π is convex-ranged (so that the mapping from $\mathcal{F}$ to $\mathcal{P}_0(\mathcal{X})$ is onto), which is the case in our model for non-extreme preferences $\succ$. The above formulation was proposed by Grant (1995), who also demanded uniqueness of μ.

Although the choices of agents with extreme preferences are consistent with a set of beliefs, $\Lambda(\mathcal{E})$, one can recover the agents' entire preference relation over all acts, even with measures that are not only not additive but also not convex-ranged. The knowledge of (one measure from) that set and the lottery preferences suffices (see Appendix 1C).

[20] Even though additivity has no bite for the beliefs of maxmin and -max individuals, there does exist an additive measure representing their beliefs. If condition (24) was adopted to define probabilistic sophistication with the additional requirement that π be unique, the extreme preferences (and only these preferences in the quantile model) would not be probabilistically sophisticated. While established in Machina and Schmeidler's (1992) and Chew and Sagi's (2006) results, uniqueness is not required in their definitions of probabilistic sophistication.

along with a weak-order structure, is then equivalent to the existence of a real-valued index on the set $\mathcal{X}$ (Debreu (1954)). Our derivation of beliefs does not impose any conditions on the outcome set. Instead, we can derive a real-valued probability measure (and a real number $\tau \in [0,1]$) without having a numerical representation of preferences. One benefit offered by our technique (presented in Section 5) is that it can be used to characterize beliefs of agents without well-defined utility functions. Unlike Machina and Schmeidler (1992) or Grant (1995), we can, for example, derive beliefs for lexicographic agents.[21]

4.2 Representation Results

Quantile Maximization with Respect to Probabilities. We now present the second main result of the paper, a complete characterization of choice behavior of quantile maximizers. Theorem 2 states that the preferences of a quantile maximizer satisfy axioms P1, P3^Q, P4^Q, P5, and P6^Q, and conversely, an individual whose preferences conform to those axioms can be viewed as a quantile maximizer.

As noted in the previous section, one novel aspect of our axiomatization is that the axioms do not impose any structure on the set of outcomes $\mathcal{X}$. To assure that the representation is numerical, in the next two theorems, we add the condition that $\mathcal{X}$ contains a countable $\succ$-order dense subset.[22] Without it, Theorems 2 and 3 could be re-cast in terms of quantiles of distributions of outcomes x rather then payoffs $u(x)$: the axioms are necessary and sufficient for there to exist a unique number τ in $[0,1]$ and probability measure(s) representing beliefs that guide the behavior of a quantile maximizer.

Theorem 2 *Consider a preference relation $\succ$ over $\mathcal{F}$. The following are equivalent:*

(1) $\succ$ satisfies: P1 (Ordering), P3^Q (Pivotal Monotonicity), P4^Q (Comparative Probability), P5 (Nondegeneracy), and P6^Q (Event Continuity).

(2) There exist:

(i) a unique number $\tau \in [0,1]$;

(ii) a probability measure π for $\tau \in (0,1)$ and a set of capacities $\Lambda(\mathcal{E})$ for $\tau \in \{0,1\}$,

[21] It might be worthwhile putting our results and the weakening of Savage's P6 in perspective. Motivated by an observation that Savage's (1954) derivation of subjective probabilities depends on axioms that lead to an expected-utility functional, Machina and Schmeidler (1992) characterized an individual whose choice is based on probabilistic beliefs, but does not necessarily comply with the Expected Utility hypothesis. Grant (1995) observed that Machina and Schmeidler's definition and proof still restrict a class of preferences by requiring that the induced lottery preferences satisfy continuity and monotonicity properties (mixture continuity and monotonicity with respect to stochastic dominance). Grant (1995) postulated that the notion of probabilistic sophistication as such should be dissociated from extraneous properties of the induced lottery preferences, and hence of the utility representation of preferences. Nevertheless, although Grant's (1995) derivation drops Savage's monotonicity, it does use a weaker continuity property (two-outcome mixture continuity). He concludes: "Ideally then, it would be nice to characterize probabilistically sophisticated preferences without requiring the induced risk preferences to exhibit any specific properties save perhaps some form of continuity ..." (p.177). Our characterization of probabilistic sophistication achieves that. In fact, the represenation functional cannot be mixture continuous in an ordinal model like ours.

[22] Natural examples include $\mathcal{X}$ being finite, countably infinite, $\mathcal{X} = \mathbb{R}$.

as characterized in Theorem 1;

(iii) a utility function on $\mathcal{X}$, u, which represents $\succ_x$, where u is unique up to strictly increasing transformations;

such that the relation $\succ$ over acts can be represented by the preference functional $\mathcal{V}(f) : \mathcal{F} \to \mathbb{R}$ given by

$$\begin{aligned} \mathcal{V}(f) &= Q^{\tau}(\Pi_f) \text{ if } \tau \in (0,1); &(27)\\ \mathcal{V}(f) &= Q^{\tau}(\Lambda_f) \text{ for any } \lambda \in \Lambda(\mathcal{E}) \text{ if } \tau \in \{0,1\}. &(28)\end{aligned}$$

The choice mechanism is thus decomposed into two factors: τ which is assured to be unique, and a probability measure, unique for all $\tau \in (0,1)$; a set of monotone measures represents beliefs held by quantile maximizers with $\tau = 0$ or $\tau = 1$.

Quantile Maximization with Respect to Capacities. In the characterization of preferences through Theorems 1 and 2, the additivity axiom (Comparative Probability, P4^Q) has no implications for the derived representation of beliefs of maxmin and -max decision makers. The next theorem establishes that when P4^Q is dispensed with, one can still uniquely pin down a number $\tau \in [0,1]$ and that, for any such τ, agent's beliefs can then be represented by capacities. Thus, the remaining four conditions axiomatize Quantile Maximization with respect to non-additive measures. A decision maker chooses as if she were evaluating an act $f \in \mathcal{F}$ by $Q^{\tau}(\Lambda_f)$ for some $\lambda \in \Lambda(\mathcal{E})$, with $\Lambda(\mathcal{E})$ being the set of representing capacities.[23,24]

Theorem 3 *Consider a preference relation $\succ$ over $\mathcal{F}$. The following are equivalent:*

(1) $\succ$ satisfies: P1 (Ordering), P3^Q (Pivotal Monotonicity), P5 (Nondegeneracy), and P6^Q (Event Continuity).

(2) There exist:

(i) a unique number $\tau \in [0,1]$;

(ii) a set of nonatomic capacities $\Lambda(\mathcal{E})$ on $\mathcal{E}$;

(iii) a utility function on $\mathcal{X}$, u, which represents $\succ_x$, where u is unique up to strictly increasing transformations;

such that the relation $\succ$ over acts can be represented by the preference functional $\mathcal{V}^C(f) : \mathcal{F} \to \mathbb{R}$ given by

$$\mathcal{V}^C(f) = Q^{\tau}(\Lambda_f) \text{ for any } \lambda \in \Lambda(\mathcal{E}). \qquad (29)$$

[23]Since we prove it through a different argument, the result is not stated as a direct corollary of Theorem 2.

[24]For $\tau \in (0,1)$ the number τ and the set $\Lambda(\mathcal{E})$ are unique as a pair.

Again, note that condition (24) holds, even if each act being compared is evaluated through a different (and non-additive) measure in $\Lambda(\mathcal{E})$. The earlier discussion questioning the aptness of (24) to capture probabilistic sophistication thus extends to all τ-maximizers, $\tau \in [0,1]$.

5 Sketch of the Proofs of Theorems 1 and 2

This section lays out in detail our axiomatization of subjective probabilities and Quantile Maximization (Theorems 1 and 2). Our goal is to separate beliefs from preferences, thus establishing probabilistic sophistication and, after we derive a measure-representation for these beliefs, to separate a quantile from probability distributions.

The proof proceeds as follows: The first step (Theorem 1) is to separate beliefs from preferences over $\mathcal{F}$, $\succ$. Formally, that step involves deriving the likelihood ranking revealed by the preference relation $\succ$, and showing that it can be represented by a probability measure π on $\mathcal{E}$. This part is the heart of the proof. Next (Theorem 2), we demonstrate that there exists a unique number τ in the unit interval $[0,1]$, such that an individual is indifferent between two acts, f and g, if and only if she is indifferent between the τ^{th}- quantile outcomes of the cumulative probability distributions induced by π and the acts f, g. To make the constructed representation of the relation $\succ$ on $\mathcal{F}$ numerical, we further derive (a family of) utility functions u over certain outcomes. Finally, we establish that axioms P1, P3^Q, P4^Q, P5, and P6^Q on the relation $\succ$, shown to be sufficient for obtaining the representation, are also necessary. The proof sketch presented here focuses on the sufficiency part.

Probability Measures. We begin by explaining why Savage's (1954) construction cannot be directly used in the Quantile Maximization model. (Section 7 clarifies why Machina and Schmeilder's [1992] and Grant's [1995] derivations cannot be applied either.) Thereby, we identify the central difficulty in constructing a probability measure in the present setting.

Briefly, in order to derive a probability measure representation for beliefs under the Expected Utility model, Savage proceeded as follows:[25] He first defined a likelihood relation – the binary relation $\succ^*$ over events in $\mathcal{E}$, formulated in (22) – induced from preferences $\succ$ over acts in $\mathcal{F}$. Savage then showed that axioms P1-P6, satisfied by the binary relation over acts, imply conditions on the likelihood relation that are necessary and sufficient for the likelihood relation to admit a unique probability-measure that (i) represents it and (ii) is convex-ranged.[26] These conditions are:

[25] The description follows Fishburn's (1970) exposition of Savage's result.

[26] Respectively, (i) $E \succ^* F$ if and only if $\pi(E) > \pi(F)$, for any $E, F \in \mathcal{E}$, and (ii) for any $E \in \mathcal{E}$, and any $\rho \in [0,1]$, there is $G \subseteq E$, such that $\pi(G) = \rho \cdot \pi(E)$. While equivalent to nonatomicity for countably additive measures, convex-rangedness is stronger for finitely additive measures, which are derived in this paper. (See Bhaskara Rao and Bhaskara Rao [1983], Ch. 5.)

A1 $\varnothing \not\succ^* E$.
A2 $\mathcal{S} \succ^* \varnothing$.
A3 $\succ^*$ *is a weak order.*
A4 $(E \cap G = F \cap G = \varnothing) \Rightarrow (E \succ^* F \Leftrightarrow E \cup G \succ^* F \cup G)$.
A5 P6$^{Q^*}$.

In the Quantile Maximization model, however, the relation $\succ^*$ does not satisfy the above set of axioms. Specifically, what fails for all $\tau \in (0,1)$ is axiom A4, which is key to establishing the additivity of the probability measure. For $\tau \in \{0,1\}$, A4 is vacuous under the relation $\succ^*$. In addition, A5 fails for $\tau \in (0,1]$, but this problem disappears when our Eventwise Continuity (P6^Q) is used instead.

A5' P6^Q

What underlies the failure of A4 is that the commonly used likelihood relation $\succ^*$ does not discriminate well between events from $\mathcal{E}$, as we now make precise. Consider a median maximizer ($\tau = 0.5$), and suppose that there does exist a probability measure, π, that represents her beliefs. Suppose further that she compares events E and F, such that $\pi(E) = 0.3$ and $\pi(F) = 0.2$. Given the measure π, each act in the definition of $\succ^*$, (22), induces a probability distribution. When the median maximizer compares these distributions, she ranks them as indifferent. Yet what this means in terms of the relation $\succ^*$ is that the decision maker ranks events E and F as equally likely. In general, under τ-maximization, the relation $\succ^*$ ranks as equally likely all pairs of events with probabilities either both smaller than $1-\tau$ or both greater than $1-\tau$[27]; for $\tau = 1$ ($\tau = 0$), the likelihood of no events both being more likely than $\varnothing$ (less likely than $\mathcal{S}$) can be ranked strictly by $\succ^*$. The following lemma demonstrates how crude the relation $\succ^*$ is: There are only two equivalence classes in the collection $\mathcal{E}$ under $\sim^*$; all events in $\mathcal{E}$ are ranked as either equally likely to the null set $\varnothing$ or to the state space $\mathcal{S}$.

Lemma 2 *$E \succ^* \varnothing \Leftrightarrow E \sim^* \mathcal{S}$; $E \prec^* \mathcal{S} \Leftrightarrow E \sim^* \varnothing$.*

Therefore, even if there is a probability measure that represents the beliefs of a quantile maximizer, the relation $\succ^*$ will not allow us to recover that measure from a data set containing all choices among acts in all possible subsets of $\mathcal{F}$.

[27] Comparing events $E, F \in \mathcal{E}$ through the likelihood ranking of their complements, as well as directly (which already imposes some additivity), would still render "equally likely" all events E and F with probabilities $\pi(E)$ and $\pi(F)$ either both smaller than $\min\{\tau, 1-\tau\}$, or both greater than $\max\{\tau, 1-\tau\}$, or both greater than $\min\{\tau, 1-\tau\}$ and smaller than $\max\{\tau, 1-\tau\}$. That such double comparison would be consistent follows from Lemma 6 in Appendix 1.

Nonetheless, we show that the structure embedded in the preference relation over acts $\succ$ is rich enough to reveal the relative likelihoods of events that can be provided with a probability measure representation. Our approach is to construct a sub-collection of "small" events, $\mathcal{E}_{**} \subset \mathcal{E}$, and define a new binary relation on events, $\succ_{**}$, which although incomplete on $\mathcal{E}$, is complete on the sub-collection $\mathcal{E}_{**}$. We then derive a unique, convex-ranged, and additive measure that represents a decision maker's beliefs about the relative likelihoods of events in $\mathcal{E}_{**}$. Next, we build up from $\mathcal{E}_{**}$ to construct a likelihood relation that is complete and satisfies A1-A5' on the whole collection of events $\mathcal{E}$. Finally, we use the new likelihood relation to uniquely extend the measure derived on $\mathcal{E}_{**}$ to the collection $\mathcal{E}$ as well as to derive a unique number τ. As explained below, in the cases leading to $\tau \in \{0, 1\}$, the information contained in $\succ$ does not suffice to permit all these steps.

The sub-collection of "small" events, $\mathcal{E}_{**} \subset \mathcal{E}$, is defined to contain all events that are ranked by the relation $\succ_*$ as less likely than their complements:

$$\mathcal{E}_{**} \equiv \{E \in \mathcal{E} | E^c \succ_* E\}. \tag{30}$$

Why this construction is helpful and the make-up of events that the sub-collection consists of will become clear after we specify a likelihood relation. On the collection $\mathcal{E}_{**}$, we define a new binary relation $\succ_{**}$:

Definition 3 *Let $E, F \in \mathcal{E}_{**}$.*

$$E \succ_{**} F \text{ if } E \cup G \succ_* F \cup G \text{ for some event } G \in \mathcal{E}, \text{ such that } (E \cup F) \cap G = \varnothing. \tag{31}$$

The idea behind the new likelihood relation $\succ_{**}$ is as follows: The events that can be strictly ranked by $\succ_{**}$ are "small" in the sense that there exists an event G in their common complement, such that the unions $E \cup G$ and $F \cup G$ can be strictly ranked by $\succ_*$. Back to our earlier example, in which a median maximizer compares events E and F, such that $\pi(E) = 0.3$ and $\pi(F) = 0.2$, we ask how will the relation $\succ_{**}$ rank these events? Take an event G disjoint with both E and F and such that $\pi(G) = 0.25$. Adding G to events E and F will enlarge the magnitudes of probabilities in an additive way (which is to be established) and, thereby, switch the evaluation of the distribution induced by act $[x \text{ if } s \notin E \cup G; \ y \text{ if } s \in E \cup G]$ in the definition of $\succ_*$, (23), from outcome x to y[28], while maintaining the evaluation of the distribution induced by $[x \text{ if } s \notin F \cup G; \ y \text{ if } s \in F \cup G]$) is unchanged at x.

[28] The earlier example referred to the (standard) relation $\succ^*$, whereas the new likelihood $\succ_{**}$ builds on the relation $\succ_*$. The change merely ensures left-continuity of the quantile representation to-be-derived. Yet, the logic behind the refined relation $\succ_{**}$ is intact. (Clearly, we could have stated the first example using $\succ_*$.)

For the relation $\succ_{**}$ to be well-defined, we need to show that an event G in (31) exists, and that there cannot be any other event G' for which the ranking is reversed. That is, for all $E, F \in \mathcal{E}_{**}$, if $E \succ_{**} F$, then $F \not\succ_{**} E$. Lemma 7 (in Appendix 1A) establishes the latter. Explaining why the former is true will also clarify the moniker "small": It is key to show that collection $\mathcal{E}_{**}$ consists of all (and only of) the events, such that for any $E, F \in \mathcal{E}_{**}$ there exists $G \subseteq (E \cup F)^c$ for which

$$E \cup G \succ_* F \cup G \text{ or } E \cup G \prec_* F \cup G \text{ or } E \cup G \sim_* F \cup G \sim_* \mathcal{S}. \tag{32}$$

(32) establishes the sense in which events in $\mathcal{E}_{**}$ can be compared through their complements by the relation $\succ_{**}$. In particular, the sub-collection $\mathcal{E}_{**}$ does not contain events for which $E \cup G \sim_* F \cup G \sim_* \varnothing$ for all events $G \subseteq (E \cup F)^c$.[29] Why this holds can be understood when we further characterize collection $\mathcal{E}_{**}$. We demonstrate that $\mathcal{E}_{**}$ is equal to the equivalence class that contains events that are ranked equally likely to the null set $\varnothing$ by one of the relations $\succ_*$ or $\succ^*$.

$$\mathcal{E}_{**} = \{E \in \mathcal{E} | E \sim_* \varnothing\} \text{ or } \mathcal{E}_{**} = \{E \in \mathcal{E} | E \sim^* \varnothing\}, \tag{33}$$

where "or" is meant exclusively. In hindsight, after the measure and τ are pinned down, one can show that the former corresponds to the representation with $\tau < 0.5$ while the latter corresponds to one with $\tau \geq 0.5$. Whether relation $\succ_*$ or $\succ^*$ is used is determined endogenously by a condition on preferences $\succ$, one that requires that the event G in (31) exists. More precisely, we establish that if there is at least one event E, such that $E \sim_* E^c \sim_* \varnothing$, then all events F in the equivalence class generated by $\varnothing$ under $\succ^*$ are such that $F \prec^* F^c$, and hence do satisfy the defining condition for the collection $\mathcal{E}_{**}$, (30). Otherwise, all events in the equivalence class generated by $\varnothing$ under $\succ_*$ satisfy the condition.[30] Thus, (32) holds, as desired.

In summary, we construct a collection of events $\mathcal{E}_{**}$ and a likelihood relation $\succ_{**}$, which is not only complete on $\mathcal{E}_{**}$ (though not on $\mathcal{E} \backslash \mathcal{E}_{**}$), but can also rank (all) events in collection $\mathcal{E}_{**}$ via their complements, in the sense of (32). The relations $\succ^*$ and $\succ_*$ should be interpreted as helpful in retrieving information from preferences in the construction of the collection $\mathcal{E}_{**}$. It is the relation $\succ_{**}$ that represents the likelihood ranking of events of a quantile maximizer.

Now comes an important difference between the extreme and not extreme preferences. In deriving the measure representation, it is essential that disjoint non-null subsets of the state

[29] This also ensures that the relation $\sim_{**}$ is meaningful.

[30] Observe that the collection $\mathcal{E}_{**}$, defined as $\{E \in \mathcal{E} | E^c \succ_* E\}$, coincides with $\{E \in \mathcal{E} | E^c \succ^* E\}$, by the definitions of $\succ^*$ and $\succ_*$ in (22) and (23).

space can be strictly ranked. This cannot be assured when preferences are extreme. For that case, we show that a decision maker's preferences over acts only depend on (and thus can only reveal) whether an event is null, or it is the state space or nested in another event, all up to differences on null sub-events. Therefore, when preferences are extreme, there cannot exist an event in the common complement of any two disjoint non-null events, so that they can be strictly ranked by $\sim_{**}$. Thus, intuitively, while all τ-maximizers, $\tau \in [0,1]$, can compare nested events, these are the only events that can be strictly ranked by 0- and 1-maximizers. It is at this point that the derivation of beliefs for $\tau = 0$ and $\tau = 1$ departs from the general proof. Specifically, we establish the following invariance properties of the extreme and non-extreme preferences:

Lemma 3

A. If the binary relation over acts, $\succ$, is not extreme, then for all events $E, F \in \mathcal{E}_{**}$, such that $E \sim_{**} F$, and all acts $h \in \mathcal{F}$,

$$\begin{bmatrix} x \text{ } \textit{if } s \in E \\ y \text{ } \textit{if } s \in F \\ h \text{ } \textit{if } s \notin E \cup F \end{bmatrix} \sim \begin{bmatrix} x \text{ } \textit{if } s \in F \\ y \text{ } \textit{if } s \in E \\ h \text{ } \textit{if } s \notin E \cup F \end{bmatrix}. \tag{34}$$

B. If the binary relation over acts, $\succ$, is extreme, then for all non-null events $E, F \in \mathcal{E}_{**}$, $E \cap F = \varnothing$, and all acts $h \in \mathcal{F}$,

$$\begin{bmatrix} x \text{ } \textit{if } s \in E \\ y \text{ } \textit{if } s \in F \\ h \text{ } \textit{if } s \notin E \cup F \end{bmatrix} \sim \begin{bmatrix} x \text{ } \textit{if } s \in F \\ y \text{ } \textit{if } s \in E \\ h \text{ } \textit{if } s \notin E \cup F \end{bmatrix}. \tag{35}$$

That is, decision makers with the non-extreme preferences are indifferent to exchanging outcomes on events equally likely according to $\sim_{**}$. Individuals whose preferences are extreme are indifferent to exchanging outcomes on disjoint, non-null events.

Furthermore, for the extreme preferences, it is essential to combine strict likelihood judgments from $\succ^*$ and $\succ_*$, since in that case, it is the extended definition that makes it possible to distinguish the likelihoods of $\varnothing$, $\mathcal{S}$ and events E, which differ from $\varnothing$ and $\mathcal{S}$ on a non-null set. The combined judgment,[31] demonstrated to be consistent in Lemma 6 (Appendix 1A), states that E is more likely than F if for all $x \succ y$,

$$\begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \notin F \\ y \text{ if } s \in F \end{bmatrix} \quad or \quad \begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \notin F \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix}. \tag{36}$$

[31] Using definitions (22) and (23), (36) can be re-stated as: E is more likely than F if $E \succ^* F$ or $E \succ_* F$. One could employ that definition for all preferences, extreme and not extreme. As we explain above, this is not necessary to characterize beliefs fully for non-extreme relation $\succ$ on $\mathcal{F}$, for which a combined judgment would generate three equivalence classes on $\mathcal{E}$.

For the non-extreme preferences, we show that axioms A1, A3, A4 and A5' hold on $\mathcal{E}_{**}\cup\overline{\mathcal{E}}$, where $\overline{\mathcal{E}}=\{E\in\mathcal{E}|\nexists F \text{ nonnull: } E\backslash F\in\mathcal{E}\backslash\mathcal{E}_{**}\}$,[32] to derive a unique, finitely additive and convex-ranged probability measure. Intuitively, collection $\mathcal{E}_{**}\cup\overline{\mathcal{E}}$ contains events whose probabilities will not be greater than $\min\{\tau,1-\tau\}$. We next show that any event from the complement of sub-collection $\mathcal{E}_{**}\cup\overline{\mathcal{E}}$ in $\mathcal{E}$ can be partitioned into events from $\mathcal{E}_{**}$. Using the properties of the measure derived on $\mathcal{E}_{**}\cup\overline{\mathcal{E}}$, we then uniquely extend the measure to all events in $\mathcal{E}$. Turning to the extreme preferences, since they only distinguish between $\varnothing$, $\mathcal{S}$ and pairs of nested events, up to differences on null sub-events, any (normalized) measure that respects monotonicity w.r.t. events will still represent the same preferences. Therefore, an extreme preference $\succ$ on $\mathcal{F}$ is consistent with a set of measures that are nonatomic but not necessarily finitely additive.

Quantile. Having derived the probability measures, we recover from preferences a unique number $\tau\in[0,1]$ that corresponds to the quantile being maximized. The following result, which relies on Lemma 3, provides the key assertion:

Lemma 4 *In the coarsest measurable partition of the state space $\mathcal{S}$ induced by act $f\in\mathcal{F}$, there is a unique pivotal event.*

With the derived measures, we can map the set of acts $\mathcal{F}$ onto the set of simple lotteries, $\mathcal{P}_0(\mathcal{X})$, through the mapping defined in Section 4.1. Then, to establish the existence and uniqueness of τ for the non-extreme preferences, we construct a sequence of equi-partitions of $\mathcal{S}$ (finite partitions whose elements are equally likely) with 2^N elements. In the $(N+1)^{th}$ partition, each element of the N^{th} partition is split into 2 equally likely elements. Next, we construct a collection of events that associates with each $N\in\mathbb{N}_{++}$ an event $\bigcup\limits_{l=1,\dots,n(N)} F_l^{2^N}$, such that[33]

$$\begin{bmatrix} x \text{ if } s\in\mathcal{S} \\ y \text{ if } s\notin\mathcal{S}\end{bmatrix}\sim\dots\sim\begin{bmatrix} x \text{ if } s\notin \bigcup\limits_{l=1,\dots,n(N)-1} F_l^{2^N} \\ y \text{ if } s\in \bigcup\limits_{l=1,\dots,n(N)-1} F_l^{2^N}\end{bmatrix}\succ\begin{bmatrix} x \text{ if } s\notin \bigcup\limits_{l=1,\dots,n(N)} F_l^{2^N} \\ y \text{ if } s\in \bigcup\limits_{l=1,\dots,n(N)} F_l^{2^N}\end{bmatrix}\sim\dots\sim\begin{bmatrix} x \text{ if } s\notin\mathcal{S} \\ y \text{ if } s\in\mathcal{S}\end{bmatrix}. \tag{37}$$

We then obtain $\tau\in(0,1)$ by approximating it with the probabilities of the union of the pivotal events $\left\{\bigcup\limits_{l=1,\dots,n(N)} F_l^{2^N}\right\}_{N\in\mathbb{N}_{++}}$.

Utilities. Given that $\succ$ is a weak order (P1), due to the ordinality property of the quantile-maximization representation, the utility on outcomes, u, depends exclusively on

[32] Adding $\overline{\mathcal{E}}$ assures that there are events in $\mathcal{E}_{**}$ that serve the role of the state space in an appropriate counterpart of A2.

[33] A corollary of Lemma 4 is used to establish that there is a unique $n(N)\in\{1,\dots,2^N\}$ for which (37) holds.

the properties of the set $\mathcal{X}$. The assumption that $\mathcal{X}$ contains a countable $\succ$-dense subset (employed only in the final step) can be used, together with the weak order, to apply Debreu's (1954) theorem and derive a real-valued utility index on $\mathcal{X}$. We note that construction of the numerical representation functional for $\succ$ does not depend on the existence of the best and worst outcomes – again, the reason is ordinality.

6 Related Literature

In this section, we relate our results to the literature on probabilistic sophistication: Machina and Schmeidler (1992), Grant (1995) and a recent contribution by Chew and Sagi (2006). This comparison serves two purposes. First, it elaborates on why the new approach to axiomatizing beliefs of quantile maximizers is required, and second, it elucidates the distinct properties of the quantile model.

Machina and Schmeidler (1992, hereafter MS). Setting as their goal the liberation of the derivation of subjective probability in the Savage world from the Expected Utility hypothesis, MS drop P2 (the Sure-Thing Principle). Dispensing with P2, however, removes more than the Marschak-Samuelson independence, associated with an expected-utility functional. As mentioned above, Savage used P2 to obtain additivity of the probability measure. To restore additivity, MS strengthen P4 (Weak Comparative Probability, to Strong Comparative Probability, P4*), otherwise, using Savage's axioms.

The preferences of the quantile maximizers do not satisfy P4*.[34] Overall, our set of conditions shares with those of MS merely P1 and P5. The functional form in the representation theorem of MS is mixture continuous and monotonic with respect to first-order stochastic dominance. While encompassing many functional forms, the conditions on the relation $\succ$ underlying these properties of the representation functional are crucially used in the derivation of probabilities. Our proof does not rely on any form of mixture continuity and monotonicity with respect to stochastic dominance holds only weakly (i.e., strict first-order stochastic dominance implies only a weak preference over distributions). Even more, mixture continuity and the stronger version of monotonicity do not hold in an ordinal model.

MS can essentially use Savage's (1954) derivation of probability, as they observe that although in general P2 fails in their case, it does hold for two-outcome acts. Since all axioms of Savage hold for such gambles, it follows from Savage's theorem that there exists a unique, finitely additive and convex-ranged probability measure. MS then apply P4* to extend the

[34]The following example illustrates that for P4*. Let $\tau = \frac{1}{3}$, $\pi(E) = \frac{1}{3} + \varepsilon$, $\pi(F) = \frac{1}{3} - \varepsilon$, $\pi(G) = \frac{1}{3}$. Then $\begin{bmatrix} 2 \text{ if } E \\ 1 \text{ if } F \\ 5 \text{ if } G \end{bmatrix} \succ \begin{bmatrix} 1 \text{ if } E \\ 2 \text{ if } F \\ 5\text{if } G \end{bmatrix}$ but $\begin{bmatrix} 2 \text{ if } E \\ 1 \text{ if } F \\ 0 \text{ if } G \end{bmatrix} \sim \begin{bmatrix} 1 \text{ if } E \\ 2 \text{ if } F \\ 0 \text{ if } G \end{bmatrix}$.

measure to the set of all acts.

Grant (1995). With a novel interpretation of probabilistic sophistication, Grant (1995) obtains a derivation of probabilities without P2 and P3 (Eventwise Monotonicity). After relaxing the latter condition, two-outcome gambles cannot be used to infer the relative likelihood of events. (The ranking of constant acts need not agree with the conditional ranking of two outcomes.) Still, with a modification of P3 to Conditional Upper or Lower Eventwise Monotonicity ($\mathrm{P3}^{CU}$, $\mathrm{P3}^{CL}$), the conditional preference between two outcomes can be used to draw an inference about the likelihood of events from the preference over conditional gambles that involve these two outcomes. This identifies a set of acts on which the hypotheses of MS (1992, Theorem 1) hold; hence a probability-measure representation can be obtained. The measure is then extended through continuity of preferences ($\mathrm{P6}^{\dagger}$) to the whole state space.

Our lottery preferences need not satisfy either of Grant's (1995) two-outcome mixture continuity[35] or conditional monotonicity.

Chew and Sagi (2006, hereafter CS). The beautiful approach proposed recently by Chew and Sagi (2006) is based on the notion of exchangeability. Two events are said to be *exchangeable* if the agent is always indifferent to permuting the payoffs assigned to these events.

Definition 4 *For any pair of disjoint events $E, F \in \mathcal{E}$, E is exchangeable with F if, for any outcomes $x, y \in \mathcal{X}$, and any act $f \in \mathcal{F}$,*

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \in F \\ f \text{ if } s \notin E \cup F \end{bmatrix} \sim \begin{bmatrix} y \text{ if } s \in E \\ x \text{ if } s \in F \\ f \text{ if } s \notin E \cup F \end{bmatrix}. \tag{38}$$

The relation of exchangeable events is then used to define the *comparability* relation, $\succsim^C$.

Definition 5 *For any events $E,\ F \in \mathcal{E}$, $E \succsim^C F$ whenever $E \backslash F$ contains a sub-event G that is exchangeable with $F \backslash E$.*

Intuitively, exchangeability carries the meaning of "equal likelihood," while comparability

[35] $V : P_0(X) \to \mathbb{R}$ is said to be *mixture continuous for two-outcome sub-lotteries* if for any pair of outcomes x, y in $\mathcal{X}$, any $\gamma \in (0,1]$, and any pair of lotteries P, Q, the sets $\{\lambda \in [0,1] |\ V(\gamma(\lambda\delta_x + (1-\lambda)\delta_y) + (1-\gamma)P) > V(Q)\}$ and $\{\lambda \in [0,1] |\ V(Q) > V(\gamma(\lambda\delta_x + (1-\lambda)\delta_y) + (1-\gamma)P)\}$ are open.

With a slight abuse of notation, we will denote by P a lottery in $\mathcal{P}_0(\mathcal{X})$ and the corresponding to it cumulative probability distribution.

$\succ_P$ is *mixture continuous for two-outcome sub-lotteries* if for all pairs of outcomes x, y in $\mathcal{X}$, all $\gamma \in (0,1]$, and distributions P, Q in $\mathcal{P}_0(\mathcal{X})$, the sets $\{\lambda \in [0,1] |\ \gamma(\lambda\delta_x + (1-\lambda)\delta_y) + (1-\gamma)P \succ_P Q\}$ and $\{\lambda \in [0,1] |\ Q \succ_P \gamma(\lambda\delta_x + (1-\lambda)\delta_y) + (1-\gamma)P\}$ are open.

conveys "greater likelihood." CS do find a set of axioms on those relations, so they can yield a likelihood relation that can be provided with a probability-measure representation.

While CS's Theorem 1 can also be used in our model,[36] it remains an open question how to derive beliefs without monotonicity and continuity of preferences without resorting to the exchangeability relation, but rather by directly defining a strict likelihood relation from preferences (as, e.g., in Machina and Schmeidler [1992], Grant [1995], and this paper). Why would such an argument be attractive? First, the link between the likelihood that CS construct and preferences over acts is only through the definition of exchangeable events, a pre-notion of "equally likely." In particular, completeness of the comparability relation (pre-notion of "more likely than") is assumed (via an axiom), and transitivity is proven without any recourse to the strict relation over acts, $\succ$. By contrast, in the direct-likelihood method, the strict "more-likely-than" relation is revealed by preferences over acts, from which it inherits its properties. Second, as CS point out, assuming independently from preferences over acts that the exchangeability-based likelihood is well defined might not be warranted. They illustrate this point using the example of "Machina's Mother" (Machina [1989]). The direct-likelihood relation, combined with the small-event approach used in the present paper, is (by construction) immune from that.[37]

7 Induced Lottery Preferences

This section takes an applied look at the quantile model. We show that Quantile Maximization stands out among ordinal choice rules that satisfy first order stochastic dominance as a unique such rule (Section 7.2); characterize risk attitudes of Quantile Maximizers (Section 7.3); in the context of an insurance problem, illustrate how the distinct properties of the model might be appealing in applied work (Section 7.4); and discuss in which applications the model might be attractive (Section 7.5). To elucidate the comparison of how the Subjective Expected Utility and Quantile Maximization perform in applications, we first delineate the differences in continuity and monotonicity properties of the induced lottery preferences between the two models (Section 7.1).

7.1 Properties of Lottery Preferences

Using Theorem 1, each act in $\mathcal{F}$ can be mapped to a lottery in $\mathcal{P}_0(\mathcal{X})$ via the mapping $f \to \pi \circ f^{-1}$. Consider the preference relation $\succsim_P$ over lotteries in $\mathcal{P}_0(\mathcal{X})$, as defined in Section 4.1.

[36] For non-extreme preferences, their axioms A, C and N follow from P1, P3^Q, P4^Q, P5, and P6^Q.

[37] For a derivation of beliefs in the "Machina's Mother" example through the direct-likelihood method, please e-mail the author.

(Lack of) Continuity. The implied lottery preferences of quantile maximizers are not mixture continuous,[38] not even for 2-outcome lotteries (as the lottery preferences are in Grant [1995]). Mixture continuity, typically implied by P6-like axiom, is too strong for an ordinal model, such as Quantile Maximization. Our weakening of P6 removes any continuity of mixture lotteries from risk preferences.

Monotonicity. Strong monotonicity with respect to first-order stochastic dominance (FOSD)[39] need not hold under Quantile Maximization. To characterize the appropriate monotonicity property for the lottery preferences, it is useful to find the analogs of axioms P1-P6^Q for lotteries, that is, for the case of a known and unique probability. We only need to look for the counterpart of Pivotal Monotonicity (P3^Q), which we call *Rankwise Monotonicity.* For a fixed lottery $P \in \mathcal{P}_0(\mathcal{X})$ and outcome x, which belongs to the support of P, $x \in supp\{P\}$, define the sums:

$$P_{x+} = \sum_{\{n|x_n \succ_P x\}} p_n, \; P_{x-} = \sum_{\{m|x_m \prec_P x\}} p_m \tag{40}$$

and let Q_{x+} and Q_{x-} be any sublotteries on P_{x+} and P_{x-} with supports, such that $supp\{Q_{x+}\} \succsim_P x$ and $supp\{Q_{x+}\} \precsim_P x$, respectively. The intuition behind the condition we state next is similar to that of P3^Q.[40]

$\overline{P3^Q}$ (**RANKWISE MONOTONICITY**): *For any simple lottery $P \in \mathcal{P}_0(\mathcal{X})$, there is an outcome $x \in supp\{P\}$, such that for all outcomes x, y, and $\lambda \in (0,1]$, there exists*

[38] $V : P_0(X) \to \mathbb{R}$ is said to be *mixture continuous* if for any lotteries P, Q and R in $\mathcal{P}_0(\mathcal{X})$, the sets $\{\lambda \in [0,1] | \; V(\lambda P + (1-\lambda)Q) > V(R)\}$ and $\{\lambda \in [0,1] | \; V(R) > V(\lambda P + (1-\lambda)P)\}$ are open.

$\succ_P$ is *mixture continuous* if for all distributions P, Q and R in $\mathcal{P}_0(\mathcal{X})$, the sets $\{\lambda \in [0,1] | \; \lambda P + (1-\lambda)Q \succ_P R\}$ and $\{\lambda \in [0,1] | \; R \succ_P \lambda P + (1-\lambda)Q\}$ are open.

[39] For an arbitrary outcome set $\mathcal{X}$, given a complete preorder over outcomes $\succsim_x$, $P = (x_1, p_1; ...; x_N, p_N)$ *weakly first order stochastically dominates* $Q = (y_1, q_1; ...; y_M, q_M)$ with respect to $\succsim_x$ if

$$\sum_{\{m|x \succsim_x x_m\}} q_m \geq \sum_{\{n|x \succsim_x x_n\}} p_n \text{ for all } x \in \mathcal{X} \tag{39}$$

and if, in addition, (39) holds with strict inequality for some $y \in \mathcal{X}$, then P *strongly first order stochastically dominates* Q with respect to $\succsim_x$.

It is said that $\succsim_P$ is weakly (strictly) *monotonic with respect to FOSD* if $P \succsim_P (\succ_P) Q$ whenever P strongly stochastically dominates Q.

[40] As shown in Appendix 4, Rankwise Monotonicity implies that for all simple lotteries $P \in \mathcal{P}_0(\mathcal{X})$, there is an outcome $x \in supp\{P\}$, such that for any outcome y, and sub-lotteries Q_{x-}, Q_{x+}, Q_{y-}, Q_{y+}:

$$x \succsim_x y \Leftrightarrow (Q_{x-}, P_{x-},; \; x, p_x; \; Q_{x+}, P_{x+}) \succsim_P (Q_{y-}, P_{x-}; \; y, p_x; \; Q_{y+}, P_{x+}) \,. \tag{41}$$

Cf. The *Ordinal Independence* axiom in Green and Jullien (1988) and *Irrelevance Axiom* in Segal (1989).

$\gamma \in [0,1]$, *such that*

$$x \succsim_x y \Leftrightarrow \gamma(1-\lambda)Q_{x-} + \lambda\delta_x + (1-\gamma)(1-\lambda)Q_{x+} \succsim_P \gamma(1-\lambda)Q_{y-} + \lambda\delta_y + (1-\gamma)(1-\lambda)Q_{y+}. \tag{42}$$

Say that ($\mathbf{L}^P$) holds if any lottery $P \in \mathcal{P}_0(\mathcal{X})$ is indifferent to the least preferred outcome in its support; that ($\mathbf{H}^P$) holds if any lottery $P \in \mathcal{P}_0(\mathcal{X})$ is indifferent to the most preferred outcome in its support; that $\succ_P$ is extreme if either ($\mathbf{L}^P$) or ($\mathbf{H}^P$) is satisfied. The counterpart of Event Continuity (P6^Q) reads:

$\overline{P6^Q}$ (**LEFT-CONTINUITY**) *If $\succ_P$ is non-extreme, then there exists an interval $[\gamma', \gamma'') \subseteq [0,1]$, such that (42) is satisfied for all $\gamma \in [\gamma', \gamma'')$. If ($\mathbf{L}^P$) holds, then (42) is satisfied if and only if $\gamma = 0$, and if ($\mathbf{H}^P$) holds, then (42), is satisfied if and only if $\gamma = 1$.*

It should be intuitive that the left-closedness of the interval of parameter γ will imply the left-continuity of the quantile representation for $\succ_P$. Formally, the following equivalence obtains:[41]

Proposition 1 *Assume P1 (Ordering) and P5 (Nondegeneracy), and that $\succ$ is probabilistically sophisticated with respect to π. Then,*

(*i*) $\succ$ *satisfies P3^Q if and only if $\succ_P$ exhibits $\overline{P3^Q}$;*
(*ii*) $\succ$ *satisfies P3^Q and P6^Q if and only if $\succ_P$ exhibits $\overline{P3^Q}$ and $\overline{P6^Q}$.*

By $\overline{P1}$ and $\overline{P5}$ denote the straightforward counterparts of weak order (P1) and nondegeneracy (P5) defined in Section 3 for the binary relation $\succ_P$. For a given lottery $P \in \mathcal{P}_0(\mathcal{X})$ and utility on outcomes u, let $Q^\tau(P)$ be the τ^{th} quantile of the cumulative probability dis-

[41] This characterization parallels the result of Grant, Kajii and Polak (1992), who established the equivalence between the monotonicity axiom of Savage (P3) and the condition (*Axiom of Degenerate Independence*):

For all simple lotteries $P \in \mathcal{P}_0(\mathcal{X})$, outcomes x, y and $\alpha \in (0,1)$,

$$x \succsim_x y \Leftrightarrow \alpha\delta_x + (1-\alpha)P \succsim_P \alpha\delta_y + (1-\alpha)P.$$

This axiom states that moving a probability mass from one outcome to another is weakly preferred (according to the induced lottery preferences) if and only if the second outcome is preferred to the first. Quantile Maximization requires that this holds only for outcomes on pivotal events.

Given that, for the restricted (compared to P3) class of sub-acts and events in condition P3^Q, the implications of Savage's P2 and P3 are equivalent, it is interesting to identify the substitution property of the induced lottery preferences that they jointly engender. Clearly, the induced lottery preferences need not obey the *Independence Axiom.* Since P3 fails, the lottery preferences need not exhibit ADI – the substitution axiom of Grant, Kajii and Polak (1992). Incidentally, failure of the latter provides one way to show that strict monotonicity w.r.t. first-order stochastic dominance does not hold in the quantile model; $\succsim_P$ respects ADI if and only if it satisfies first-order stochastic dominance. (See, for example, Grant [1995].)

tribution corresponding to P.[42] The following Corollary of Theorem 2 axiomatizes Quantile Maximization under risk.

Corollary 1 *A binary relation on the set of lotteries $\mathcal{P}_0(\mathcal{X})$ satisfies $\overline{P1}$ (Ordering), $\overline{P3^Q}$ (Rankwise Monotonicity), $\overline{P5}$ (Nondegeneracy), and $\overline{P6}$ (Left-Continuity) if and only if there exists a unique number $\tau \in [0,1]$ and a function $u : \mathcal{X} \to \mathbb{R}$, such that the relation $\succ_P$ over simple probability distributions can be represented by the preference functional $W : \mathcal{P}_0(\mathcal{X}) \to \mathbb{R}$ given by*

$$W(P) = Q^{\tau}(P). \tag{43}$$

where u is unique up to strictly increasing transformations.

7.2 Quantile Maximization As a Unique Ordinal Choice Rule

We now build on Corollary 1 to demonstrate that although by itself monotonicity with respect to first-order stochastic dominance does not imply Quantile Maximization, assuming $\overline{P1}$, $\overline{P5}$ and $\overline{P6^Q}$, it is, however, equivalent to Rankwise Monotonicity if one also requires that the following condition of *Ordinal Invariance* holds: Write each distribution $R \in \mathcal{P}_0(\mathcal{X})$ as a pair of ordered vectors of outcomes and probabilities $(\mathbf{x}_R, \mathbf{p}_R)$; for all pairs $P, Q \in \mathcal{P}_0(\mathcal{X})$,

$$(\mathbf{x}_P, \mathbf{p}_P) \succsim_P (\mathbf{x}_Q, \mathbf{p}_Q) \Leftrightarrow (\phi \circ \mathbf{x}_P, \mathbf{p}_P) \succsim_P (\phi \circ \mathbf{x}_Q, \mathbf{p}_Q) \tag{44}$$

for any mapping $\phi : \mathcal{X} \to \mathcal{X}$, such that if $x \succ_x y$ then $\phi \circ x \succ_x \phi \circ y$, where $\phi \circ \mathbf{x}_R$ is defined element by element.

Proposition 2 *Assume $\overline{P1}$ (Ordering), $\overline{P5}$ (Nondegeneracy), and $\overline{P6^Q}$ (Left-Continuity) hold. The following axioms are equivalent for a binary relation on the set of lotteries $\mathcal{P}_0(\mathcal{X})$, $\succ_P$:*

(1) $\overline{P3^Q}$ (Rankwise Monotonicity);

(2) Weak monotonicity with respect to FOSD and Ordinal Invariance.

The implication of Proposition 2 is threefold. First, it provides an equivalent to Corollary 1 characterization of our model in risk settings. Second, it justifies Quantile Maximization as a unique ordinal decision rule (that is, a rule satisfying ordinal invariance). Alternatively, what can perhaps be more directly seen from the equivalence established jointly by Lemmas 1 and 8, Quantile Maximization is the unique ordinal decision rule in the class of

[42] Again, to assure that the representation is real-valued, we assume that $\mathcal{X}$ contains a countable $\succ_x$-dense subset.

probabilistically sophisticated[43] rules for uncertain environments.[44] Finally, Proposition 2 shows that the notion of monotonicity we used in deriving probabilities is, in a sense, the weakest monotonicity property. While monotonicity w.r.t. FOSD is not a tight notion of monotonicity in that it is not equivalent to Rankwise Monotonicity, a local version of it which compares τ^{th} quantiles exclusively is tight in the presence of other axioms.[45]

7.3 Risk Attitudes

In applications, one would like to be able to characterize the attitudes of the quantile maximizers toward risk. Clearly, in the quantile model, characterization of risk attitudes through concavity of utility functions is no longer available. Do quantile maximizers then exhibit any consistent attitudes towards risk?[46] We now show that the model admits a notion of comparative risk attitude that also permits a complete ranking of agents. One needs two definitions: Risk and risk attitude. Say that distribution $P \in \mathcal{P}_o(\mathcal{X})$ is *more risky than* distribution $Q \in \mathcal{P}_o(\mathcal{X})$ if Q crosses P from below; that is, there exists $u(x) \in \mathbb{R}$, such that (i) $Q(u(y)) \leq P(u(y))$ for all y, such that $u(y) < u(x)$ and (ii) $Q(u(y)) \geq P(u(y))$ for all y, such that $u(y) > u(x)$. Consider the class of all pairs of distributions with the single-crossing property, $\mathcal{SC} = \{(P,Q) \in \mathcal{P}_o(\mathcal{X}) \times \mathcal{P}_o(\mathcal{X}) : Q \text{ crosses } P \text{ from below}\}$. Individual 1 is *more risk averse than* individual 2 if, for all pairs of distributions $(P,Q) \in \mathcal{SC}$, whenever 2 weakly prefers a less risky distribution, so does 1.
With these two definitions, one can show that τ itself provides a measure of risk attitude.

Proposition 3 *In the Quantile Maximization model, $\tau' < \tau$ if and only if τ'-maximizer is weakly more risk averse than the τ-maximizer.*

[43] And satisfying the weakest monotonicity for distributions (defined for a given τ^{th} quantile and a fixed $\tau \in [0,1]$).

[44] To aid intuition, we note the relation to results in social choice literature, which our axiomatization yields as a special case. Corollary 1 generalizes the model of *rank-dictatorship* (Gevers [1979]), also known as *positional dictatorship* (Roberts [1980]), which predicts that the person with the k^{th} level of wealth in a society will be a dictator. The social-choice result, which obtains here for the uniform distribution, also characterizes choice behavior based on the ranking of outcome vectors (e.g., order statistics). Our result for risk settings can be usefully interpreted for classes of citizens, ranked according to wealth.

[45] Even if the two distributions P and Q coincide only at one quantile and P first-order stochastically dominates Q otherwise, it may be that $Q \sim_P P$.

[46] Using a model-free definition, one can still ask how a quantile maximizer chooses between a lottery $P \in \mathcal{P}_o(\mathcal{X})$ and the expected return from this lottery:

$$P \begin{matrix} \succ \\ \sim \\ \prec \end{matrix} \int x_P \pi(x) dx. \tag{45}$$

Except in the extreme cases, $\tau = 0$ (risk aversion) and $\tau = 1$ (risk loving), quantile maximizers do not exhibit any global (i.e., for all lotteries P) risk attitude in the above sense. For example, consider a half-half bet between outcomes 1 and 3 versus a certain outcome of 2. All τ-maximizers with $\tau \leq \frac{1}{2}$ will strictly prefer the gamble's average face value of 2, while those with $\tau > \frac{1}{2}$ will choose to gamble. However, when probabilities are modified to $\frac{1}{3}$ on 1 and $\frac{2}{3}$ on 3, all quantile maximizers with $\tau \in (\frac{1}{3}; \frac{1}{2}]$ will switch to preferring the gamble.

Thus, the lower τ, the weakly more risk averse the decision maker is, with maxmin being the most and maxmax – the least risk averse. Proposition 3 suggests two ways in which the model studied in the present paper contributes to the description of choice behavior. First, maxmin agents have been commonly, though informally, referred to as *cautious* (e.g., in game theory). With the general quantile representation, the intuited notion of cautiousness can be formalized as a risk attitude.[47] Second, maxmin and maxmax have alternatively been interpreted as representing extreme optimism and pessimism. The quantile model also admits more moderate levels of optimism and pessimism.

It is worth emphasizing that, in the present model, risk attitudes can be analyzed even if outcomes are not measurable on an interval scale (as they are, e.g., for categorical variables). Indeed, the above definitions of risk and risk aversion do not impose any such assumptions on the set $\mathcal{X}$.

To provide a unified interpretation of risk attitudes for Quantile Maximization and the Expected Utility models, we will invoke a result familiar in econometrics. Doing so will also elucidate where the qualitative differences between the two models originate. Consider an individual who is choosing between two lotteries. It is useful to think about this individual as trying to assess (or, to estimate) the outcome of each lottery and then to choose the preferred lottery based on those estimates. We can then draw an analogy between the Expected Utility and Quantile Maximization in decision theory, and the Least-Squares (mean based) estimation and the Quantile Regression in econometrics. As is well known (see e.g. Koenker [2005, Ch. 1]), just as the least-squares estimator can be obtained by minimizing the loss function defined as a sum of squared errors, the median minimizes the sum of absolute errors. This evaluation can be extended to any quantile, by weighting mispredictions appropriately. In evaluating a given distribution P, the τ^{th} quantile solves the minimization problem for the loss function that assigns the weight of τ to overpredictions and $(1-\tau)$ to underpredictions. The agent's resulting prediction about (or, estimation of) the realization from the lottery

[47] Proposition 3 is also in Manski (1988), who did not, however, have in mind the connection to maxmin and maxmax. Manski (1988) also defined two other measures that can be readily used to compare agent risk attitudes. These measures parallel the concepts used for the Subjective Expected Utility and have very simple expressions in the Quantile Maximization model: The *τ-certainty equivalent of P*, $CE^{\tau}(P)$, is the amount of money for which a decision maker is indifferent between lottery P and $CE^{\tau}(P)$ with certainty. Clearly, the τ-certainty equivalent is the quantile outcome,

$$CE^{\tau}(P) = Q^{\tau}(P). \tag{46}$$

For any fixed amount of money $x \in \mathbb{R}$ and distribution $P \in \mathcal{P}_0(\mathcal{X})$, Manski (1988) defines a *risk premium* as the value $\mu_P(x, \tau) \in \mathbb{R}$ solving

$$u(x - \mu_P(x, \tau)) = Q^{\tau}(P). \tag{47}$$

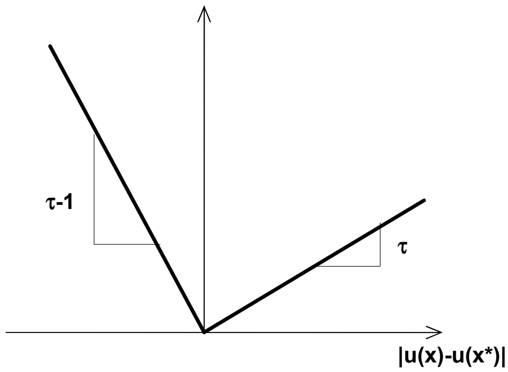
Letting $u(x) = x$, we have that the unique and finite risk premium is equal to $\mu_P(x, \tau) = x - Q^{\tau}(P)$.

is given by

$$(1-\tau)\int_{-\infty}^{u(x^*)}[u(x^*)-u(x)]dP(x)+\tau\int_{u(x^*)}^{\infty}[u(x)-u(x^*)]dP(x). \tag{48}$$

An asymmetric weighting function is depicted in Figure 2. The case of symmetric piecewise linear value function corresponds to the median.

Figure 2. A weighting function for the loss function (48)



When predicting the realization of a given lottery, the lower the agent's τ, the more that agent cares about the lower-tail outcomes relative to the upper-tail outcomes; hence, the more the individual is concerned about underpredictions relative to overpredictions, and the more cautious she is.

That the quantile solves the minimization of an absolute rather than a quadratic loss function makes it less affected by outliers. This characteristic of quantiles has been explored in econometrics using robust estimation. Mean-based estimators, such as OLS, are very sensitive to large errors and to asymmetric distributions that are often met in practice. A popular alternative estimator, which is more robust, is based on the median or quantile (Least Absolute Deviations, or LAD, method; Koenker and Bassett [1978b]).

7.4 Quantile Maximization in Practice

Using an insurance problem as an example, we point to several features of the model that might be attractive in applied work. Consider a quantile maximizer that chooses among alternative insurance plans, say, car, travel or medical insurance. The insurance company offers the following menu: The individual will choose the amount of money up to which she wants to be fully insured, $C \in \mathbb{R}_{++}$. Any losses up to C will be covered in full; for losses exceeding C, the individual will receive C. Thus, the family of menus offered can be written as $\{C, P(C)\}_{C\in\mathbb{R}_{++}}$, where C is the maximum coverage and $P(C)$ is the price for a contract with the maximal coverage of C. Let $P'(C) > 0$.

The Expected Utility requires that in order to choose optimally, the insuree must know all (insurable) levels of losses and their respective probabilities. Instead, we assume that the agent compares, say, the median levels of the distributions of wealth (e.g. "I am a median driver and I will incur a median level of losses"). Let π be the (subjective) distribution of losses. A τ-maximizer, $\tau \in [0,1]$, solves

$$\max_{\{C,C(P)\}} Q^{\tau}(R), \tag{49}$$

where R is the (cumulative distribution of) the random variable $R = W - L(d) + C(d) - P(C)$, W - the agent's wealth level, d - the damage variable distributed according to π, $L(d)$ - the level of loss incurred, $C(d)$ - the coverage provided by the contract. If the insurance company charges the fair price,[48] our model predicts that a τ-maximizer will insure up to the $(1-\tau)^{th}$ level of losses, $\tau \in [0,1]$. Thus, more risk averse individuals will choose higher levels of full coverage. The "full insurance" prediction parallels the one yielded by the Expected Utility framework. This is not, however, the present section's punchline.

Think about the insurance company taking the Quantile Maximization model to data. Unlike using the Expected Utility, with the quantile model: (1) There is no need to make any parametric assumptions about the client's utility function. (2) To compare agent risk attitudes, one does not have to first recover the concavity of utility function from data. As shown in Section (7.3), the model allows for the analyzing of attitudes toward risk even though the utilities need not be continuous, let alone concave. (3) To make policy recommendations based on the quantile model, recovering a unique parameter τ suffices to pin down the entire preference ordering $\succ_P$ over lotteries in $\mathcal{P}_0(\mathcal{X})$. This should be compared to recovering the cardinal Bernoulli utility function – in principle, an infinitely dimensional object. (4) Finally, the quantile model is robust to fat tails and works well with distributions that do not possess finite moments, a circumstance often encountered in non-life insurance.

7.5 Applications

"No theory exists to show that VaR is the appropriate measure upon which to build optimal decision rules." Schachter (1997, p. 19) As an application, our axiomatization provides the foundations for decision making based on Value-at-Risk. The present paper also suggests that VaR can serve not only as a measure of risk, but also as a (comparative) risk attitude. In this section, we revisit the central and distinct features of the quantile choice rule –

[48] The result holds for a wide range of prices and is robust to the assumption of "fair" price (equal to expected coverage).

robustness and ordinality – to illustrate how Quantile Maximization can complement and extend the class of applications of Expected Utility, or cardinal models more generally.

• Single-Event Decision Making: Consider an agent who is selecting a particular cancer treatment. By contrast to repeated decisions, such as adhering to a weight-loss program or maintaining smoking abstinence, the decision about medical treatment involves no opportunity for revisions: the choice is one-time.[49] That decision making differs qualitatively between one-time as opposed to repeated problems has been recognized in the medical sciences (and acknowledged by NIH) and artificial intelligence (see, e.g., review in Dubois *et al.* [2000], Dubois *et al.* [2002]). When the choice among uncertain alternatives (e.g., treatments) is made only once, using an expectation-based evaluation may not appear suitable for at least two reasons. First, since only a single outcome (e.g., survival rate) will be experienced, no compensation through averaging that comes from repetition will actually take place. Second, one-time decisions involve little opportunity for learning. Therefore, individuals may not have enough precise information about the potential extreme realizations, or may not know their preferences sufficiently well enough to be able to quantify them.[50] Many decisions of economic interest are made once or at most infrequently: choosing a retirement program, making a career choice, purchasing a durable good, selecting (life, car, travel or medical) insurance, to mention just a few.

• Agency Problems/Expert Recommendations: Consider a principal delegating a task, for which the realization is risky, to an agent or a group of agents. Since the task is hard to monitor, the principal aims to set a standard of performance that is independent of agent's(s') idiosyncratic preferences for money – either because of a lack of such knowledge in the case of a single agent or because of heterogeneity of risk preferences in the case of multiple agents. The quantile model similarly addresses such concerns in expert recommendations (e.g., advising customers on an optimal investment or insurance).

• Robust Policy Choice: The literature on robustifying economic and policy design has focused on relaxing the assumption that decision makers know or act as if they know the true probability distribution (e.g., Hansen and Sargent [2004] applying the model of Gilboa and Schmeidler [1989], Klibanoff, Marinacci and Mukerji [2005]). Another and less explored robustness test involves relaxing the assumption that decision makers have cardinal (as well as ordinal) rankings of outcomes; or that cardinal, parametric assumptions about utilities

[49] Again, the standard expectation-based framework presupposes that an agent knows all of the possible outcomes and their respective probabilities. This arguably assumes a lot on the part of decision makers, considering the costs and benefits of alternative prevention methods (e.g., genetic testing, clinical trials) or treatments (e.g, surgery, chemotherapy, radiotherapy). In the model proposed in this paper, an individual would only need to assess his accident chances relative to the population; e.g., "I am of median health, and hence I anticipate that median-level statistics (e.g., risk rates, survival rates) will apply to me."

[50] In decision-theoretic terms, the individuals may not be able to relate outcomes from different equivalence classes beyond ranking them (e.g., in the sense of a concatenation condition). Using the robustness interpretation, the agents might want to make a decision based on a criterion that is robust to their own utility (over money).

do affect decisions.[51]

• Scenario-Based Analysis: One prominent example of the existing applications of Quantile Maximization is choosing according to the "worst case scenario" (often occurs in policy design). This is typically justified by arguments that support precautionary policy (e.g., the "worst-case scenario" rule in the Environmental Protection Agency and the Department of Justice; the *Precautionary Principle* in the European Commission's food and agricultural biotechnology policy; and Walsh [2004] in the context of monetary policy). Critics have argued that such an extreme criterion places too much importance to outcomes that may be very unlikely. Moreover, while the "worst case scenario" is a natural criterion for when no information about probabilities can be obtained, that information is often available. The quantile representation studied in this paper generalizes the best- (and worst-)case scenario analysis to intermediate scenarios.[52]

• Categorical Variables: Many economic and social variables are categorical (e.g. jobs, the A-F grading scheme, qualities in online ratings). Applying a cardinal-utility model to studying choice for categorical variables imposes measurability on an interval scale (as well as a concatenation condition on preferences) – thereby denying the defining characteristic of such variables. Quantile Maximization offers an alternative that respects that characteristic and provides well defined risk measures.[53]

8 Concluding Remarks

For some applications, it may be desirable to extend the model proposed in this paper to more than one quantile. In particular, a choice rule may depend on the "focal" worst-, best-, and typical- case scenarios; or, only a range of quantiles that are higher or lower than some

[51] Another advantage brought by ordinality in the policy context is that the decisions are independent from how the policy maker values money.

[52] In the policy context, distributional consequences of policies are of interest much beyond just average statistics. Quantiles have been used to assess social policies and treatment effects, to compare unemployment duration and distributions of wages, etc. Their use in formal empirical studies has been spurred by Quantile Regression (Koenker and Bassett [1978a]), in which the classical least squares estimation of conditional mean is replaced by an estimation of conditional quantile functions.

[53] As another example, sometimes information is naturally or optimally given in the comparative rather than the absolute form; e.g., when information must be conveyed, but restricting an expert's incentive to exaggerate in absolute statements is desired. Rubinstein (1996) notes that comparative statements are relatively more common in natural language. He justifies the optimality of such statements by formalizing their three properties: The ability to indicate elements of the set, e.g., by means of order; the ability to accurately convey information; and the easiness by which a comparative statement can communicate the content to others.

Chakraborty and Harbaugh (2007) demonstrated that comparative cheap-talk statements can be credible when absolute statements are not (e.g., a professor ranking students for a prospective employer; a seller auctioning goods; an analyst's claims about the likely returns to a stock might not be credible, but the statement that one stock is better than another might be). In the context of multi-object auction, Chakraborty, Gupta, and Harbaugh (2002) show that a seller's incentive to lie may be diminished or eliminated when only comparative and not absolute statements are allowed.

threshold may be of interest. For instance, a policy may be targeted for a specific range of income distribution, school attainment, test performance, etc.

Another direction one could pursue would be to model Quantile Maximization with multiple priors. One compelling motivation comes from the recent interest in formalizing robustness concerns to address model uncertainty. The novelty here would be twofold: The multiple-prior model would allow for heterogenous ambiguity attitudes, nesting Gilboa and Schmeidler's (1989) maxmin as a special case. It would, thus, address the criticism of extremeness raised against the seminal multiple-prior maxmin (e.g., Sims [2001], Svensson [2007]).[54] Moreover, ordinality of the multiple-prior quantile model with respect to the second-order beliefs would ensure that the choice does not depend on a policy-maker utility function over the possible models, thereby, complementing recent contributions to modelling ambiguity (e.g., *Smooth Model* by Klibanoff, Mukerjii and Marinacci [2005]; *Variational Preferences* by Maccheroni, Marinacci and Rustichini [2006]).

[54] The extremeness of maxmin raises additional concerns in the context of multiple priors. For example, Svensson (2007) argues that "the worst possible model is on the boundary of a feasible set of models, and hence depends crucially on the assumed feasible set of models. If the worst possible model somehow ended up in the interior of the feasible set of models, one could perhaps argue that the outcome is less sensitive to the assumptions about the feasible set of models" (p. 7).

Note: Appendices 1, 2 and 3 contain proofs of Theorems 1, 2, and 3, respectively, with auxiliary lemmas. Appendix 4 provides proofs of other results.

Appendix 1: Proof of Theorem 1

In part 1A, we establish several auxiliary results that will be frequently used throughout the proof. Part 1B presents the proof for the non-extreme preferences (definition 2). Although the general line of the proof is essentially the same for the extreme preferences, the derived properties of the relation over acts are distinct and, therefore, the representation results require that alternative arguments be employed. In order to highlight the differences, we present the proof for the extreme preferences separately in part 1C. Part 1D contains the proofs of Lemmas 3 and 4.

1A. Auxiliary Results

The lemmas in this section characterize the binary relations $\succ^*$, $\succ_*$, and $\succ_{**}$ over events in $\mathcal{E}$ (defined in (22), (23) and (36), respectively) used in the paper. P5 ensures that the relations are nontrivial.

Lemma 5 *$E \succ_* F \Leftrightarrow (\mathcal{S} \sim_* E$ and $F \sim_* \varnothing)$; $E \succ^* F \Leftrightarrow (\mathcal{S} \sim^* E$ and $F \sim^* \varnothing)$.*

Proof. Since the arguments for $\succsim^*$ and $\succsim_*$ are symmetric, we only prove the assertion for $\succsim_*$. Let $E \succ_* F$, that is by definition (23) and P5,

$$f = \begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \notin F \\ y \text{ if } s \in F \end{bmatrix} = g, \text{ for all } x \succ y. \tag{50}$$

Then, it must be that event E is pivotal for act f and F^c – for g. By P3^Q,

$$f \sim y \sim \begin{bmatrix} x \text{ if } s \notin \mathcal{S} \\ y \text{ if } s \in \mathcal{S} \end{bmatrix} \text{ and } g \sim x \sim \begin{bmatrix} x \text{ if } s \in \mathcal{S} \\ y \text{ if } s \notin \mathcal{S} \end{bmatrix}. \tag{51}$$

That is, by the definition of $\succsim_*$, P1 and P4^Q, $\mathcal{S} \sim_* E$ and $F \sim_* \varnothing$.

For the converse, assume $\mathcal{S} \sim_* E,\ F \sim_* \varnothing$. Using the definition of $\succ_*$ in (23) and P5, for all $x \succ y$

$$\begin{array}{ccc} \begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} & & \begin{bmatrix} x \text{ if } s \notin F \\ y \text{ if } s \in F \end{bmatrix} \\ \wr & & \wr \\ \begin{bmatrix} x \text{ if } s \notin \mathcal{S} \\ y \text{ if } s \in \mathcal{S} \end{bmatrix} & & \begin{bmatrix} x \text{ if } s \in \mathcal{S} \\ y \text{ if } s \notin \mathcal{S} \end{bmatrix} \\ \wr & & \wr \\ y & \prec & x \end{array} \tag{52}$$

Then, by P1, P4^Q and (23), $E \succ_* F$, as desired. ■

Lemma 2 in Section 5 is a direct corollary of Lemma 5:

Lemma 2 $E \succ_* \varnothing \Leftrightarrow E \sim_* \mathcal{S}$; $E \prec_* \mathcal{S} \Leftrightarrow E \sim_* \varnothing$; $E \succ^* \varnothing \Leftrightarrow E \sim^* \mathcal{S}$; $E \prec^* \mathcal{S} \Leftrightarrow E \sim^* \varnothing$.

The next result establishes consistency between the relations $\succsim_*$ and $\succsim^*$.

Lemma 6 *$E \succ_* F \Rightarrow E \not\prec^* F$; $E \succ^* F \Rightarrow E \not\prec_* F$.*

Proof. There are three relevant cases to the proof of the assertion about the relation $\succsim_*$:

If E and F are disjoint: The result follows by P4^Q and Lemma 1.

If $F \subset E$ (symmetric for $E \subset F$): Define $H = E \backslash F$ and let H be non-null. The assertion will be implied if we show that it cannot be the case that $E \prec^* F$ or $E \prec_* F$. Suppose $E \prec_* F$. Using the definition of $\succ_*$ and Lemma 1,

$$\begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in H \\ y \text{ if } s \in F \end{bmatrix} \sim_{by\ P3^Q} \begin{bmatrix} x \text{ if } s \notin F \\ y \text{ if } s \in F \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} = \begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in H \\ y \text{ if } s \in F \end{bmatrix}. \tag{53}$$

Contradiction to P1. Hence, if $F \subset E$, then $F \precsim_* E$ and $F \precsim^* E$.

If $F \backslash E \neq \varnothing$, $E \backslash F \neq \varnothing$ and $F \cap E \neq \varnothing$: Define $I = F \cap E$. The result follows by P4^Q applied to events $F \backslash E$ and $E \backslash F$, and Lemma 1.

An analogous argument can be provided for the relation $\succsim^*$. ■

An important corollary of Lemma 6 is the following result:

Lemma 7 *If $E \sim_* F$ and there is a non-null event $G \subseteq (E \cup F)^c$, such that $E \cup G \succ_* F \cup G$, then there is no event $G' \subseteq (E \cup F)^c$, such that $E \cup G' \prec_* F \cup G'$.*

Lemma 1 establishes an important property of preferences that will be key to defining a likelihood relation over events. Following it Lemma 8 demonstrates that, assuming P1 and P5, the condition from Lemma 1 is not only necessary but also sufficient for P3^Q and could replace the latter throughout the axiomatization.

Lemma 1 *Assume Weak Order (P1), Pivotal Monotonicity (P3^Q) and Nondegeneracy (P5). For all events E and F, all pairs of outcomes $x^* \succ x$ and $y^* \succ y$, and all subacts g_{x^*+}, $g_{x^*-,x+}$, g_{x-}, h_{y^*+}, $h_{y^*-,y+}$, and h_{y-},*

$$\begin{bmatrix} g_{x^*+} \text{ if } s \in G_1 \\ x^* \text{ if } s \in E \\ g_{x^*-,x+} \text{ if } s \in G_2 \\ x \text{ if } s \in F \\ g_{x-} \text{ if } s \in G_3 \end{bmatrix} \succ \begin{bmatrix} g_{x^*+} \text{ if } s \in G_1 \\ x^* \text{ if } s \in F \\ g_{x^*-,x+} \text{ if } s \in G_2 \\ x \text{ if } s \in E \\ g_{x-} \text{ if } s \in G_3 \end{bmatrix} \Rightarrow \begin{bmatrix} h_{y^*+} \text{ if } s \in G_1 \\ y^* \text{ if } s \in E \\ h_{y^*-,y+} \text{ if } s \in G_2 \\ y \text{ if } s \in F \\ h_{y-} \text{ if } s \in G_3 \end{bmatrix} \succ \begin{bmatrix} h_{y^*+} \text{ if } s \in G_1 \\ y^* \text{ if } s \in F \\ h_{y^*-,y+} \text{ if } s \in G_2 \\ y \text{ if } s \in E \\ h_{y-} \text{ if } s \in G_3 \end{bmatrix}. \tag{54}$$

Proof. Let $x^* \succ x$ and $y^* \succ y$. Assume P1, P3^Q and P5, and that the left hand side of (54) holds; call the preferred act f and the less preferred act f'. Using P3^Q, the pivotal events for acts f and f' could be, respectively, E and E, or E and G_2, or G_2 and E, or G_2 and G_2.

Consider the first pair of pivotal events, E and E. We will prove that the implication in (54) must be satisfied. If the set of outcomes $\mathcal{X}$ consists of two nonindifferent events, the implication (54) is vacuous; we will separately consider the cases in which the outcome set consists of three and (at least) four pairwise nonindifferent outcomes.

Suppose $\mathcal{X} = \{x, y, z\}$ and let (w.l.o.g.) $x \succ y \succ z$. The following six cases need to be analyzed:

x^* and x in the left-hand-side inequality in (54) equal to:	x, y	x, z	y, z
y^* and y in the right-hand-side inequality in (54) equal to:	x, z *or* y, z	x, y *or* y, z	x, y *or* x, z

We will argue for the two pairs in the first column. Consider the implication (54), taking outcomes x, y for the left-hand-side implication and x, z for the right-hand-side one. Let $G_1 \cup G_2 \cup G_3$ be null (w.l.o.g.). Then,

$$\begin{array}{ccc} \begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \in F \end{bmatrix} & \succ & \begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \in E \end{bmatrix} \\ \wr_{\text{By P3}^Q} & & \curlyvee_{\text{By P3}^Q} \\ \begin{bmatrix} x \text{ if } s \in E \\ z \text{ if } s \in F \end{bmatrix} & & \begin{bmatrix} x \text{ if } s \in F \\ z \text{ if } s \in E \end{bmatrix} \end{array} \tag{55}$$

and that (54) holds follows by P1. Now consider the pairs x, y and y, z. Applying P3^Q to event E

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \in F \end{bmatrix} \succ_{\text{By P3}^Q} \begin{bmatrix} y \text{ if } s \in E \\ y \text{ if } s \in F \end{bmatrix} \sim_{\text{By P3}^Q} \begin{bmatrix} y \text{ if } s \in E \\ z \text{ if } s \in F \end{bmatrix} \sim_{\text{By P3}^Q} \begin{bmatrix} y \text{ if } s \in E \\ y \text{ if } s \in F \end{bmatrix} \tag{56}$$

and to event E

$$\begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \in E \end{bmatrix} \sim_{\text{By P3}^Q} \begin{bmatrix} y \text{ if } s \in F \\ y \text{ if } s \in E \end{bmatrix} \succ_{\text{By P3}^Q} \begin{bmatrix} y \text{ if } s \in F \\ z \text{ if } s \in E \end{bmatrix} \sim_{\text{By P3}^Q} \begin{bmatrix} z \text{ if } s \in F \\ z \text{ if } s \in E \end{bmatrix}, \tag{57}$$

the implication in (54) follows by P1 and $y \succ z$.

The argument is analogous for the remaining pairs of pivotal events and is, therefore, omitted.

Now let the outcome set consist of (at least) four outcomes x, y, z, w, such that $x \succ y \succ z \succ w$. Having proved that (54) holds for all three-outcome sets, it suffices to consider the cases in which the outcome sets mapped from events E and F, respectively, are disjoint. For example, to prove

$$f = \begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \in F \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \in E \end{bmatrix} = f' \Rightarrow \begin{bmatrix} z \text{ if } s \in E \\ w \text{ if } s \in F \end{bmatrix} \succ \begin{bmatrix} z \text{ if } s \in F \\ w \text{ if } s \in E \end{bmatrix}, \tag{58}$$

one can apply P3^Q to acts f and f' separately, mimicking the logic of (56) and (57). Analogous arguments apply to the remaining cases. ■

Lemma 8 *Assume Weak Order (P1) and Nondegeneracy (P5). The property established in Lemma 1 implies Pivotal Monotonicity (P3^Q).*

Proof. The proof restricts attention to the assertion that involves strict inequalities in P3^Q. Assume P1, P5 and that the condition shown in Lemma 1 holds. Consider an act $f \in \mathcal{F}$ and the coarsest partition with respect to which it is measurable. Let E be a non-null event in that partition and let x^* be the outcome to which it maps. Write f as $[g_{x^*+}, H_1; x^*, E; g_{x^*-}, H_2]$, where $H_1 \cup H_2 = E^c$. Suppose that

$$f = \begin{bmatrix} g_{x^*+} \text{ if } s \in H_1 \\ x^* \text{ if } s \in E \\ g_{x^*-} \text{ if } s \in H_2 \end{bmatrix} \succ \begin{bmatrix} g_{x+} \text{ if } s \in H_1 \\ x \text{ if } s \in E \\ g_{x-} \text{ if } s \in H_2 \end{bmatrix} = f'. \tag{59}$$

By Lemma 1, (59) holds for any subacts on events H_1 and H_2 that weakly preserve the ranking with respect to x^* and x. In particular, it must be that $x^* \succ x$, and thus the "if" part of P3^Q is established. That there exists a non-null event E, such that (59) holds is implied by the definition of nullness and P5.

Take again the act f and its coarsest measurable partition. Let x^* be the outcome on a non-null event E in the partition. For an outcome $x \prec x^*$ (it exists by P5), construct an act f' from (59), such that $g_{x^*+} = g_{x+}$ and $g_{x^*-} = g_{x-}$. It follows from the definition of nullness and P5 that the coarsest measurable partition of f must contain at least one such event E for which, in addition, the strict inequality from (59) holds. The "only if" implication of P3^Q then obtains through Lemma 1. ■

1B. Proof of Theorem 1 for Non-extreme Preferences

Proof. Assume that preferences are not extreme (definition 2).

The proof consists of a series of steps. Step 1 demonstrates that $\succ_*$ and $\succ^*$ are weak orders. Step 2 characterizes the set of equivalence classes of $\mathcal{E}$ under $\sim_*$ and $\sim^*$. They are used in Step 3 to derive a subset of $\mathcal{E}$, $\mathcal{E}_{**}$, on which a new and complete likelihood relation is defined, $\succ_{**}$. Step 4 verifies that axioms A1, A3, A4 and A5' hold on $\mathcal{E}_{**}$, which is then employed in Step 5 to derive a unique, convex-ranged and finitely additive probability-measure representation of $\succ_{**}$ on $\mathcal{E}_{**}$, $\tilde{\pi}$. Next, Step 6 constructs a likelihood relation which is complete on the entire set of events, $\mathcal{E}$, and shows that measure $\tilde{\pi}$ on $\mathcal{E}_{**}$ uniquely extends to $\mathcal{E}$; we call the extended measure π. Finally, Step 7 establishes that $\succ$ is probabilistically sophisticated w.r.t. π. We will repeatedly invoke Lemma 1 without mentioning; it assures that the likelihood relations used in the axiomatization can be defined as revealed from preferences over acts.

Step 1 ($\succ_*$ **AND** $\succ^*$ **ARE WEAK ORDERS**):

1. We prove that $\succ_*$ is a weak order. Asymmetry is implied by the definition of $\succ_*$ (23), P1 and P5. To show negative transitivity, suppose $E \not\succ_* F$ and $F \not\succ_* G$. With $x \succ y$ (P5), (23) and Lemma 1 give

$$\begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} \not\prec \begin{bmatrix} x \text{ if } s \notin F \\ y \text{ if } s \in F \end{bmatrix} \text{ and } \begin{bmatrix} x \text{ if } s \notin F \\ y \text{ if } s \in F \end{bmatrix} \not\prec \begin{bmatrix} x \text{ if } s \notin G \\ y \text{ if } s \in G \end{bmatrix}. \tag{60}$$

Then, P1 yields

$$\begin{bmatrix} x \text{ if } s \notin E \\ y \text{ if } s \in E \end{bmatrix} \not\prec \begin{bmatrix} x \text{ if } s \notin G \\ y \text{ if } s \in G \end{bmatrix}, \tag{61}$$

hence $E \not\succ_* G$. A symmetric argument proves that $\succ^*$ is a weak order.

2. $\mathcal{S} \succ_* \varnothing$ and $\mathcal{S} \succ^* \varnothing$. Suppose otherwise, then the definitions of $\succ$ and $\succ_*$ lead to a contradiction.

Step 2 (**CHARACTERIZATION OF EQUIVALENCE CLASSES OF** $\mathcal{E}$ **UNDER** $\sim_*$ **AND** $\sim^*$):

1. Since $\succ_*$ on $\mathcal{E}$ is a weak order, $\sim_*$ is an equivalence relation. By Lemma 5, there are only two equivalence classes on $\mathcal{E}$ under $\sim_*$: $\mathcal{E}|_{\sim_*\varnothing} = \{F \in \mathcal{E}|F \sim_* \varnothing\}$ and $\mathcal{E}|_{\sim_*\mathcal{S}} = \{F \in \mathcal{E}|F \sim_* \mathcal{S}\}$. Similarly, there are only two equivalence classes on $\mathcal{E}$ under $\sim^*$: $\mathcal{E}|_{\sim^*\varnothing} = \{F \in \mathcal{E}|F \sim^* \varnothing\}$ and $\mathcal{E}|_{\sim^*\mathcal{S}} = \{F \in \mathcal{E}|F \sim^* \mathcal{S}\}$.

2. That the sets $\mathcal{E}|_{\sim_*\varnothing}$, $\mathcal{E}|_{\sim_*\mathcal{S}}$ and $\mathcal{E}|_{\sim^*\varnothing}$, $\mathcal{E}|_{\sim^*\mathcal{S}}$ all contain non-null events, follows from the assumption of the non-extreme preferences.

Step 3 (**CONSTRUCTION OF** $\mathcal{E}_{**}$):

1. Define $\mathcal{E}_{**} = \{\bar{E} \in \mathcal{E}|\bar{E} \prec_* \bar{E}^c\}$. Fix $E \in \mathcal{E}_{**}$. Then, by P3^Q, for any $F \in \mathcal{E}_{**}$ disjoint (w.l.o.g.) with E (and hence such that $F \subset E^c$), there exists $G \subseteq (E \cup F)^c$, such that $F \cup G \succ_* \varnothing$. Since the above is true for disjoint E and F, hence for any $E, F \in \mathcal{E}_{**}$ there exists $G \subseteq (E \cup F)^c$, such that $E \cup G \succ_* F \cup G$ or $E \cup G \prec_* F \cup G$ or $E \cup G \sim_* F \cup G \sim_* \mathcal{S}$.

2. We prove that either $\mathcal{E}_{**} = \mathcal{E}|_{\sim_*\varnothing}$ or $\mathcal{E}_{**} = \mathcal{E}|_{\sim^*\varnothing}$. Assume first that there exists an event $E \in \mathcal{E}|_{\sim_*\varnothing}$, such that $E \sim_* E^c \sim_* \varnothing$. Consider an event $F \in \mathcal{E}|_{\sim^*\varnothing}$. We will show that for all such events F, $F \sim_* \varnothing$. By the definitions of $\succ_*$ and $\succ^*$ and Lemma 2, $F^c \succ_* \varnothing$. Using the definitions of $\succ_*$ and $\succ^*$ again, $F \prec^* E$ and $F \prec^* E^c$. Then, by Lemma 6, $F \precsim_* E$ and $F \precsim_* E^c$. Since $\succ_*$ is a weak order (Step 1), it follows that $F^c \succ_* F \sim_* \varnothing$. The event $F \in \mathcal{E}|_{\sim^*\varnothing}$ was picked arbitrarily, and hence, $\mathcal{E}_{**} = \mathcal{E}|_{\sim^*\varnothing}$.

Furthermore, if there is no $E \in \mathcal{E}|_{\sim_*\varnothing}$ for which $E \sim_* E^c \sim_* \varnothing$, then $\mathcal{E}_{**} = \mathcal{E}|_{\sim_*\varnothing}$.

3. On the collection $\mathcal{E}_{**}$, define a binary relation over events, $\succ_{**}$, as in definition 3.

4. By Step 2.2, and axioms P6^{Q_*} and P6$^{Q^*}$, $\mathcal{E}_{**}$ and the relation $\succ_{**}$ are nondegenerate.

Step 4 (**AXIOMS A1, A3, A4, A5' HOLD ON** $\mathcal{E}_{**}$): Let $E, F, H \in \mathcal{E}_{**}$.

(**A1**) Consider $E \sim_* \varnothing$. By an argument analogous to the one in Lemma 6, there cannot exist an event $G \subseteq E^c$, such that

$$f = \begin{bmatrix} x \text{ if } s \notin E \cup G \\ y \text{ if } s \in E \cup G \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \notin G \\ y \text{ if } s \in G \end{bmatrix} = g. \tag{62}$$

Hence $E \not\succ_{**} \varnothing$. If E is null, then for all $G' \subseteq E^c$

$$\begin{bmatrix} x \text{ if } s \notin E \cup G' \\ y \text{ if } s \in E \cup G' \end{bmatrix} \sim \begin{bmatrix} x \text{ if } s \notin G' \\ y \text{ if } s \in G' \end{bmatrix}. \tag{63}$$

Again, $E \not\succ_{**} \varnothing$.

(**A3**) By Lemma 7, $\succ_{**}$ is asymmetric. Condition (i) in negative transitivity follows from the transitivity of $\sim_*$ ($E, F, H \in \mathcal{E}_{**}$). Suppose that:

(1) There does not exist $G \subseteq E^c \cap F^c$ non-null, such that $E \cup G \succ_* F \cup G$; and

(2) There does not exist $G' \subseteq F^c \cap H^c$ non-null, such that $F \cup G' \succ_* H \cup G'$.

We need to show that for no $G" \subseteq E^c \cap H^c$ non-null, $E \cup G" \succ_* H \cup G"$.

Observe that (1) can be extended to all subsets of F^c: for all $G \subseteq F^c$,

$$\begin{bmatrix} x \text{ if } s \notin E \cup G \\ y \text{ if } s \in E \cup G \end{bmatrix} \succsim \begin{bmatrix} x \text{ if } s \notin F \cup G \\ y \text{ if } s \in F \cup G \end{bmatrix}. \tag{64}$$

Similarly, for all $G' \subseteq H^c$,

$$\begin{bmatrix} x \text{ if } s \notin F \cup G' \\ y \text{ if } s \in F \cup G' \end{bmatrix} \succsim \begin{bmatrix} x \text{ if } s \notin H \cup G' \\ y \text{ if } s \in H \cup G' \end{bmatrix}. \tag{65}$$

Thus, for all $G'' \subseteq F^c \cap H^c$ (where $F^c \cap H^c \neq \varnothing$ by Step 3.1),

$$\begin{bmatrix} x \text{ if } s \notin E \cup G \\ y \text{ if } s \in E \cup G \end{bmatrix} \succsim \begin{bmatrix} x \text{ if } s \notin H \cup G \\ y \text{ if } s \in H \cup G \end{bmatrix}. \tag{66}$$

Applying the above argument again, (66) holds for all $G'' \subseteq H^c$, and hence for all $G'' \subseteq E^c \cap H^c$. By P1, P5, Lemma 1 and Step 3.1, this proves the assertion.

(**A4**) Assume $E \cap H = F \cap H = \varnothing$, $E \cup H, F \cup H \in \mathcal{E}_{**}$, and $x \succ y$.

($\Leftarrow$) Assume that there is event $G \subseteq E^c \cap F^c \cap H^c$, such that $E \cup H \cup G \succ_* F \cup H \cup G$. Taking $G' = G \cup H$ immediately gives event G' non-null, such that $E \cup G' \succ_* F \cup G'$.

($\Rightarrow$) Assume now that there is event $G'' \subseteq E^c \cap F^c$, such that $E \cup G'' \succ_* F \cup G''$. By Lemma 7, $E \cup H \cup G''' \succsim_* F \cup H \cup G'''$ for all $G''' \subseteq E^c \cap F^c \cap H^c$. It suffices to show that it is not the case that for all $G''' \subseteq E^c \cap F^c \cap H^c$, $E \cup H \cup G''' \sim_* F \cup H \cup G'''$ (note that, by Step 3.1, $E^c \cap F^c \cap H^c$ is nonempty and there exists $\tilde{G} \subseteq E^c \cap F^c \cap H^c$ for which $E \cup H \cup \tilde{G} \sim_* \mathcal{S}$, or $F \cup H \cup \tilde{G} \sim_* \mathcal{S}$, or both). Suppose than that $E \cup H \cup G'' \not\succ_* F \cup H \cup G''$, for otherwise the assertion is delivered. Since, by assumption of $E \cup G'' \succ_* F \cup G''$ and Lemma 5, $E \cup G'' \sim_* \mathcal{S}$, we have that $E \cup H \cup G'' \sim_* F \cup H \cup G'' \sim_* \mathcal{S}$. Observe that, by P6^{Q_*}, event H can be partitioned into two non-null events H_1 and H_2 such that

$$\begin{bmatrix} x \text{ if } s \notin F \cup G'' \cup H_1 \\ y \text{ if } s \in F \cup G'' \cup H_1 \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \notin F \cup G'' \cup H \\ y \text{ if } s \in F \cup G'' \cup H \end{bmatrix} \sim \begin{bmatrix} x \text{ if } s \notin E \cup G'' \cup H \\ y \text{ if } s \in E \cup G'' \cup H \end{bmatrix} \sim \begin{bmatrix} x \text{ if } s \notin E \cup G'' \\ y \text{ if } s \in E \cup G'' \end{bmatrix} \tag{67}$$

Using P6^Q and non-nullness of H_1 and H_2, there is a subset of G'', $\tilde{G}$, such that

$$\begin{bmatrix} x \text{ if } s \notin F \cup \tilde{G} \cup H \\ y \text{ if } s \in F \cup \tilde{G} \cup H \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \notin F \cup G'' \cup H \\ y \text{ if } s \in F \cup G'' \cup H \end{bmatrix} \sim \begin{bmatrix} x \text{ if } s \notin E \cup \tilde{G} \cup H \\ y \text{ if } s \in E \cup \tilde{G} \cup H \end{bmatrix}. \tag{68}$$

(**A5'**) It follows by P6^Q and the definition of $\succ_*$.

Step 5 (**DERIVATION OF** π **ON** $\mathcal{E}_{**}$):

1. Axioms A1, A3, A4, and A5' hold for all subsets of $\mathcal{E}_{**}$. Define a sub-collection of events, $\overline{\mathcal{E}}$, by $\overline{\mathcal{E}} = \{E \in \mathcal{E} | \nexists F \text{ non-null: } E \backslash F \in \mathcal{E} \backslash \mathcal{E}_{**}\}$. By construction, for any event $E \in \overline{\mathcal{E}}$, $E \succ_{*} \varnothing$ if $\mathcal{E}_{**} = \mathcal{E}|_{\sim_{*}\varnothing}$ and $E \succ_{*} \varnothing$ if $\mathcal{E}|_{\sim^{*}\varnothing}$. Call the latter property A2', a counterpart of Savage's A2 (See Section 5). We will argue that using A1, A2', A3, A4 and A5', Fishburn's (1970, Ch.14) proof can be applied to construct a unique, finitely additive and convex-ranged probability measure $\tilde{\pi}$ that represents the likelihood relation $\succ_{**}$ on the collection $\mathcal{E}_{**} \cup \overline{\mathcal{E}}$. For any event $E \in \overline{\mathcal{E}}$, set $\tilde{\pi}(E) = \tau$ if $\mathcal{E}_{**} = \mathcal{E}|_{\sim_{*}\varnothing}$ and set $\tilde{\pi}(E) = 1 - \tau$ if $\mathcal{E}_{**} = \mathcal{E}|_{\sim^{*}\varnothing}$. We need to show that Fishburn's argument delivers the desired result when the collection $\mathcal{E}$ is replaced by $\mathcal{E}_{**} \cup \overline{\mathcal{E}}$ (with $\overline{\mathcal{E}}$ serving the role of $\mathcal{S}$).

2. Take an event $E \in \overline{\mathcal{E}}$ and let $\mathcal{E}_E$ be the collection of all of its subsets. Applied to the relation $\succ_{**}$ and the collection $\mathcal{E}_E$, Fishburn's result yields a unique, finitely additive and convex-ranged probability measure $\tilde{\pi}$ that represents $\succ_{**}$ on $\mathcal{E}_E$. By Step 5.1, the measure $\tilde{\pi}$ on $\mathcal{E}_E$ is normalized to τ if the collection $\mathcal{E}_{**}$, which induced $\overline{\mathcal{E}}$, is equal to $\mathcal{E}|_{\sim_{*}\varnothing}$ and to $1 - \tau$ if that collection is equal to $\mathcal{E}|_{\sim^{*}\varnothing}$.

3. Using that $\succ_{**}$ is well defined on all of $\mathcal{E}_{**}$, we extend the measure $\tilde{\pi}$ to the remaining events in $\mathcal{E}_{**} \cup \overline{\mathcal{E}}$ as follows: For an event F in $\left(\mathcal{E}_{**} \cup \overline{\mathcal{E}}\right) \backslash \mathcal{E}_E$, let

$$\tilde{\pi}(F) = \begin{cases} \pi(H) \text{ if } F \in \mathcal{E}_{**}, \text{ where } H \in \mathcal{E}_E \text{ and } H \sim_{**} F; \\ \pi(E) \text{ if } F \in \overline{\mathcal{E}}. \end{cases} \tag{69}$$

For the extension (69), to be well defined, we need to show that for any event $F \in \mathcal{E}_{**}$ there exists an event H as specified in (69). This follows by convex-rangedness of $\tilde{\pi}$ on $\mathcal{E}_E$ and Debreu (1954) ($\succ_{**}$ is a weak order and, by Step 4 (A5'), the collection $\left(\mathcal{E}_{**} \cup \overline{\mathcal{E}}\right) \backslash \mathcal{E}_E$ contains a countable $\succ_{**}$-dense subset). Thus, one can map all equivalence classes in $\left(\mathcal{E}_{**} \cup \overline{\mathcal{E}}\right) \backslash \mathcal{E}_E$ to those in $\mathcal{E}_E$.

4. We finally demonstrate that the extension of the measure $\tilde{\pi}$ to all events in $\mathcal{E}_{**} \cup \overline{\mathcal{E}}$, normalized as in Step 5.1, is unique, and preserves finite additivity and convex-rangedness. Consider an event $F \in \left(\mathcal{E}_{**} \cup \overline{\mathcal{E}}\right) \backslash \mathcal{E}_E$ and its two partitions:[55] $\{F_1, ..., F_M\}$ and $\{H_1, ..., H_N\}$. Let $\{I_1, ..., I_L\}$ be the coarsest common refinement of those partitions. Uniqueness of summations $\sum_{n=1,...,M} \tilde{\pi}(F_m)$ and $\sum_{m=1,...,N} \tilde{\pi}(H_n)$ follows immediately from their each being equal to $\sum_{l=1,...,L} \tilde{\pi}(I_l) = \sum_{l=1,...,L} \tilde{\pi}(E_l)$ where, for every $l \in \{1, ..., L\}$, $I_l \sim_{**} E_l \in \mathcal{E}_E$, and since each event is finitely decomposed, $\tilde{\pi}$ is finitely additive on $\mathcal{E}_{**} \cup \overline{\mathcal{E}}$. To see that it is convex-ranged, for any $\rho \in [0, 1]$ and $F \in \mathcal{E}_{**} \cup \overline{\mathcal{E}}$, observe that $\rho \cdot \tilde{\pi}(F) = \rho \cdot \tilde{\pi}(H) = \tilde{\pi}(H')$, where we used that there is an event H, such that $F \sim_{**} H \in \mathcal{E}_E$ and that, by convex-rangedness of $\tilde{\pi}$ on $\mathcal{E}_E$, there is a sub-event $H' \subseteq H$, such that $\tilde{\pi}(H') = \rho \cdot \tilde{\pi}(H)$. It remains to show that $\tilde{\pi}(H') = \tilde{\pi}(F')$ for some $F' \subseteq F \in \left(\mathcal{E}_{**} \cup \overline{\mathcal{E}}\right) \backslash \mathcal{E}_E$. To this end, apply P6^Q to construct a partition of $\mathcal{S}$, $\{G_1, ..., G_K\}$, every event $k \in \{1, ..., K\}$ of which is such that $\tilde{\pi}(G_k) \leq \tilde{\pi}(E)$.

[55] Such finite partitions exist by P6^{Q_*} applied to $\mathcal{S}$ if $E \in \mathcal{E}_{**}$, and applied directly to E if $E \in \overline{\mathcal{E}}$.

On each of these events, Fisburn's theorem generates a measure with desired properties. Normalizing measures of events G_k according to (69) completes the argument.

Step 6 (**EXTENDING $\tilde{\pi}$ TO $\mathcal{E}$**): From the relation $\succ_{**}$ on $\mathcal{E}_{**}$, we derive a complete binary relation $\succ_{***}$ on $\mathcal{E}$.

1. We first show that all events in $\mathcal{E}\backslash\mathcal{E}_{**}$ can be partitioned into finitely many events, each of which is in $\mathcal{E}_{**}$. We consider the cases $\mathcal{E}_{**} = \mathcal{E}|_{\sim_*\varnothing}$ and $\mathcal{E}_{**} = \mathcal{E}|_{\sim^*\varnothing}$ separately. First, assume that $\mathcal{E}_{**} = \mathcal{E}|_{\sim_*\varnothing}$ and consider an event $E \in \mathcal{E}|_{\sim_*\mathcal{S}}$. By P6^{Q_*}, there exists an N-partition of $\mathcal{S}$, $\{G_1, ..., G_N\}$, such that $E \succ_* G_n,$ for all $n = 1, ..., N$. By Lemma 5, for all $n = 1, ..., N$, $G_n \sim_* \varnothing$ and, hence, $G_n^N \in \mathcal{E}_{**}$. Write $E = \left(\bigcup\limits_{\tilde{n}=1}^{m} G_{\tilde{n}}^N\right) \cup \left(E\backslash \bigcup\limits_{\tilde{n}=1}^{m} G_{\tilde{n}}^N\right)$, where $m \in \{1, ..., N\}$ is such that $\bigcup\limits_{\tilde{n}=1}^{m} G_{\tilde{n}}^N \subseteq E \subset \bigcup\limits_{\tilde{n}=1}^{m+1} G_{\tilde{n}}^N$. By construction, for all $\tilde{n} = 1, ..., m$, $G_{\tilde{n}}^N \in \mathcal{E}_{**}$, while $E\backslash \bigcup\limits_{\tilde{n}} G_{\tilde{n}}^N \subset G_{m+1}^N$ and hence also $E\backslash \bigcup\limits_{\tilde{n}} G_{\tilde{n}}^N \in \mathcal{E}_{**}$. Thus, the events in the collection $\mathcal{E}\backslash\mathcal{E}_{**}$ can be decomposed into events from the collection $\mathcal{E}_{**}$.

Next, assume that $\mathcal{E}_{**} = \mathcal{E}|_{\sim^*\varnothing}$. By P6$^{Q^*}$, the state space $\mathcal{S}$ can be partitioned into finitely many events $\{H_1, ..., H_M\}$, such that $\mathcal{S} \succ^* H_m$ (and hence, by Lemma 5, $H_m \sim^* \varnothing$) for every $m = 1, ..., M$. Using the argument analogous to the one for $\mathcal{E}_{**} = \mathcal{E}|_{\sim_*\varnothing}$, each event in $\mathcal{E}|_{\sim^*\mathcal{S}}$ can be partitioned into finitely many events in $\mathcal{E}|_{\sim^*\varnothing}$.

2. For events E and F in $\mathcal{E}$, define a binary relation $\succ_{***}$ as follows: $E \succ_{***} F$ if there exists N-partitions of E and F, such that for all $n = 1, ..., N$, $E_n \succ_{**} F_n$. The condition necessary (and, as demonstrated below, sufficient) for the existence of such partitions can be found using the convex-rangedness of $\tilde{\pi}$ on $\mathcal{E}_{**} \cup \overline{\mathcal{E}}$: Consider $E, F \notin \mathcal{E}_{**}$ and let $\{E_1, ..., E_N\}$ and $\{F_1, ..., F_N\}$ be partitions of E and F, respectively, into elements in $\mathcal{E}_{**}$. By convex-rangedness of $\tilde{\pi}$, those partitions can be made equi-numbered and such that if $\sum\limits_{n=1,...,N} \tilde{\pi}(E_n) > \sum\limits_{n=1,...,N} \tilde{\pi}(F_n)$, then for each $n = 1, ..., N$, $\tilde{\pi}(E_n) > \tilde{\pi}(F_n)$. This also shows that if there exists an N-partition of E and F, such that for all $n = 1, ..., N : E_n \succ_{**} F_n$, then it cannot hold for any N'-partition that for all $n' = 1, ..., N' : E_{n'} \prec_{**} F_{n'}$.

3. We show that the probability measure $\tilde{\pi}$ can be uniquely extended to a finitely additive and convex-ranged probability measure representing the relation $\succ_{***}$ on the collection $\mathcal{E}$. Define an extension of $\tilde{\pi}$ to $\mathcal{E}\backslash\mathcal{E}_{**}$, π: For each $E \in \mathcal{E}\backslash\mathcal{E}_{**}$ and its finite partition $\{E_1, ..., E_N\}$, $E_n \in \mathcal{E}_{**}$ for $n \in \{1, ..., N\}$, let $\pi(E) = \sum\limits_{n=1,...,N} \tilde{\pi}(E_n)$.

Consider two partitions of a given event $E \in \mathcal{E}\backslash\mathcal{E}_{**}$: $\{E_1, ..., E_N\}$, $\{F_1, ..., F_M\}$, $E_n \in \mathcal{E}_{**}$ for $n \in \{1, ..., N\}$ and $F_m \in \mathcal{E}_{**}$ for $m \in \{1, ..., M\}$. Let $\{H_1, ..., H_L\}$ be the coarsest common refinement of those partitions. Uniqueness of summations $\sum\limits_{n=1,...,N} \tilde{\pi}(E_n)$ and $\sum\limits_{m=1,...,M} \tilde{\pi}(F_m)$ follows immediately from each being equal to $\sum\limits_{l=1,...,L} \tilde{\pi}(H_l)$, and since each event is finitely decomposed, π is finitely additive on $\mathcal{E}$. To see that it is convex-ranged, for any $E \in \mathcal{E}$ and

$\rho\in[0,1]$, take $\rho\cdot\pi(E)=\rho\cdot\sum_{n=1,...,N}\tilde{\pi}(E_n)=\sum_{n=1,...,N}\rho\cdot\tilde{\pi}(E_n)=\sum_{n=1,...,N}\tilde{\pi}(G_n)=\pi\left(\bigcup_{n=1}^{N}G_n\right)$, where we used that for each $n=1,...,N$, there is a $G_n\subseteq E_n$, such that $\tilde{\pi}(G_n)=\rho\cdot\tilde{\pi}(E_n)$.

Step 7 ($\succ$ **IS PROBABILISTICALLY SOPHISTICATED W.R.T.** π):

1. Establishing condition (24) is an application of the argument in Machina and Schmeidler (1992, Theorem 1, Step 5). It suffices to show that the construction employed there can be used. This follows from Lemma 3A.

2. Given that π is convex-ranged, for any $P\in\mathcal{P}_0(\mathcal{X})$, there exists an act $f\in\mathcal{F}$, such that $\pi\circ f^{-1}=P$. Therefore, using in addition that $\succ$ is a weak order, the stronger version of probabilistic sophistication from Section 4.1 is also satisfied. ■

1C. Proof of Theorem 1 for Extreme Preferences

Assume that preferences are extreme (definition 2).

Proof. Note: To aid in contrasting the arguments with those in the proof for the non-extreme preferences, each step is assigned the same number as its counterpart step in Appendix 1B. Some steps are left out as no longer relevant.

Denote the binary relation defined on $\mathcal{E}$ in (36) by $\succ_*^*$; that is, $E\succ_*^* F$ if $E\succ^* F$ or $E\succ_* F$.

Step 1: By an argument as in Step 1, Appendix 1B, and Lemma 6, the relation $\succsim_*^*$ is a weak order on $\mathcal{E}$.

Step 2:

1. Given that $\succsim_*^*$ on $\mathcal{E}$ is a weak order, Lemma 5 defines three equivalence classes of $\mathcal{E}$ under $\sim_*^*$: $\mathcal{E}|_{\sim_*^*\varnothing}=\{F\in\mathcal{E}|F\sim_*^*\varnothing\}$, $\mathcal{E}|_{\succ_*^*\varnothing}^{\prec_*^*\mathcal{S}}=\{F\in\mathcal{E}|\mathcal{S}\succ_*^* F\succ_*^*\varnothing\}$ and $\mathcal{E}|_{\sim_*^*\mathcal{S}}=\{F\in\mathcal{E}|F\sim_*^*\mathcal{S}\}$.

2. We show that under (**H**), the equivalence classes $\mathcal{E}|_{\sim_*^*\varnothing}$ and $\mathcal{E}|_{\sim_*^*\mathcal{S}}$ contain events that differ from $\varnothing$ and $\mathcal{S}$, respectively, only on a null sub-event, that is

$$\begin{aligned}\mathcal{E}|_{\sim_*^*\varnothing} &= \{E\in\mathcal{E}|E\text{ is null}\} \qquad (70)\\ \mathcal{E}|_{\sim_*^*\mathcal{S}} &= \{F\in\mathcal{E}|F=\mathcal{S}\backslash H,\ H\text{ null}\}.\end{aligned}$$

Consider a non-null event E, such that $\mathcal{S}\backslash E$ is non-null (possible by P6^{Q_*}). Then, given the assumption (**H**),

$$\begin{bmatrix} x\text{ if }s\notin E\\ y\text{ if }s\in E\end{bmatrix}\succ\begin{bmatrix} x\text{ if }s\notin \mathcal{S}\\ y\text{ if }s\in \mathcal{S}\end{bmatrix}\text{ and }\begin{bmatrix} x\text{ if }s\in E\\ y\text{ if }s\notin E\end{bmatrix}\succ\begin{bmatrix} x\text{ if }s\notin \mathcal{S}\\ y\text{ if }s\in \mathcal{S}\end{bmatrix},\ x\succ y. \qquad (71)$$

By Lemma 1 and the definitions of $\succ_*$ and $\succ^*$, it follows accordingly that $E \prec_* \mathcal{S}$ and $E \succ^* \varnothing$. Since E is an arbitrary non-null event with a non-null complement, using the definition of $\succsim_*^*$, we conclude that the equivalence classes $\mathcal{E}|_{\sim_*^*\varnothing}$ and $\mathcal{E}|_{\sim_*^*\mathcal{S}}$ are as characterized in (70). The proof for the case (**L**) is analogous and is, therefore, omitted.

3. By Lemma 3B, all events that differ from each $\varnothing$ and $\mathcal{S}$ on a non-null sub-event are ranked as equally likely by $\sim_*^*$ and are, thus, contained in a single equivalence class, $\mathcal{E}|_{\succ_*^*\varnothing}^{\prec_*^*\mathcal{S}}$:

$$\mathcal{E}|_{\succ_*^*\varnothing}^{\prec_*^*\mathcal{S}} = \{G \in \mathcal{E}|G \text{ is non-null and } G^c \text{ is non-null}\}. \tag{72}$$

Step 3: Suppose that we wish to define a counterpart of the binary relation $\succ_{**}$ (definition 3), now based on the relation $\succsim_*^*$ rather than $\succ_*$, on the subset of the collection $\mathcal{E}$ containing all events E, such that $E \prec_*^* E^c$. When either of the assumptions (**H**) or (**L**) holds, the subset of events in $\mathcal{E}|_{\succ_*^*\varnothing}^{\prec_*^*\mathcal{S}}$ that could possibly be strictly ranked through such a relation contains only nested events that differ on non-null sub-events. For example, under (**H**), consider two events $E_1, E_2 \in \mathcal{E}|_{\succ_*^*\varnothing}^{\prec_*^*\mathcal{S}}$, such that $E_1 \subset E_2$ and $E_2 \backslash E_1$ is non-null. Then, by Step 2, $E_1 \sim_*^* E_2$ and for $G = \mathcal{S}\backslash E_2$:

$$\begin{bmatrix} x \text{ if } s \notin E_1 \cup G \\ y \text{ if } s \in E_1 \cup G \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \notin \mathcal{S} \\ y \text{ if } s \in \mathcal{S} \end{bmatrix} = \begin{bmatrix} x \text{ if } s \notin E_2 \cup G \\ y \text{ if } s \in E_2 \cup G \end{bmatrix} \tag{73}$$

and the conclusion obtains by Lemma 1. Observe that since the strict ranking of E_1 and E_2 can only be achieved by adding the complement of the nesting event E_2 (up to null differences), there can be no event G' for which the ranking would be reversed.

However, the strict relation cannot be extended to non-nested events F_1 and F_2 for which $F_1\backslash F_2$ and $F_2\backslash F_1$ are non-null. What fails is that, under (**H**) or (**L**), any event can be strictly ranked by the relation $\succ_*^*$ only with the events in $\mathcal{E}|_{\sim_*^*\mathcal{S}}$, and hence, by Step 2, there cannot exist a non-null event in the common complement of any two non-nested events that would enable their strict comparison. Therefore, the strict ranking cannot be extended beyond nested events. In other words, there are no non-null events E, such that $E \prec_*^* E^c$.

Step 4: Demonstrating that A1, A2, A3 and A5 hold for the relation $\succ_*^*$ on $\mathcal{E}$ can be done through arguments analogous to those for the non-extreme relation $\succ$. To establish A4, assume $E \cap H = F \cap H = \varnothing$ and $x \succ y$. Since by Step 3 the relation $\succ_*^*$ can satisfy A4 only for nested events (with non-null set differences), let $F \subset E$ with E/F non-null.

($\Leftarrow$) For $G = E^c \cap F^c \cap H^c = E^c \cap H^c$ we have $E \cup H \cup G \succ_*^* F \cup H \cup G$. Taking $G' = G \cup H$ directly yields G' non-null for which $E \cup G' \succ_*^* F \cup G'$.

($\Rightarrow$) Since $F \subset E$, for $G'' = E^c \cap F^c = E^c$ we have $E \cup G'' \succ_*^* F \cup G''$. Taking $G''' = G'' \backslash H$

gives G''' non-null for which $E \cup G''' \succ_*^* F \cup G'''$.

Step 5:

1. Define a set function $\lambda : \mathcal{E} \to [0,1]$ as follows: let $\lambda(\varnothing) = 0$, $\lambda(\mathcal{S}) = 1$; and whenever $F \subseteq E$, let $\lambda(F) \leq \lambda(E)$ with a strict inequality if E/F is non-null.

2. Notice that any function λ that satisfies the conditions from Step 5.1 represents the relation $\succ_*^*$ on $\mathcal{E}$ under (**H**) or (**L**): For all $E \in \mathcal{E}$, $E \succ_*^* F \Leftrightarrow \lambda(E) > \lambda(F)$. Denote by $\Lambda(\mathcal{E})$ the set of all measures λ that represent $\succsim_*^*$ under (**H**) or (**L**).

3. Each measure $\lambda \in \Lambda(\mathcal{E})$ is nonatomic. We will prove this for (**H**). Fix $\lambda \in \Lambda(\mathcal{E})$ and consider an event $E \in \mathcal{E}|_{\succ_*^*\varnothing}^{\prec_*^*\mathcal{S}}$. By Steps 2 and 5.1, $\lambda(E) > 0$ and $\lambda(E^c) > 0$. Using P5, construct a pair of acts

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \notin \mathcal{S} \\ y \text{ if } s \in \mathcal{S} \end{bmatrix}. \tag{74}$$

Applying the definition of $\succ_*$, $E^c \prec_* \mathcal{S}$, and by P6^{Q_*} event E can be partitioned into F and $E \backslash F$, such that $E^c \cup F \prec_* \mathcal{S}$. Both F and $E \backslash F$ are necessarily non-null, for otherwise $E^c \cup F \sim_* \mathcal{S}$ or $E^c \cup (E \backslash F) \sim_* \mathcal{S}$ – a contradiction to P6^{Q_*}. By Step 3 and Step 5.1, $\lambda(E) > \lambda(F)$ and $\lambda(E) > \lambda(E \backslash F)$. This completes the proof.[56]

Step 7:

1. Fix a (possibly non-additive) measure $\lambda \in \Lambda(\mathcal{E})$. Consider two lotteries $P, Q \in \mathcal{P}_0(\mathcal{X})$, such that $P = Q$ and $\lambda \circ f^{-1} = P$, $\lambda \circ g^{-1} = Q$ for some $f, g \in \mathcal{F}$. Since the least preferred (w.r.t. $\succ_P$) outcomes assigned some positive probability by lotteries P and Q are identical, and equal to the least preferred (w.r.t. $\succ$) outcomes mapped from non-null events by acts f and g, condition (24) follows for (**L**). Similarly, the most preferred (w.r.t. $\succ_P$) outcomes in the supports of P and Q are equal and coincide with the most preferred (w.r.t. $\succ$) outcomes assigned to non-null events by acts f and g. Hence, condition (24) follows for (**H**).

2. If λ is convex-ranged, then for any $P \in \mathcal{P}_0(\mathcal{X})$, there exists an act $f \in \mathcal{F}$, such that $\lambda \circ f^{-1} = P$. Hence, using in addition that $\succ$ is a weak order, the stronger version of probabilistic sophistication from Section 4.1 is also satisfied. ■

[56] The representing measures need not be convex-ranged. Consider a monotone measure $\tilde{\lambda}$ that assigns $\tilde{\lambda}(\varnothing) = 0$, $\tilde{\lambda}(\mathcal{S}) = 1$, and a maximum of 0.9 to any event E for which $\mathcal{S} \backslash E$ is non-null. That measure represents the relation $\succ_*^*$ on $\mathcal{E}$ as desired; but $\tilde{\lambda}$ is clearly not convex-ranged.

1D. Proofs of Lemmas 3 and 4

Lemma 3A *If the binary relation over acts, $\succ$, is not extreme, then for all events $E, F \in$* $\mathcal{E}_{**}$, such that $E \sim_{**} F$, and all acts $h \in \mathcal{F}$,

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \in F \\ h \text{ if } s \notin E \cup F \end{bmatrix} \sim \begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \in E \\ h \text{ if } s \notin E \cup F \end{bmatrix}. \tag{75}$$

Proof. Assume that the relation $\succ$ is not extreme. Let π be the probability measure derived in Theorem 1. Take a pair of equal-probability events: $E, F \in \mathcal{E}_{**}$, $\pi(E) = \pi(F)$. Define $\widetilde{E} = E \backslash F$ and $\widetilde{F} = F \backslash E$. By additivity of π, $\pi(\widetilde{E}) = \pi(\widetilde{F})$. Consider acts

$$f = \begin{bmatrix} x \text{ if } s \in \widetilde{F} \\ y \text{ if } s \in \widetilde{E} \\ h \text{ if } s \notin \widetilde{E} \cup \widetilde{F} \end{bmatrix} = \begin{bmatrix} h_{x+} \text{ if } s \in G_1 \\ x \text{ if } s \in \widetilde{F} \\ h_{x-,y+} \text{ if } s \in G_2 \\ y \text{ if } s \in \widetilde{E} \\ h_{y-} \text{ if } s \in G_3 \end{bmatrix} \text{ and } g = \begin{bmatrix} x \text{ if } s \in \widetilde{E} \\ y \text{ if } s \in \widetilde{F} \\ h \text{ if } s \notin \widetilde{E} \cup \widetilde{F} \end{bmatrix} = \begin{bmatrix} h_{x+} \text{ if } s \in G_1 \\ x \text{ if } s \in \widetilde{E} \\ h_{x-,y+} \text{ if } s \in G_2 \\ y \text{ if } s \in \widetilde{F} \\ h_{y-} \text{ if } s \in G_3 \end{bmatrix}. \tag{76}$$

Using the definition of $\sim_{**}$,

$$\widetilde{E} \sim_{**} \widetilde{F} \text{ if for any } x \succ y, \begin{bmatrix} x \text{ if } s \notin \widetilde{E} \cup G \\ y \text{ if } s \in \widetilde{E} \cup G \end{bmatrix} \sim \begin{bmatrix} x \text{ if } s \notin \widetilde{F} \cup G \\ y \text{ if } s \in \widetilde{F} \cup G \end{bmatrix}, \text{ for any } G \subseteq (\widetilde{E}^c \cap \widetilde{F}^c). \tag{77}$$

Lemma 1, invoked to define the likelihood relation $\sim_{**}$ and to reduce the cardinality of the outcome sets $f(\mathcal{S})$ and $g(\mathcal{S})$ to two, implies $f \sim g$. ■

Lemma 3B *If the binary relation over acts, $\succ$, is extreme, then for all non-null events* $E, F \in \mathcal{E}_{**}$, $E \cap F = \varnothing$, and all acts $h \in \mathcal{F}$,

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \in F \\ h \text{ if } s \notin E \cup F \end{bmatrix} \sim \begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \in E \\ h \text{ if } s \notin E \cup F \end{bmatrix}. \tag{78}$$

Proof. Pick two non-null events E' and F', $E' \cap F' \neq \varnothing$ and assume that (**H**) holds. Consider the following acts:

$$f = \begin{bmatrix} x \text{ if } s \in F' \\ y \text{ if } s \in E' \\ h \text{ if } s \notin E' \cup F' \end{bmatrix} \text{ and } g = \begin{bmatrix} x \text{ if } s \in E' \\ y \text{ if } s \in F' \\ h \text{ if } s \notin E' \cup F' \end{bmatrix}, \; x \succ y, \tag{79}$$

where $h(\mathcal{S}) \prec x$. Since event F' is pivotal in act f, event E' is pivotal in act g, and $f \sim g$, as desired. The argument is analogous under (**L**). ∎

In the next lemma and its corollary, we only analyze the nontrivial case of the non-extreme preferences.

Lemma 4 *In the coarsest measurable partition of the state space $\mathcal{S}$ induced by act $f \in \mathcal{F}$, there is a unique pivotal event.*

Proof. We first show that the property of being pivotal is state-independent. Next, we demonstrate that the pivotal event is (2) unique to an act, given the coarsest measurable partition induced by that act.

1. We first state the key assertion, implied by Lemma 3A,B: Let π be the probability measure derived in Theorem 1. Consider act $f \in \mathcal{F}$, such that E is the pivotal event of f, and for a disjoint with E event F, $\pi(E) = \pi(F)$ and $f^{-1}(E) = x \nsim f^{-1}(F)$ for some $x \in f(\mathcal{S})$. (Such a pair of events with $\pi(E) = \pi(F)$ and $\pi(E) + \pi(F) \leq 1$ can be constructed by convex-rangedness of π.) Then, swapping outcomes between events E and F yields act g, such that $g^{-1}(F) = x$, $g^{-1}(E) = f^{-1}(F)$ and $g \sim f \sim x$. That is, the property of being pivotal is state-independent.

2. Take act $f \in \mathcal{F}$ and let E and F be disjoint measurable events that map to non-indifferent outcomes: $f^{-1}(E) = x \nsim y = f^{-1}(F)$. Suppose these events are both pivotal to act f. Applying P3^Q twice to f and invoking P1 yields a contradiction:

$$x \sim \begin{bmatrix} x \text{ if } s \in E \\ x \text{ if } s \in E^c \end{bmatrix} \sim_{\text{Pivotal } E} f = \begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \in F \\ h \text{ if } s \notin E \cup F \end{bmatrix} \sim_{\text{Pivotal } F} \begin{bmatrix} y \text{ if } s \in F^c \\ y \text{ if } s \in F \end{bmatrix} \sim y, \quad (80)$$

where $h \in \mathcal{F}$ and $h \nsim x$, $h \nsim y$. ∎

Corollary 2 *Let E be the pivotal event of act $f \in \mathcal{F}$ and $f(E) = x$ for some $x \in f(\mathcal{S})$. Then,*

$$\textit{either } f = \begin{bmatrix} g_{x+} \textit{ if } s \in E_{fx+} \\ x \textit{ if } s \in E \\ g_{x-} \textit{ if } s \in E_{fx-} \end{bmatrix} \succ \begin{bmatrix} g_{x+} \textit{ if } s \in E_{fx+} \\ x \textit{ if } s \in E_1 \\ g_{x-} \textit{ if } s \in E_2 \cup E_{fx-} \end{bmatrix} \textit{ or } f \sim \begin{bmatrix} g_{x+} \textit{ if } s \in E_{fx+} \\ x \textit{ if } s \in E_1 \\ g_{x-} \textit{ if } s \in E_2 \cup E_{fx-} \end{bmatrix}. \quad (81)$$

Proof. Pick an act $f \in \mathcal{F}$, let $x \in f(\mathcal{S})$ be the outcome on the pivotal event of f, E'. Using P3^Q, reduce the cardinality of $f(\mathcal{S})$ to two so that $f = [x, E; y, E^c]$, $E' \subseteq E'$.

Partition event E^c into non-null events F_1 and F_2. Since $f \sim x$, it cannot be the case that

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \in E \\ x \text{ if } s \in F_1 \\ y \text{ if } s \in F_2 \end{bmatrix} \text{ or } \begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \prec \begin{bmatrix} x \text{ if } s \in E \\ x \text{ if } s \in F_1 \\ x \text{ if } s \in F_2 \end{bmatrix}. \tag{82}$$

Hence, for E_1 and E_2, such that $E_1 \cup E_2 = E$,

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \in E_1 \\ y \text{ if } s \in E_2 \\ y \text{ if } s \notin E \end{bmatrix} \text{ or } \begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \sim \begin{bmatrix} x \text{ if } s \in E_1 \\ y \text{ if } s \in E_2 \\ y \text{ if } s \notin E \end{bmatrix}. \tag{83}$$

More can be said if the cardinality of the set of outcomes $\mathcal{X}$ is at least four[57]. Applying P6^Q, partition the pivotal event E' of f into non-null events E'_1 and E'_2. Consider act g:

$$\begin{bmatrix} g_{x+} \text{ if } s \in E_{x+} \\ x \text{ if } s \in E'_1 \\ x' \text{ if } s \in E'_2 \\ g_{x'-} \text{ if } s \in E_{x-} \end{bmatrix}, \tag{84}$$

where $x \succ x'$, $g_{x+}(\mathcal{S}) \succ x$ and $g_{x'-}(\mathcal{S}) \prec x'$. By P3^Q, it must be that either E'_1 or E'_2 are pivotal in g and by Lemma 4, only one of E'_1 and E'_2 can be pivotal in (84). ■

[57]With a three-outcome $\mathcal{X}$, let $z \succ x \succ y$. Take act $f \in \mathcal{F}$ and let event E be pivotal to f, with $f^{-1}(x) = E$. Next, patrition E into E_1 and E_2. By P3^Q and P1, it must be that

$$f = \begin{bmatrix} z \text{ if } s \in E_{x+} \\ x \text{ if } s \in E \\ y \text{ if } s \in E_{x-} \end{bmatrix} \sim \begin{bmatrix} z \text{ if } s \in E_{x+} \\ x \text{ if } s \in E_1 \\ y \text{ if } s \in E_2 \cup E_{x-} \end{bmatrix} \prec \begin{bmatrix} z \text{ if } s \in E_1 \cup E_{x+} \\ x \text{ if } s \in E_2 \\ y \text{ if } s \in E_{x-} \end{bmatrix}$$

or

$$f \sim \begin{bmatrix} z \text{ if } s \in E_1 \cup E_{x+} \\ x \text{ if } s \in E_2 \\ y \text{ if } s \in E_{x-} \end{bmatrix} \succ \begin{bmatrix} z \text{ if } s \in E_{x+} \\ x \text{ if } s \in E_1 \\ y \text{ if } s \in E_2 \cup E_{x-} \end{bmatrix}.$$

Appendix 2: Proof of Theorem 2

Proof. Sufficiency: (1) $\Rightarrow$ (2)

Step 1 of the sufficiency part defines a preference relation $\succ_P$ over probability distributions induced by the relation $\succ$ over acts. Step 2 establishes the existence and uniqueness of a number τ. Step 3 then constructs a preference functional over probability distributions that represents $\succ_P$ as a left-continuous τ^{th} quantile. Step 4 builds on the derived preference functional for distributions, $\succ_P$, to derive a representation for the relation over acts, $\succ$.

Step 1 (**DERIVATION OF** $\succ_P$):

1. Let π be the probability measure derived for the non-extreme preferences in Theorem 1. The measure π can be used to map each act $f \in \mathcal{F}$ to some probability distribution $P \in \mathcal{P}_0(\mathcal{X})$ through the mapping $\pi \circ f^{-1} = P$. This mapping induces a relation over probability distributions in $\mathcal{P}_0(\mathcal{X})$, $\succ_P$, from the relation over acts, $\succ$, as defined in Section 4.1. Probabilistic sophistication (Step 7 of Theorem 1), convex-rangedness of π and P1 imply that the mapping from acts to simple probability distributions is onto. Hence, the relation $\succ_P$ is asymmetric and negatively transitive, and such that for all P, Q in $\mathcal{P}_0(\mathcal{X})$ and all f, g in $\mathcal{A}$,

$$\left(P \succsim_P Q,\ \pi \circ f^{-1} = P,\ \pi \circ g^{-1} = Q\right) \Leftrightarrow f \succsim g. \tag{85}$$

2. For the extreme preferences, fix a measure $\lambda \in \Lambda(\mathcal{E})$ and map acts in $\mathcal{F}$ to lotteries in $\mathcal{P}_0(\mathcal{X})$ through the mapping $f \rightarrow \lambda \circ f^{-1}$, defined in Section 4.1.

Step 2 (**EXISTENCE AND UNIQUENESS OF** τ):

Assume first that the relation $\succ$ is not extreme.

1. We will repeatedly use that whenever $F \succ_{***} \varnothing$,[58] then for any $N \in \mathbb{N}_{++}$, there exists a 2^N-partition of F, $\{F_1^{2^N}, ..., F_{2^N}^{2^N}\}$, such that $F_1^{2^N} \sim_{***} ... \sim_{***} F_n^{2^N} \sim_{***} ... \sim_{***} F_{2^N}^{2^N}$. Given that axioms A1-A5' hold on the set $\mathcal{E}$, such a partition, referred to as a *uniform* 2^N*-partition* of F, can be derived by applying the argument in Fishburn (1970, Ch.14.2).[59]

2. Fix $N \in \mathbb{N}_{++}$ and consider a uniform 2^N-partition of the state space $\mathcal{S}$ and associate with it a sequence of acts $\{f(n|N)\}_{n=1,...,2^N}$, where

$$f(n|N) = \left[\begin{array}{l} x \text{ if } s \in \bigcup\limits_{l=n+1,...,2^N} F_l^{2^N} \\ y \text{ if } s \in \bigcup\limits_{l=1,...,n} F_l^{2^N} \end{array}\right],\ n = 1, ..., 2^N,\ x \succ y. \tag{86}$$

Corollary 2 can be applied recursively to the sequence of acts $\{f(n|N)\}_{n=1,...,2^N}$ to es-

[58]The relation $\succ_{***}$ is defined in Step 6.2 in the proof of Theorem 1.

[59]Alternatively, the convex-rangedness of π could be used directly to derive 2^N-partitions.

tablish that there is a unique $n(N) \in \{1, ..., 2^N\}$ for which

$$\begin{bmatrix} x \text{ if } s \in \mathcal{S} \\ y \text{ if } s \notin \mathcal{S} \end{bmatrix} \sim ... \sim \begin{bmatrix} x \text{ if } s \notin \bigcup\limits_{l=1,...,n(N)-1} F_l^{2^N} \\ y \text{ if } s \in \bigcup\limits_{l=1,...,n(N)-1} F_l^{2^N} \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \notin \bigcup\limits_{l=1,...,n(N)} F_l^{2^N} \\ y \text{ if } s \in \bigcup\limits_{l=1,...,n(N)} F_l^{2^N} \end{bmatrix} \sim ... \sim \begin{bmatrix} x \text{ if } s \notin \mathcal{S} \\ y \text{ if } s \in \mathcal{S} \end{bmatrix}. \tag{87}$$

Now, construct a sequence of such events, one for every N, $\left\{ \bigcup\limits_{l=1,...,n(N)} F_l^{2^N} \right\}_{N \in \mathbb{N}_{++}}$. Applying Lemma 4 again, these events are nested and weakly decreasing in N. By properly choosing a subsequence, we may define $\tau = \lim_{N\to\infty} \pi \left(\bigcup\limits_{l=1,...,n(N)} F_l^{2^N} \right) = \bigcap\limits_N \pi \left(\bigcup\limits_{l=1,...,n(N)} F_l^{2^N} \right)$, $\tau \in (0, 1)$.

3. For the extreme preferences, set $\tau = 0$ if (**L**) holds and $\tau = 1$ if (**H**) holds.

Step 3 (**REPRESENTATION FUNCTIONAL FOR** $\succ_P$):

1. By P3^Q, there is a one-to-one mapping between the sets of equivalence classes on the set of acts $\mathcal{F}$ w.r.t. $\succ$ and the outcome set $\mathcal{X}$ w.r.t. $\succ_x$; that is, for all pairs of acts $f, g \in \mathcal{F}$,

$$f \succ g \Leftrightarrow x \succ y \tag{88}$$

where x is the outcome mapping from the pivotal event by act f, while y – by act g.

Assume that preferences are not extreme. Using the derived preference relation $\succ_P$ over simple distributions in $\mathcal{P}_0(\mathcal{X})$ and the properties of π, (88) is equivalent to

$$\pi \circ f^{-1} \succ_P \pi \circ g^{-1} \Leftrightarrow \delta_x \succ_P \delta_y. \tag{89}$$

The set of equivalence classes on $\mathcal{P}_0(\mathcal{X})$ w.r.t. $\succ_P$ can thus be mapped onto the set of equivalence classes on $\mathcal{X}$ (understood as the set of constant distributions) w.r.t. $\succ_P$. Hence, the certainty-equivalence mapping for distributions – each simple distribution is $\succ_P$-indifferent to an outcome in its support – can be used to construct a representation for the relation $\succ_P$ on $\mathcal{P}_0(\mathcal{X})$. The latter, in turn, can then be used to provide a representation for $\succ$ on $\mathcal{F}$.

If preferences are extreme, for any measure $\lambda \in \Lambda(\mathcal{E})$, (88) is equivalent to

$$\lambda \circ f^{-1} \succ_P \lambda \circ g^{-1} \Leftrightarrow \delta_x \succ_P \delta_y. \tag{90}$$

For a fixed measure $\lambda \in \Lambda(\mathcal{E})$, define a subset of lotteries in $\mathcal{P}_0(\mathcal{X})$:

$$\mathcal{P}_0(\mathcal{X}, \lambda) = \left\{ P \in \mathcal{P}_0(\mathcal{X}) | P = \lambda \circ f^{-1} \text{ for some } f \in \mathcal{F} \right\}. \tag{91}$$

The set of equivalence classes on $\mathcal{P}_0(\mathcal{X}, \lambda)$ w.r.t. $\succ_P$ can now be mapped onto the set of equivalence classes on $\mathcal{X}$. Again, the certainty-equivalence mapping for distributions can be used to construct a representation for the relation $\succ_P$ on $\mathcal{P}_0(\mathcal{X},\lambda)$, and back up from it a representation for $\succ$ on $\mathcal{F}$.

2. The remaining steps characterize the certainty equivalence map between lotteries and outcomes as a (generalized) inverse of a distribution and establish that it represents the preference relation $\succ_P$. The unique number $\tau\in[0,1]$ derived in Step 2 will be used in defining the inverse equal to the τ^{th} quantile of the distribution.

By Step 2, for any act $f\in\mathcal{F}$,

$$f\sim x,\text{ where } x \text{ is such that } \pi(f(s)\precsim x)\geq\tau \tag{92}$$

and x is (one of) the least preferred outcome(s) in $\{y\in f\left(\mathcal{S}\right)|\pi(f(s)\precsim y)\geq\tau;\ \pi\left(f^{-1}\left(y\right)\right)>0\}$. Further, using the definition of $\succ_P$, P1 and P3^Q, it is straightforward to show that acts that imply indifferent τ^{th} outcomes are indifferent.

3. We verify that the inverse to-be-defined (in Step 3.5) should be left continuous. Let τ be the number from $[0,1]$, derived in Step 2. Consider the sequence of acts $\{f(n|N)\}_{n=1,\dots,2^N}$ constructed in Step 2.2. For any $N\in\mathbb{N}_{++}$, define $\tau_n=\pi\left(\bigcup_{l=1,\dots,n(N)-1}F_l^{2^N}\right)$. By properly choosing a subsequence, we obtain $\tau_n\to\tau$ as $N\to\infty$; and for any given $N\in\mathbb{N}_{++}$, $f\left(\bigcup_{l=1,\dots,n(N)-1}F_l^{2^N}\right)=y$ and $f\left(F_{n(N)}^{2^N}\right)=y$. Together with the fact that the sequence of events $\left\{\bigcup_{l=1,\dots,n(N)}F_l^{2^N}\right\}_{N\in\mathbb{N}_{++}}$ used to derive τ is weakly decreasing, that gives that the inverse to-be-defined should be left continuous.

4. For the preference functional on lotteries to be real-valued, it suffices to ensure that there exists a real-valued utility function on outcomes, $u:\mathcal{X}\to\mathbb{R}$. Given that $\succ$ on $\mathcal{X}$ is a weak order (P1) and $\mathcal{X}$ contains a countable $\succ$-order dense subset, a standard argument (Debreu [1954]) delivers a real-valued utility function $u(\cdot)$ on $\mathcal{X}$, unique up to a strictly increasing transformation. Let $\mathcal{U}^O$ be the set of all such functions u that represent $\succ_x$.

5. Fix utility $u\in\mathcal{U}^O$, the number $\tau\in[0,1]$, the measure π for the non-extreme preferences and the set of capacities $\Lambda\left(\mathcal{E}\right)$ for the extreme preferences. Define ν as $\nu=\pi$ if $\succ_P$ is non-extreme and $\nu=\lambda$ if $\succ_P$ is extreme. For any $P=\nu\circ f$, with an outcome set $\{x_1,\dots,x_N\}$ define $V:\mathcal{P}_0(\mathcal{X})\to\mathcal{X}$ as

$$V\left(P\right)=\begin{cases}\inf\{z\in\mathbb{R}|\pi[u(x_n)\leq z|x_n\in f\left(\mathcal{S}\right)]\geq\tau\}\text{ if }\tau\in(0,1);\\ \sup\{z\in\mathbb{R}|\lambda[u(x_n)\leq z|x_n\in f\left(\mathcal{S}\right)]\leq 0\}\text{ if }\tau=0;\\ \inf\{z\in\mathbb{R}|\lambda[u(x_n)\leq z|x_n\in f\left(\mathcal{S}\right)]\geq 1\}\text{ if }\tau=1.\end{cases} \tag{93}$$

where the definition for the extreme preferences holds for all $\lambda\in\Lambda\left(\mathcal{E}\right)$.

Step 4 (**REPRESENTATION FUNCTIONAL FOR** $\succ$): We can now combine the above steps to define a functional $\mathcal{V} : \mathcal{F} \to \mathcal{X}$ that represents the relation $\succ$ on acts: By Step 1 for all $f, g \in \mathcal{F}$, and all $P, Q \in \mathcal{P}_0(\mathcal{X})$, such that $P = \pi \circ f$ and $Q = \pi \circ g$,

$$f \succ g \Leftrightarrow P \succ Q, \tag{94}$$

which by Steps 2 and 3 is equivalent to

$$\mathcal{V}(f) = V(P) > V(Q) = \mathcal{V}(g). \tag{95}$$

That is, as desired, the preference relation $\succ$ on $\mathcal{F}$ can be represented by evaluating each act $f \in \mathcal{F}$ by the τ^{th} quantile of the distribution induced by act f and measures π for the non-extreme and $\Lambda(\mathcal{E})$ for the extreme preferences:

$$\mathcal{V}(f) = \begin{cases} \inf\{z \in \mathbb{R} | \pi[u(f(s)) \leq z] \geq \tau\} \text{ if } \tau \in (0,1); \\ \sup\{z \in \mathbb{R} | \pi[u(f(s)) \leq z] \leq 0\} \text{ if } \tau = 0; \\ \inf\{z \in \mathbb{R} | \lambda[u(f(s)) \leq z] \geq 1\} \text{ if } \tau = 1. \end{cases} \tag{96}$$

■

From the analysis so far, a natural local notion of monotonicity suggests itself for the Quantile Maximization model.

Definition 6 *Fix* $\tau \in [0,1]$. *Given a complete preorder over outcomes* $\succsim_x$, *distribution* $Q = (y_1, q_1; ...; y_M, q_M)$ τ-first-order stochastically dominates *(*τ-FOSD*) distribution* $R = (y_1, r_1; ...; y_M, r_M)$ *with respect to* $\succsim_x$ *if*

$$V(Q) > V(R), \tag{97}$$

where $V(\cdot)$ *is as defined in (93).*

$\succsim_P$ is said to satisfy τ*-first order stochastic dominance* if $P \succ_P Q$ whenever P τ-FOSD Q with respect to $\succsim_x$. Unlike FOSD, τ-FOSD gives rise to a complete ranking of distributions.

Proof. Necessity: Assume that the representation $\mathcal{V}(f)$ holds for $\succ$.

(2) $\Rightarrow$ (1) Fix a number $\tau \in [0,1]$, a measure π for $\tau \in (0,1)$ and a set of capacities $\Lambda(\mathcal{E})$ for τ equal to 0 or 1; for the extreme quantiles, the arguments below hold for all capacities in $\Lambda(\mathcal{E})$. The proof assumes a given utility $u \in \mathcal{U}^O$. Showing that conditions (**L**) and (**H**) hold for the representation with $\tau = 0$ and $\tau = 1$, respectively, is straightforward and is omitted here.

P1 (**ORDERING**): This holds, since there is a real-valued representation of $\succ$.

P3^Q (**PIVOTAL MONOTONICITY**):

(*only if*) Pick an act $f \in \mathcal{F}$. By τ-FOSD, $f \sim x$ for an outcome $x \in \mathcal{X}$, such that $\mathcal{V}(f) = x$. That the event to which x is mapped by act f, $f^{-1}(x) = E$, is non-null follows from the representation. By τ-FOSD, E is such that $f \sim [g_{x+} \text{ if } E_{fx+}; x \text{ if } E;\ g_{x-} \text{ if } E_{fx-}]$. Consider outcome $y \precsim x$. Then, appealing to τ-FOSD again yields $[g_{x+} \text{ if } E_{fx+}; x \text{ if } E; g_{x-} \text{ if } E_{fx-}] \succsim [g_{y+} \text{ if } E_{fx+}; y \text{ if } E; g_{y-} \text{ if } E_{fx-}]$ for any subacts $g_{x+},\ g_{x-},\ g_{y+},$ and g_{y-}, as desired.

(*if*) Implied by τ-FOSD.

P4^Q (**COMPARATIVE PROBABILITY**): Assume $\tau \in (0,1)$. Pick disjoint events E and F, outcomes $x^* \succ x$, and the following acts

$$\begin{bmatrix} x^* \text{ if } s \in E \\ x \text{ if } s \in F \\ g \text{ if } s \notin E \cup F \end{bmatrix} \succ \begin{bmatrix} x^* \text{ if } s \in F \\ x \text{ if } s \in E \\ g \text{ if } s \notin E \cup F \end{bmatrix}. \tag{98}$$

Define $(E \cup F)^c_{gx-} = \{s \in \mathcal{S} | g(s) \prec x\}$. Then, for (98), τ-FOSD implies that $\pi((E \cup F)^c_{gx-}) + \pi(F) < \tau \leq \pi((E \cup F)^c_{gx-}) + \pi(E)$, and hence, $\pi(E) > \pi(F)$. Reversing the above argument, it must be that

$$\begin{bmatrix} x^* \text{ if } s \in E \\ x \text{ if } s \in F \\ h \text{ if } s \notin E \cup F \end{bmatrix} \succsim \begin{bmatrix} x^* \text{ if } s \in F \\ x \text{ if } s \in E \\ h \text{ if } s \notin E \cup F \end{bmatrix}, \tag{99}$$

where $(E \cup F)^c_{hx-} = \{s \in \mathcal{S} | h(s) \prec x\}$, and we used that $\pi((E \cup F)^c_{hx-}) + \pi(F) < \pi((E \cup F)^c_{hx-}) + \pi(E)$.

For $\tau \in 0$, (98) can hold only if event F is null, event E is non-null and $g(s) \succ x$ for every $s \in (E \cup F)^c$; similarly, for $\tau \in 1$, (98) can hold only if event F is null, event E is non-null and $g(s) \prec x^*$ for every $s \in (E \cup F)^c$. In either case, the assertion follows immediately.

P5 (**NONDEGENERACY**): This follows, since the functional $\mathcal{V} : \mathcal{F} \to \mathbb{R}$ is nonconstant.

P6^Q (**SMALL-EVENT CONTINUITY OF** $\succsim_l$): Let $\tau \in (0,1)$. Suppose that for all $x \succ y$,

$$\begin{bmatrix} x \text{ if } s \in E \\ y \text{ if } s \notin E \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \in F \\ y \text{ if } s \notin F \end{bmatrix} \tag{100}$$

By the definition of the relation $\succ^*$, $E \succ^* F$, and hence, given the measure π and the definition of the relation $\succ_{***}$ which it represents, $\pi(E) > \pi(F)$. Further, the representation of $\succ$ implies that $1 - \pi(F) \geq \tau > 1 - \pi(E)$. By nonatomicity of the measure π, we can partition the state space $\mathcal{S}$ into N events $\{H_1, ..., H_N\}$ and choose N, such that $\tau > 1 - \pi(E) + \pi(H_n)$ for all $n = 1, ..., N$. The definitions of $\succ^*$ and $\succ_{***}$, then, imply that $E \succ^* F \cup H_n$, for all $n = 1, ..., N$. By a similar argument, P6$^{Q^*}$ holds for any event $E \in \mathcal{E}$ and $\varnothing$.

For (**L**), (100) is only satisfied when F^c is non-null and E^c is null. Then, $\lambda(E) = 1 >$

$\lambda(F)$ for any $\lambda \in \Lambda(\mathcal{E})$. Nonatomicity of measures λ in $\Lambda(\mathcal{E})$ completes the proof. So does it for (**H**), in which case (100) is only satisfied when E is non-null and F is null; that is, $\lambda(F^c) = 1 > \lambda(E^c)$. ■

Appendix 3: Proof of Theorem 3

The necessity part ((2) $\Rightarrow$ (1)) is analogous to the proof of Theorem 2, and therefore we focus on the sufficiency part ((1) $\Rightarrow$ (2)) of Theorem 3. Since the proof of the result for the extreme preferences is provided in Appendices 1 and 2, here we assume that the relation $\succ$ is non-extreme. The numbering of (as well as the logic behind) steps follows that of the proof of Theorem 1.

Proof. Steps 1-7 derive and characterize beliefs induced by the relation $\succ$ under axioms P1, P3^Q, P5 and P6^Q. Step 8 constructs the representation functional for $\succ$.

Steps 1-3: As in the proof of Theorem 1. However, in the absence of P4^Q, Lemma 7 does not hold. More precisely, it only holds for nested events. The relation $\succ_{**}$ is, thus, not well defined on all of $\mathcal{E}_{**}$. Therefore, we proceed differently and define a new binary relation $\succ'_{**}$ for nested events in $\mathcal{E}$.

Definition 7 *For $E, F \in \mathcal{E}$, such that $F \subset E$,*

$$\begin{array}{ll} E \succ'_{**} F \text{ if } & \text{there is an event } G \subseteq (E \cup F)^c = E^c, \text{ such that } E \cup G \succ_* F \cup G \\ & \text{or if there is an event } H \subseteq E \cap F = F, \text{ such that } E \backslash H \succ_* F \backslash H. \end{array} \tag{101}$$

To argue that the relation $\succ'_{**}$ is well-defined, we first observe that Lemma 7 does hold for nested events in $\mathcal{E}$:

$$\begin{array}{c} \text{For all } F, E \in \mathcal{E}, \text{ such that } F \subset E, \text{ if } E \sim_* F \text{ and there is a non-null event } G \subseteq (E \cup F)^c = E^c, \\ \text{such that } E \cup G \succ_* F \cup G, \text{ then there is no event } G' \subseteq (E \cup F)^c = E^c \\ \text{for which } E \cup G' \prec_* F \cup G'. \end{array} \tag{$*$}$$

By an argument similar to the one used in proving Lemma 7, $(*)$ only has a bite for events in $\mathcal{E}|_{\sim_*\varnothing}$: If $F, E \in \mathcal{E}|_{\sim_*\mathcal{S}}$ and $F \subset E$, then for all $G'' \subseteq E^c$, $F \cup G'' \sim_* F \cup G'' \sim_* \mathcal{S}$. Nonetheless, one can prove the following:

$$\begin{array}{l} \text{For all } F, E \in \mathcal{E}, \text{ such that } F \subset E, \text{ if } E \sim_* F \text{ and there is a non-null event } H \subseteq E \cap F = F, \\ \text{such that } E \backslash H \succ_* F \backslash H, \text{ then there is no event } H' \subseteq E \cap F = F, \text{ such that } E \backslash H' \prec_* F \backslash H'. \end{array} \tag{$**$}$$

Again, although $(**)$, like $(*)$, holds for nested events in all of $\mathcal{E}$, it clearly only has a bite for events in $\mathcal{E}|_{\sim_*\mathcal{S}}$: If $F, E \in \mathcal{E}|_{\sim_*\varnothing}$ and $F \subset E$, then for all $H'' \subseteq F$, $F \backslash H'' \sim_* F \backslash H'' \sim_* \varnothing$. Furthermore, it is immediate that if $F \subset E$, $E \sim_* F$ and $E \cup G \succ_* F \cup G$ for some

$G \subseteq (E \cup F)^c = E^c$, then there is no $H \subseteq E \cap F = F$, such that $E \backslash H \prec_* F \backslash H$, and vice versa.

Step 4: Showing that A1-A5' hold for the relation $\succ'_{**}$ can proceed analogously to the proof of Theorem 1, and is therefore omitted. Observe that part "if" of A4 is non-trivial only for events in $\mathcal{E}|_{\sim_*\varnothing}$, while part "only if" – for events in $\mathcal{E}|_{\sim_*\mathcal{S}}$.

Step 5:

1. Define a set function $\lambda : \mathcal{E} \to [0, 1]$ as follows: let $\lambda(\varnothing) = 0$, $\lambda(\mathcal{S}) = 1$; and whenever $F \subseteq E$, let $\lambda(F) \leq \lambda(E)$ with a strict inequality if E/F is non-null. The function λ is additive for nested events: If $F \subset E$, then $\lambda(F) + \lambda(E \backslash F) = \lambda(E)$. Notice that any function λ that satisfies these conditions represents the relation $\succ'_{**}$ on $\mathcal{E}$: For all $E, F \in \mathcal{E}$ with $F \subset E$, $E \succ'_{**} F \Leftrightarrow \lambda(E) > \lambda(F)$. Denote by $\Lambda(\mathcal{E})$ the set of all measures λ that represent $\succ'_{**}$ in that sense.

2. We show that each measure $\lambda \in \Lambda(\mathcal{E})$ is nonatomic. For events $E \in \mathcal{E}|_{\sim_*\mathcal{S}}$, it follows directly from $\mathrm{P6}^{Q_*}$ that E can be partitioned into non-null events F and $E \backslash F$, such that $\lambda(E) > \lambda(F)$ and $\lambda(E) > \lambda(E \backslash F)$. Next, pick an event H from $\mathcal{E}|_{\sim_*\varnothing}$. Using that $H^c \in \mathcal{E}|_{\sim^*\mathcal{S}}$, and applying $\mathrm{P6}^{Q^*}$ to H^c and $\varnothing$, take $G \in \mathcal{E}|_{\sim_*\varnothing}$, such that

$$\begin{bmatrix} x \text{ if } s \notin H \\ y \text{ if } s \in H \end{bmatrix} \succ \begin{bmatrix} x \text{ if } s \notin H \cup G \\ y \text{ if } s \in H \cup G \end{bmatrix} \tag{102}$$

(that this is possible follows from the non-nullness of H.) Then, $\mathrm{P6}^{Q_*}$ implies that H can be partitioned into non-null events F' and $H \backslash F'$, such that $\lambda(H) > \lambda(F')$ and $\lambda(H) > \lambda(H \backslash F')$.

Step 7: Establishing that the relation $\succ$ is probabilistically sophisticated can proceed in a manner analogous to Step 7 of Theorem 1 for the extreme preferences.

Step 8: This final step constructs the representation functional for $\succ$.

1. Fix a measure $\lambda \in \Lambda(\mathcal{E})$ and map acts in $\mathcal{F}$ to lotteries in $\mathcal{P}_0(\mathcal{X})$ through the mapping $f \to \lambda \circ f^{-1}$.

2. For any act $f \in \mathcal{F}$, the certainty equivalents of a lottery $\lambda \circ f^{-1}$ coincide for all measures in $\Lambda(\mathcal{E})$. Taking an additive measure from $\Lambda(\mathcal{E})$, one can apply the derivation of $\tau \in (0, 1)$ from the proof of Theorem 2 for the non-extreme preferences.

3. The argument constructing a left-continuous quantile functional (29) that represents $\succ$ parallels the one for the extreme preferences and is, therefore, omitted. ■

Appendix 4: Other Results

Proposition 1*Assume P1 (Ordering) and P5 (Nondegeneracy), and that $\succ$ is probabilistically sophisticated with respect to π. Then,*

(i) $\succ$ *satisfies* P3^Q *if and only if* $\succ_P$ *exhibits* $\overline{P3^Q}$*;*

(ii) $\succ$ *satisfies* P3^Q and $P6^Q$ *if and only if* $\succ_P$ *exhibits* $\overline{P3^Q}$ *and* $\overline{P6^Q}$.

Proof. Assume P1 (Ordering) and P5 (Nondegeneracy), and that $\succ$ is probabilistically sophisticated with respect to π.

(i) (P3^Q $\Rightarrow$ $\overline{P3^Q}$) Pick an act $f \in \mathcal{F}$, let E be its pivotal event and let $f(E) = x$ for some $x \in f(\mathcal{S})$. Assume that or any outcomes $x, y,$ and subacts $g_{x+},\ g_{x-},\ g_{y+},\ g_{y-}$:

$$\begin{bmatrix} g_{x+} \text{ if } s \in E_{fx+} \\ x \text{ if } s \in E \\ g_{x-} \text{ if } s \in E_{fx-} \end{bmatrix} \succsim \begin{bmatrix} g_{y+} \text{ if } s \in E_{fx+} \\ y \text{ if } s \in E \\ g_{y-} \text{ if } s \in E_{fx-} \end{bmatrix} \Leftrightarrow x \succsim y. \tag{103}$$

From the proof of Theorem 1, a stronger version of probabilistic sophistication (defined in Section 4.1) is satisfied for extreme as well not preferences. By probabilistic sophistication,

$$x \succsim y \Leftrightarrow \delta_x \succsim_P \delta_y \tag{104}$$

and

$$f = \begin{bmatrix} g_{x+} \text{ if } s \in E_{fx+} \\ x \text{ if } s \in E \\ g_{x-} \text{ if } s \in E_{fx-} \end{bmatrix} \succsim \begin{bmatrix} g_{y+} \text{ if } s \in E_{fx+} \\ y \text{ if } s \in E \\ g_{y-} \text{ if } s \in E_{fx-} \end{bmatrix} = f' \Leftrightarrow \tag{105}$$

$$\Leftrightarrow \quad \begin{array}{ccc} \pi \circ f^{-1} & \succsim_P & \pi \circ f'^{-1} \\ \parallel & & \parallel \\ \pi(E_{fx-})G_{x-} + \pi(E)\delta_x + \pi(E_{fx+})G_{x+} & & \pi(E_{fx-})G_{y-} + \pi(E)\delta_y + \pi(E_{fx+})G_{y+} \end{array}$$

Since the pivotal event E is non-null, $\pi(E) > 0$. Setting $\pi(E) = \lambda$ and $\pi(E_{fx-}) = \gamma(1-\lambda)$ completes the proof.

($\overline{P3^Q}$ $\Rightarrow$P3^Q) Take a simple lottery $P = [G_{x-}, P_{x-}, ..., x, p_x, G_{x+}, P_{x+}]$. If $\succ$ is non-extreme, by convex-rangedness of π, there exist events E_{fx-}, E and E_{fx+}, such that $\pi(E_{fx-}) = P_{x-}$, $\pi(E) = p_x$ and $\pi(E_{fx+}) = P_{x+}$. The assertion follows from (104) and (105). For extreme $\succ$, the proof is immediate.

(ii) For the non-extreme $\succ$, the argument mimics the proof of Theorem 2 (Step 2); for extreme $\succ$, the proof is straightforward and is, therefore, omitted. ■

Proposition 2 *Assume* $\overline{P1}$ *(Ordering),* $\overline{P5}$ *(Nondegeneracy), and* $\overline{P6^Q}$ *(Left-Continuity) hold. The following axioms are equivalent for a binary relation on the set of lotteries* $P_0(X)$,

$\succ_P$:

(1) $\overline{P3^Q}$ *(Rankwise Monotonicity);*

(2) Weak Monotonicity with respect to FOSD and Ordinal Invariance.

Proof. ((2)⇒(1)) Using Debreu's (1954) theorem to derive utility representation for outcomes u, as in the proof of Theorem 2, we can assume without loss of generality that $\mathcal{X} = \mathbb{R}$. Then monotonicity with respect to FOSD implies the following: Letting $\mathbf{x}_R$ be the ranked outcome vector of a lottery $R \in \mathcal{P}_0(\mathcal{X})$, we have that for all $P, Q \in \mathcal{P}_0(\mathcal{X})$ if $\mathbf{x}_P \geq \mathbf{x}_Q$, then $\mathbf{x}_P \succsim_P \mathbf{x}_Q$. Then, the proof of the assertion is an application of the result in Gevers (1979), with the anonymity condition implied by monotonicity with respect to FOSD.

((1)⇒(2)) Rankwise monotonicity implies that for each distribution P there is an outcome x, such that for any Q_{x-} and Q_{x+}:

$$P \sim (Q_{x-}, P_{x-},;\ x, p_x;\ Q_{x+}, P_{x+}), \tag{106}$$

in particular, $P \sim x$. Hence, it is straightforward that Rankwise Monotonicity implies monotonicity with respect to FOSD and Ordinal Invariance. ■

References

[1] Arrow K. J. and L. Hurwicz L. (1972): "An Optimality Criterion for Decision-Making under Ignorance," in: C. F. Carter, J. L. Ford, eds., *Uncertainty and Expectations in Economics*, Fairfield, NJ: Augustus M. Kelley.

[2] Barbera, S. and M. Jackson (1988): "Maximin, Leximin and the Projective Criterion: Characterizations and Comparisons," *Journal of Economic Theory*, 46: 34-44.

[3] Bhaskara Rao, K. P. S. and M. Bhaskara Rao (1983): *Theory of Charges. A Study of Finitely Additive Measures*, London: Academic Press.

[4] Bhattacharya, D. (2007): "Inferring Optimal Resource Allocation from Experimental Data," Dartmouth College.

[5] Boutilier, C. (1994): "Towards a Logic for Qualitative Decision Theory", in: *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR'94, Bonn, Germany, May 24-27)*: 75-86.

[6] Börgers, T. (1993): "Pure Strategy Dominance," *Econometrica*, 61, 2: 423-30.

[7] Buchinsky, M. (1995): "Recent Advances in Quantile Regression Models," *Journal of Human Resources*, 33, 1: 88-126.

[8] Chakraborty, A. and R. Harbaugh (2007): "Comparative Cheap Talk," *Journal of Economic Theory*, 132: 70-94.

[9] Chakraborty, A., N. Gupta and R. Harbaugh (2002): "Ordinal Cheap Talk in Common Value Auctions," working paper.

[10] Chamberlain, G. (1994): "Quantile Regression, Censoring and the Structure of Wages," in: *Advances in Econometrics*, Christopher Sims, ed., New York: Elsevier: 179-209.

[11] Chambers, C. P. (2007): "Ordinal Aggregation and Quantiles," forthcoming in *Journal of Economic Theory.*

[12] Chateauneuf, A., J. Eichberger and S. Grant (2006): "Choice under Uncertainty with the Best and Worst in Mind: Neo-additive Capacities," working paper.

[13] Chernozhukov, V., G. W. Imbens and W. K. Newey (2007): "Nonparametric Identification and Estimation of Nonseparable Models," *Journal of Econometrics*, 139: 4-14.

[14] Chesher, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71: 245-61.

[15] Chew, S. H. and J. S. Sagi (2006): "Event Exchangeability: Probabilistic Sophistication without Continuity or Monotonicity," *Econometrica*, 74, 3: 771-86.

[16] Cohen, M. (1992): "Security Level, Potential Level, Expected Utility: a Three-criteria Decision Model under Risk," *Theory and Decision*, 33: 101-34.

[17] Cohen, M. and J.-Y. Jaffray (1980): "Rational Behavior under Complete Ignorance," *Econometrica*, 48, 5: 1281-99.

[18] Debreu, G. (1954): "Representation of a Preference Ordering by a Numerical Function," in R. M. Thrall, C. H. Coombs, R. L. Davis, eds., *Decision Processes*, New York: Wiley.

[19] de Finetti, B. (1937): "Le Prévision: Ses Lois Logiques, ses Sources Subjectives," *Annales de l'Institute Henri Poincaré*, 7: 1-68. English translation: "Foresight: Its Logical Laws, Its Subjective Sources," in H. Kyburg and H. Smokler, eds., *Studies in Subjective Probability*, Huntington, New York: Krieger, 1980.

[20] Denneberg, D. (1994): *Non-Additive Measure and Integral*, Dordrecht, Boston: Kluwer.

[21] Dubois, D., H. Fargier, H. Prade and P. Perny (2002): "Qualitative Decision Theory: From Savage's Axioms to Nonmonotonic Reasoning," *Journal of the ACM*, 49, 4: 455-95.

[22] Dubois, D., M. M. Grabisch, F. Modave and H. Prade (2000): "Relating Decision under Uncertainty and Multicriteria Decision Making Models," *International Journal of Intelligent Systems*, 15, 10: 967-79.

[23] Fishburn, P.C. (1970): *Utility Theory for Decision Making*, New York: Wiley, 1970, reprinted in Huntington, NY: Robert E. Kreiger Publishing Company.

[24] Gevers, L. (1979): "On Interpersonal Comparability and Social Welfare Orderings," *Econometrica*, 47: 75-89.

[25] Gilboa, I. (1988): "A Combination of Expected Utility and Maxmin Decision Criteria," *Journal of Mathematical Psychology*, 32: 405-20.

[26] Gilboa, I. and D. Schmeidler (1989): "Maxmin Expected Utility with Non-unique Prior," *Journal of Mathematical Economics*, 18: 141-53.

[27] Grant, S. (1995): "Subjective Probability Without Monotonicity: Or How Machina's Mom May Also Be Probabilistically Sophisticated," *Econometrica*, 63, 1: 159-191.

[28] Grant, S., A. Kajii and B. Polak (1992): "Many Good Choice Axioms: When Can Many-Good Lotteries be Treated as Money Lotteries?," *Journal of Economic Theory*, 56: 313-337.

[29] Green, J. R. and B. Jullien (1988): "Ordinal Independence in Nonlinear Utility Theory," *Journal of Risk and Uncertainty*, 1: 355-387.

[30] Hansen, L. P. and T. J. Sargent (2004): *Robust Control and Economic Model Uncertainty*, Princeton, NJ: Princeton University Press.

[31] Hurwicz, L. (1951): "Optimality Criteria for Decision Making under Ignorance," Cowles Commission Discussion Paper, 370.

[32] Jaffray, J.-Y. (1988): "Choice under Risk and the Security Factor: An Axiomatic Model," *Theory and Decision*, 24: 169-200.

[33] Klibanoff, P., M. Marinacci and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73, 6: 1849-92.

[34] Koenker, R. (2005): *Quantile Regression*, Econometric Society Monograph Series, Cambridge University Press.

[35] Koenker, R. and G. Bassett (1978a): "Regression Quantiles," *Econometrica*, 46, 1: 33-50.

[36] Koenker, R. and G. Bassett (1978b): "The Asymptotic Distribution of the Least Absolute Error Estimator," *Journal of the American Statistical Association*, 73: 618-22.

[37] Koenker, R. and K. F. Hallock (2001): "Quantile Regression," *Journal of Economic Perspectives*, 15, 4: 143-56.

[38] Maccheroni, F., M. Marinacci and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 6: 1447-98.

[39] Machina, M. J. (1989): "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *Journal of Economic Literature*, 27, 4: 1622-68.

[40] Machina, M. and D. Schmeidler (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, 4: 745-80.

[41] Manski, C. F. (1988): "Ordinal Utility Models of Decision Making Under Uncertainty," *Theory and Decision*, 25, 1: 79-104.

[42] Maskin, E. (1979): "Decision-Making under Ignorance with Implications for Social Choice," *Theory and Decision*, 11: 319-37.

[43] Milnor, J. (1954): "Games against Nature," in: *Decision Processes*, R. M. Thrall, C. H. Coombs, and R. L. Davis, eds., Wiley: New York.

[44] Ramsey, F. P. (1931): "Truth and Probability," in: *Foundations of Mathematics and Other Logical Essays*, London: K. Paul, Trench, Trubner, and Co., reprinted in Ramsey (1978), Kyburg and Smokler (1980), and Gärdenfors and Sahlin (1988).

[45] Rawls, J. (1971): *A Theory of Justice*, Cambridge, MA: Harvard Press.

[46] Roberts, K. W. S. (1980): "Possibility Theorems with Interpersonally Comparable Welfare Levels," *Review of Economic Studies*, 47: 409-20.

[47] Roy, A. D. (1952): "Safety First and the Holding of Assets," *Econometrica*, 20, 3: 431-49.

[48] Rubinstein, A. (1996): "Why are Certain Properties of Binary Relations Relatively More Common in Natural Language?," *Econometrica*, 64: 343-56.

[49] Savage, L. J. (1954): *The Foundations of Statistics*, New York: Wiley, revised and enlarged, New York: Dover Publications, 1972.

[50] Schachter, B. (1997): "An Irreverent Guide to Value-at-Risk," *Financial Engineering News*, 1, 1, reprinted in: *Risks and Rewards*, 1998: 17-8.

[51] Segal, U. (1989): "Anticipated Utility Theory: A Measure Representation Approach," *Annals of Operations Research*, 19: 359-73.

[52] Segal, U. and J. Sobel (2002): "Min, Max, and Sum," *Journal of Economic Theory*, 106, 1: 126-50.

[53] Sims, C. A. (2001): "Pitfalls of a Minimax Approach to Model Uncertainty," Princeton University.

[54] Svensson L. E. O. (2007): "Robust Control Made Simple," Princeton University.

[55] Walsh, C. E. (2004): "Precautionary Policies," FRBSF Economic Letter, 2004-05.

[56] Yildiz, M. (2007): "Wishful Thinking in Strategic Environments," *Review of Economic Studies*, 74: 319-44.