CORRELATION MISPERCEPTION IN CHOICE

ANDREW ELLIS AND MICHELE PICCIONE

Abstract. We present a decision-theoretic analysis of an agent's understanding of the interdependencies in her choices. We provide the foundations for a simple and flexible model that allows the misperception of correlated risks. We introduce a framework in which the decision maker chooses a portfolio of assets among which she may misperceive the joint returns, and present simple axioms equivalent to a representation in which she attaches a probability to each possible joint distribution over returns and then maximizes subjective expected utility using her (possibly misspecified) beliefs.

"The debt collectors at Deutschebank sensed the bond traders at Morgan Stanley misunderstood their own trade. They weren't lying; they genuinely failed to understand the nature of the subprime CDO. The correlation among triple-B-rated subprime bonds was not 30 percent; it was 100 percent. When one collapsed, they all collapsed, because they were all driven by the same broader economic forces."

–Michael Lewis, *The Big Short*

1. Introduction

Many risky decisions, such as constructing a portfolio of securities, involve the interaction of many distinct variables, the correlation among which may be missed or misunderstood. The issue of misperception of correlated risks has recently acquired a renewed prominence, as authors such as Brunnermeier (2009), Coval et al. (2009), Hellwig (2009), and, more informally, Lewis (2010), have examined the significance of the inadequate understanding of correlations and the resulting mispricing of assets for the events surrounding the financial crisis in 2008. We present a decision-theoretic analysis of a decision maker's understanding of the interdependencies in her choices.

Date: October, 2016.

Department of Economics, London School of Economics and Political Science. Email: a.ellis@lse.ac.uk and m.piccione@lse.ac.uk. We would like to thank Larry Samuelson (the editor), three anonymous referees, David Ahn, Eddie Dekel, Philippe Jehiel, Bart Lipman, Larry Epstein, Erik Eyster and Rani Spiegler as well as participants at BU, BRIC 2016, CIREQ, Cowles 2015, ITAM, LBS, LSE, PSE, Queen Mary, SAET 2015, SOCRATES, Southampton, Tel Aviv and Texas for helpful comments and discussions.

Our aim is to provide the foundations of a simple and flexible model that represents misperception of correlated outcomes in a intuitive manner.[1]

Our analysis departs from the standard device of modeling the decision maker (DM) as choosing among individual, mutually exclusive alternatives. Instead, we propose a framework that explicitly considers her preferences over portfolios of assets, or more generally, profiles of actions. The DM's comparisons of portfolios allow us to define "misperception" of an uncertain environment in a very simple and straightforward way: misperception occurs when she is not indifferent between two different portfolios that "objectively" always result in the same outcomes. Unlike the standard model, our framework allows the DM to express a strict preference between owning a portfolio of the 500 underlying stocks of the S&P 500 (in the right proportions) and owning an S&P 500 index-tracking fund even with no transaction costs. We identify the DM's misperception of the correlation among the stocks of the S&P 500 when indifference between the portfolio and the tracking fund fails.[2]

We propose a model of decisions that maintains the classic assumptions of subjective expected utility but for those that conflict with the DM's possibly incorrect perception of correlation. Our main result describes the behavioral regularities necessary and sufficient for representing a DM who assigns probabilistic beliefs to correlations and maximizes expected utility. The next section provides a simple illustration of our setting as well as the new assumptions, and then describes our main conclusions.

1.1. **Illustration and overview.** The basic primitives of our model, formally described in Section 2, are a set of states, a set of assets, and a preference relation $\succsim$ over *portfolios of assets*. Each asset a returns a real number $a(\omega)$ in state ω. A portfolio is a finite collection of assets, denoted by $\langle a_1, a_2, ..., a_n \rangle$, that yields a payoff equal to the sum of the returns of all the underlying assets. The total return of a portfolio $\langle a_1, a_2, ..., a_n \rangle$ in state ω is $\sum_{i=1}^{n} a_i(\omega)$. The set of states thus describes the "objective" or "correct" structure of joint returns, i.e. the objectively possible returns of every profile. If a DM understands that returns have this structure, then her beliefs over the set of states fully describe her beliefs over portfolio returns.

The DM's preference over portfolios reveals when she does not understand the structure of returns; for instance, if she strictly prefers the index fund to its underlying stocks. More concretely, fix any two assets b and c and consider a third asset a

[1] Throughout, we use the more colloquial term correlation interchangeably with the more accurate "joint distribution" of returns or outcomes.

[2] We thank a referee for suggesting this example and another referee for pointing out related interpretive issues that we discuss in Section 4.

satisfying $a(\omega) = b(\omega) + c(\omega)$ for every state ω. The profiles $\langle b, c\rangle$ and $\langle a\rangle$ give the same return in every state, so one expects that a DM who understands the objective structure of returns would be indifferent between $\langle b, c\rangle$ and $\langle a\rangle$. In the language of decision theory, such indifference is implied by the standard *Monotonicity* axiom: state-by-state dominance by one portfolio implies a preference for it. However, if she misperceives the correlation between b and c, then she may have a strict preference for one or the other, e.g. $\langle b, c\rangle \succ \langle a\rangle$. To formalize our opening quote, if $\langle a\rangle$ is one triple-B-rated subprime bond and $\langle b, c\rangle$ is a portfolio with two halves of similar bonds in a synthetic CDO, then a DM may strictly prefer $\langle b, c\rangle$ to $\langle a\rangle$, believing it to be significantly less risky despite the default of these bonds being "driven by the same broader economic forces."

As a DM may violate Monotonicity for a number of reasons, we propose a novel "Weak Monotonicity" axiom that limits her violations to those directly attributable to misperception of correlation. Deferring a formal statement to Section 3, it roughly requires that whenever one portfolio always yields a better outcome than a second for *every conceivable correlation between* their returns, she prefers the first to the second. In the context of comparing $\langle b, c\rangle$ with $\langle a\rangle$, it requires that $\langle b, c\rangle \succsim \langle a\rangle$ whenever

$$\min_{\omega} b(\omega) + \min_{\omega} c(\omega) \geq \max_{\omega} a(\omega).$$

Thus, we can attribute a given violation of Monotonicity to the DM believing that a particular joint realization of returns, possibly inconsistent with the objective structure of returns, is sufficiently likely.

We show that a DM whose behavior satisfies Weak Monotonicity as well as the other axioms of expected utility *acts as if* she forms beliefs about the correlation between assets and then maximizes expected utility. We represent her choices by specifying beliefs on an enlarged state space, rich enough to express any perceived correlation between assets. For instance, our model explains the agent's preference for the synthetic CDO $\langle b, c\rangle$ over the bond $\langle a\rangle$ by assigning positive probability to b and c being less than perfectly correlated.

Our DM can also be represented *as if* she ranks portfolios according to a seemingly ad hoc yet intuitive procedure, which we call a *probabilistic correlation representation* or PCR. First, the assets are divided into *understanding classes*. The DM chooses *as if* she correctly understands the structure of returns for the assets within an understanding class. For instance, with assets a, b, and c as above, she is indifferent between $\langle b, c\rangle$ and $\langle a\rangle$ when all are in the same class, whereas a strict preference indicates that at least one of the assets belongs to a different class than the others. Second, she forms beliefs about the correlation between classes of assets. Formally, beliefs are defined on a multi-dimensional state space where each dimension corresponds to a class – if two assets belong to the same understanding class, then they depend on the same coordinate. Finally, the DM maximizes subjective expected utility with these beliefs. In addition to providing a parsimonious and intuitive model of decisions, the PCR permits a tight connection between parameters and behavior. With enough diversity in each class, we show the existence of a unique collection of understanding classes corresponding tightly to the DM's accuracy of perception of joint occurrences and, for generic risk preferences, uniqueness of beliefs.

A brief discussion of some simple yet significant implications of our model for asset pricing and, in particular, for the evaluation of complex securities with tranching such as CDO's concludes are analysis.

1.2. **Related Literature.** Misperception of correlations is a broad concept that has been studied in various guises, ranging from the limited understanding of patterns, as in Piccione and Rubinstein (2003), Eyster and Piccione (2013), and Levy and Razin (2015a), to the inability to derive some logical implications (Lipman, 1999). Evidence of misperception of correlations has been found in several experimental studies such as Eyster and Weizsäcker (2010), Enke and Zimmerman (2013), and Rubinstein and Salant (2015). In particular contexts, misperception of correlation has been shown to lead to a range of behaviors, including social influence (DeMarzo et al., 2003), overconfidence (Ortoleva and Snowberg, 2015), and polarization (Levy and Razin, 2015b). When applied to incomplete information games, our framework nests the behavior in solution concepts such as Cursed Equilibrium (Eyster and Rabin, 2005) and Analogy Based Expectations Equilibrium (Jehiel, 2005; Jehiel and Koessler, 2008). A key feature of our approach is that, being based on preferences, it is neutral with respect to the psychological biases and limitations that cause agents to perceive correlations incorrectly.

Framing can also be viewed as a proximate reason for misperception: different framings of the same action can make understanding correlations harder; see Example 1. More fittingly, different portfolios that yield the same outcomes can be viewed as different framings that affect the DM's choices. Choice theoretic works that highlight other aspects of framing include Salant and Rubinstein (2008) who study the conditions under which choice data can be rationalized as resulting from choice from a menu under different frames, and Ahn and Ergin (2010) who axiomatize a formal model where the framing of an act affects the probabilities used by the DM.

Our results also relate to a body of literature on boundedly rational choice theory. Lipman (1999) introduces a decision-theoretic model for relating an agent's logic to preferences. Al-Najjar et al. (2003) explicitly model the effects of complex environments on decision making as a preference for flexibility. Kochov (2015) develops a model of a DM with imperfect foresight, which can be interpreted as misperception of the auto-correlation between actions, where failure of Monotonicity also plays a role. Lastly, our representation can admit an interpretation in which different, endogenous sources of uncertainty (the understanding classes) determine beliefs, as in Chew and Sagi (2008) or Gul and Pesendorfer (2015).

Esponda (2008), Spiegler (2015), and Levy and Razin (2016) have developed approaches to misperception that are motivated by similar behavioral insights, but that do not in general admit PCR representations. In the first two, the DM forms her perception as a "fixed point", permitting the marginal distribution of a fixed payoff-relevant variable (such as the other player's strategy) to be affected by her own choice. In the last, the DM has ambiguity about the correlations and thus violates our Independence axiom.

2. Primitives

There is a set $\mathcal{A}$ of *actions*, with typical elements a, a_i, b, b_i.[3] Each action results in an *outcome* or consequence in $X = \mathbb{R}$, with typical elements x, y, z. This outcome is determined by a *state of the world* drawn from the finite set Ω.[4] We interpret the state space Ω as a description of the "objectively possible" joint realizations of the outcomes of any set of actions, against which the DM's subjective perceptions of joint realizations are evaluated.

A map $\rho : \mathcal{A} \times \Omega \to X$ describes the relationship between actions, states, and outcomes, with the action a yielding the outcome $\rho(a, \omega)$ in state ω. Slightly abusing notation, we write $a(\omega)$ for $\rho(a, \omega)$. Note that we allow for distinct actions a and b with $a(\omega) = b(\omega)$ for any $\omega \in \Omega$. Thus, each action implicitly includes a description that can affect how its relationship with other actions is understood; for instance, a and b could yield outcomes that depend on temperature, where a is described in Fahrenheit and b in Celsius, as in Example 1 below. We assume that $\mathcal{A}$ includes every constant action, i.e. for any $x \in X$ there is an action, also denoted by x, yielding the

[3]While we focus on the interpretation of actions as an active choice by the DM, actions can more generally be interpreted as different dimensions of an alternative affecting its outcome.

[4]We assume that $X = \mathbb{R}$ and that Ω is finite for ease of exposition. With minor adjustments, our results remain true for any Ω and many other outcome spaces on which we can define an appropriate addition operation, including $\mathbb{N}$, $\mathbb{R}_+$, and $\mathbb{R}^n$; see Ellis and Piccione (2016).

outcome x for every state in Ω. We write $\sigma(a)$ ($\sigma(a,b)$) for the coarsest σ-algebra by which a is (both a and b are) measurable.

From the set of actions, we derive a set $\mathcal{F}$ of *action profiles* (or profiles). Each element of $\mathcal{F}$ is a finite vector of actions for which the order does not matter – i.e., a multiset of actions. This allows the agent to take the same action multiple times, such as buying many shares of the same security while making irrelevant where an asset is listed in the description of a portfolio. A profile that consists of taking the n actions $a_1, ..., a_n$ is denoted $\langle a_1, ..., a_n\rangle$ or $\langle a_i\rangle_{i=1}^n$. To save notation, we do not write out the index when the number of actions is unimportant, i.e. we write $\langle a_i\rangle$ instead of $\langle a_i\rangle_{i=1}^n$. An agent who chooses the profile $\langle a_i\rangle_{i=1}^n$ receives the outcomes of all n actions $a_1, ..., a_n$, that is, she receives $\sum_{i=1}^n a_i(\omega)$ is state ω. We emphasize one key difference in interpretation from the menu-choice literature (Kreps, 1979; Dekel et al., 2001; Gul and Pesendorfer, 2001): the agent receives all of the actions in her chosen profile and does not make a second choice from the profile at a later point in time.

Our goal is to provide a simple and intuitive model of decisions among action profiles, starting from axioms on preferences. To this end, we adopt the widely used approach popularized by Anscombe and Aumann (1963) of expanding the set of alternatives to incorporate lotteries with objective probabilities. Specifically, the DM chooses among probability distributions over $\mathcal{F}$ having finite support, the set of which we denote by $\Delta\mathcal{F}$, rather than among profiles in $\mathcal{F}$. A typical element of $\Delta\mathcal{F}$ is $p = (p_1, \langle a_i^1\rangle; ...; p_n, \langle a_i^n\rangle)$, interpreted as the lottery where profile $\langle a_i^m\rangle$ occurs with probability p_m, and we write $p(\langle a_i\rangle) > 0$ for a profile $\langle a_i\rangle$ in the support of p. We discuss the reasons for including lotteries in our setup in Remark 1 and their role in the proofs in Remark 3, which can be skipped without any loss of continuity.

The DM chooses by maximizing a preference relation $\succsim$ over $\Delta\mathcal{F}$, with the symbol $\sim$ denoting indifference and $\succ$ strict preference. Naturally, the profile $\langle a_i\rangle_{i=1}^n$ in $\mathcal{F}$ corresponds to the lottery in $\Delta\mathcal{F}$ in which $\langle a_i\rangle_{i=1}^n$ has probability equal to one, and lotteries over X to lotteries over profiles containing a single, constant action. In this manner, $\Delta\mathcal{F}$ "contains" both all profiles and all lotteries over X.

Remark 1. As first shown by Anscombe and Aumann (1963), lotteries with objective probabilities facilitate the elicitation of subjective beliefs and utility by "convexifying" the choice domain, that is, by turning it into a mixture space. The most widely used procedure for generating a mixture space, introduced by Fishburn (1970), places additional structure on the outcome space to which actions map, usually lotteries over the original outcome space X; that is, horse race-roulette actions in which the roulette wheel is spun after the race ends. Thus, a mixture of actions becomes equivalent to an action that mixes among outcomes. In our paper, we adopt instead the reverse order while maintaining the original outcome space: the wheel is spun before running the race. As pointed out by Kreps (1988) and shown by Battigalli et al. (2013) in a setting with actions (but not profiles), the latter approach is equivalent to the former one when it includes an appropriate, yet less elegant, monotonicity axiom that explicitly rather than implicitly incorporates reduction of compound uncertainty.[5] In our setting, spinning the wheel first provides a natural way of mixing profiles that avoids thorny issues of interpretation. In particular, as the identities of the actions in a mixture are central to our model, establishing an exogenous equivalence between a mixture and another action imposes ad hoc restrictions on how the joint realizations of the actions involved are understood.

3. Foundations

We first introduce some standard assumptions. We then move to the key axioms of our approach.

3.1. **Standard Assumptions.**

Given two lotteries $p, q \in \Delta\mathcal{F}$, a mixture $\alpha p + (1-\alpha)q$, $\alpha \in [0,1]$, is the lottery in $\Delta\mathcal{F}$ in which the probability of each profile in the support of p and q is determined by compounding the probabilities in the obvious way.

Axiom 1 (Weak order)**.** The preference relation $\succsim$ is complete and transitive.

Axiom 2 (Continuity)**.** The sets $\{\alpha \in [0,1] : \alpha p + (1-\alpha)q \succsim r\}$ and $\{\alpha \in [0,1] : r \succsim \alpha p + (1-\alpha)q\}$ are closed for all $p, q, r \in \Delta\mathcal{F}$.

Axiom 3 (Independence)**.** For any $p, q, r \in \Delta\mathcal{F}$ and any $\alpha \in (0,1]$, $p \succsim q$ if and only if

$$\alpha p + (1-\alpha)r \succsim \alpha q + (1-\alpha)r.$$

These are classic axioms and their interpretation is standard. While one may plausibly argue that the DM's misperception does or should cause violations of some of them, we show that a meaningful and rich model of decisions can be constructed when they hold. In fact, most commonly used models featuring misperception do not violate any of them.

[5] We note that recent work incorporating ex-ante (as well as ex-post) lotteries and studying a DM who fails to reduce uncertainty has yielded interesting behavior in contexts such as uncertainty (e.g. Seo (2009)) and social preference (e.g. Saito (2013)).

3.2. **Weak Monotonicity.** In the standard approach, a profile $\langle a_i\rangle_{i=1}^n$ corresponds to an act $f:\Omega\to X$ yielding the consequence $f(\omega)=\sum_{i=1}^n a_i(\omega)$ in state ω. Whenever the DM reduces action profiles to acts, the map f is a sufficient description of the profile. In particular, if $\langle a_i\rangle_{i=1}^n$ and $\langle b_i\rangle_{i=1}^m$ correspond to the same act, then $\langle a_i\rangle_{i=1}^n \sim \langle b_i\rangle_{i=1}^m$. Within the expected utility framework, reduction to acts is implied by *Monotonicity*: if for any $\omega\in\Omega$

$$\sum_{i=1}^n a_i(\omega) \succsim \sum_{i=1}^m b_i(\omega),$$

then $\langle a_i\rangle_{i=1}^n \succsim \langle b_i\rangle_{i=1}^m$. We return to Monotonicity in Section 5.1; the following example illustrates the type of violations this paper is concerned with.

Example 1. A DM must choose between bets that depend on τ, tomorrow's high temperature. The DM can have either \$100 or the sum of the outcomes of bets b_C and b_F, where b_C pays \$100 if τ is less than 30 degrees Celsius (\$0 otherwise) and b_F pays \$100 if τ is at least 86 degrees Fahrenheit (\$0 otherwise). As 30° Celsius equals 86° Fahrenheit, a DM who knows this and easily converts Fahrenheit to Celsius expresses indifference between the sum of b_C and b_F and \$100 for sure. However, a DM who does not know how to convert from one unit to the other may not exhibit such indifference and reasonably prefer \$100 for sure to holding both b_C and b_F or vice versa.

A DM who expresses the preference $\langle 100\rangle \succ \langle b_C, b_F\rangle$ contradicts Monotonicity: both $\langle b_C, b_F\rangle$ and $\langle 100\rangle$ yield 100 in every state. Our novel axiom, *Weak Monotonicity*, relaxes this property by considering objectively impossible joint realizations of outcomes that would be possible with alternative joint distributions. To motivate it, consider why a DM might prefer $\langle 100\rangle$ to $\langle b_C, b_F\rangle$. She can "plausibly" conceive four possible joint realizations of $\langle b_C, b_F\rangle$: $(100,0)$, $(0,100)$, $(0,0)$, $(100,100)$. If the DM prefers 100 to $\langle b_C, b_F\rangle$, then she must think it is sufficiently likely that both bets return 0. Weak Monotonicity subsumes such considerations by strengthening the conditions under which a lottery dominates another. In particular, it requires that if the lottery over the outcomes generated by p is preferred to that generated by q for *every* "plausibly" conceived joint realizations of the outcomes of the actions in the supports, then $p \succsim q$.

Formally, for any finite subset of actions $\{c_1, ..., c_n\} = C \subset \mathcal{A}$, the set of all *plausible realizations* of C equals

$$range(c_1)\times range(c_2)\times ...\times range(c_n).$$

Thus, each plausible realization $\vec{x} = (x^{c_1}, x^{c_2}, ..., x^{c_n})$ is a vector of outcomes indexed by C such that each action c_i could, in isolation, result in x^{c_i}: for every $c_i \in C$, there exists $\omega \in \Omega$ so that $x^{c_i} = c_i(\omega)$. In Example 1, the plausible realizations of $\{b_C, b_F\}$ are $(100, 0)$, $(0, 100)$, $(100, 100)$, and $(0, 0)$, which we interpret below as $\langle b_C, b_F\rangle$ yielding four possible aggregate outcomes, namely, 100, 100, 200, or 0. Similarly, the profile $\langle b_C, b_F, 100\rangle$ could yield the aggregate outcomes 200, 200, 300, or 100. Naturally, the outcomes of $\langle b_C, b_F, 100\rangle$ dominate the outcomes of $\langle b_C, b_F\rangle$ regardless of any uncertainty about the conversion of temperature, and thus Weak Monotonicity will require that $\langle b_C, b_F, 100\rangle \succsim \langle b_C, b_F\rangle$.

A vector of outcomes $\vec{x}$ is a *plausible realization of lotteries p and q* if it is a plausible realization of the set of all the actions included in profiles that are assigned positive probability by either p or q. Fixing any such $\vec{x}$, each action a is assigned the outcome x^a and a profile $\langle a_i\rangle_{i=1}^n$ in the support of either p or q is assigned the aggregate outcome $\sum_{i=1}^n x^{a_i}$. Hence, for a plausible realization $\vec{x}$ of p and q, p *induces the lottery* $p_{\vec{x}}$ defined as

$$\left(p\Big(\langle a_i\rangle_{i=1}^n\Big), \langle\sum_{i=1}^n x^{a_i}\rangle\right)_{p(\langle a_i\rangle)>0}$$

in which the constant action yielding the outcome $\sum_{i=1}^n x^{a_i}$ has probability $p\left(\langle a_i\rangle_{i=1}^n\right)$, i.e. the probability of the profile to which the outcome is assigned. The lottery induced by q, denoted $q_{\vec{x}}$, is defined and interpreted similarly. Note that given a plausible realization $\vec{x}$ of p and q, if an action a occurs in both p and q, then its outcome is x^a in both induced lotteries.

For example, suppose that p randomizes equally between $\langle b_C, b_F\rangle$ and $\langle b_F\rangle$ and q selects $\langle b_F\rangle$ with certainty. For the plausible realizations $(100, 0)$ and $(100, 100)$, p induces a lottery that randomizes equally between two outcomes (100 and 0 in the first plausible realization, 200 and 100 in the second) while q induces a lottery that selects the worse of the two outcomes with certainty (0 and 100, respectively). Similarly, for the plausible realizations $(0, 100)$ and $(0, 0)$ p and q induce identical lotteries that yield with certainty 100 in the first plausible realization and 0 in the second.

As each realization corresponds to an outcome of the lottery under *some* joint distribution, these induced lotteries provide a natural way to compare p and q. When the lottery induced by p is preferred to that induced by q for each plausible realization, then p is better than q for any possible joint distribution. Weak Monotonicity relates this comparison to preference in the natural way.

Axiom 4 (Weak Monotonicity)**.** For any $p, q \in \Delta\mathcal{F}$, if $p_{\vec{x}} \succsim q_{\vec{x}}$ for every plausible realization $\vec{x}$ of p and q, then $p \succsim q$.

In words, if the DM prefers the lottery induced by p better than to that induced by q for any of their plausible realizations, then she prefers p to q. If she prefers larger payoffs, then it reduces to the following two implications when comparing $\langle a, b\rangle$ with $\langle c\rangle$. First, $\min a + \min b \geq \max c$ implies $\langle a, b\rangle \succsim \langle c\rangle$. Second, $\min c \geq \max a + \max b$ implies $\langle c\rangle \succsim \langle a, b\rangle$. Neither implication has bite when comparing $\langle b_C, b_F\rangle$ with $\langle 100\rangle$, so Weak Monotonicity does not restrict the ranking of these two profiles.

Remark 2. Weak Monotonicity has some formal similarities with the Dominance axiom of Seo (2009). Roughly, Seo's axiom states that for two compound lotteries over acts, if the compound lottery generated by the first is better than the second for any subjective probability distribution over states, then the first is preferred to the second. As in our axiom, it considers all possible beliefs, though over Ω rather than over plausible realizations. Seo's explicitly does not require reduction of compound lotteries, while ours explicitly assumes reduction. The Mixed Consequentialism axiom in Battigalli et al. (2013) is also related, but only considers the outcomes that are possible according to Ω.

3.3. **Understanding and Richness.** The above axioms suffice for our representation theorems but not for the identification of basic parameters of the model such as the DM's beliefs. To this end, our final assumption relies on identifying the DM's understanding of correlations among actions and ensures that there exist sufficiently diverse and understood sets of actions.

Our definition of understanding extends the logic of Weak Monotonicity. A DM who perceives the correlations within a subset of the actions correctly rules out some plausible realizations. Specifically, if she understands the correlations of the actions in a set C, then she should consider irrelevant any plausible realization of p and q that fails to "synchronize" the outcomes for the actions in C as for the joint occurrences that are determined by Ω. That is, she should only consider a plausible realization $\vec{x}$ if there exists $\omega \in \Omega$ such that $x^a = a(\omega)$ for all $a \in C$; we call any such plausible realization *C-synchronous*. For example, if $C = \{a, b\}$, $p = \langle a, b\rangle$ and $q = \langle c\rangle$, then a plausible realization (x^a, x^b, x^c) of p and q is C-synchronous if there exists $\omega \in \Omega$ so that $x^a = a(\omega)$ and $x^b = b(\omega)$. We say that the DM understands C if C-synchronous plausible realizations suffice to determine her preference.

Definition 1. The preference $\succsim$ *understands* $C \subseteq \mathcal{A}$ if for any $p, q \in \Delta\mathcal{F}$, $p \succsim q$ whenever $p_{\vec{x}} \succsim q_{\vec{x}}$ for all C-synchronous plausible realizations $\vec{x}$ of p and q.

In order to identify the main parameters of our representations, we assume that each action belongs to a suitably diverse, understood set of actions. A sufficient condition is that this understood set is rich, defined as follows.

Definition 2. A set $B \subset \mathcal{A}$ is *rich* if, for any $a, b \in B$ and any $\sigma(a, b)$-measurable function $f : \Omega \to X$, there exists $c \in B$ with $c(\omega) = f(\omega)$ for all ω.

Thus, a rich set containing bets on the events E and F contains all possible bets on E, F, $E \bigcup F$ and $E \bigcap F$. For two polar examples, a set containing an action for every possible mapping between states and outcomes is rich, but a singleton set is never rich. We can now state our assumption.

Assumption 1 (Non-Singularity)**.** Each $a \in \mathcal{A}$ belongs to a rich, understood subset of actions.

Non-Singularity is in the spirit of the Savage (1954) assumption that the domain of preference contains all possible acts. It is a joint assumption on both the preference $\succsim$ and the set $\mathcal{A}$.

4. Representation

A DM who violates Monotonicity acts as if she perceives uncertainty that is not entirely captured by the state space Ω. Under our axioms, we derive a state space sufficiently rich to express any additional uncertainty resulting from the misperception of the joint realizations of outcomes, and show that the DM acts as if maximizing expected utility on this enriched state space. While one can plausibly increase the dimensionality of uncertainty in many ways, we do so by considering multiple copies of the state space.

4.1. **Basic Correlation Representation.** Our first result shows that, under Axioms 1-4, one can obtain a basic representation in which each action is assigned its own copy. This representation is sufficiently flexible to encompass a wide variety of subjective perceptions of correlation. In particular, it can explain choices such as $100 \succ \langle b_C, b_F \rangle$ in Example 1 by assigning positive probability to b_C and b_F simultaneously returning 0.

To state our first representation formally, we introduce some notation. Let $\Omega^{\mathcal{A}} = \prod_{a \in \mathcal{A}} \Omega$ be the Cartesian product where one copy of Ω is assigned to each action in

$\mathcal{A}$, $\Sigma_{\mathcal{A}} = \otimes_{a\in\mathcal{A}}\sigma(a)$ be the product σ-algebra for $\Omega^{\mathcal{A}}$, Ω^a be the copy of Ω assigned to $a \in \mathcal{A}$, and for any $\vec{\omega} \in \Omega^{\mathcal{A}}$, ω^a be the component of $\vec{\omega}$ in Ω^a.

Definition 3. The preference relation $\succsim$ has a *basic correlation representation* if there exist a utility index $u : X \to \mathbb{R}$ and a probability measure π over $\Sigma_{\mathcal{A}}$ such that $p \succsim q$ if and only if $U(p) \geq U(q)$ where

$$U(p) = \sum_{p(\langle a_i \rangle)>0} p(\langle a_i\rangle_{i=1}^n)V(\langle a_i\rangle_{i=1}^n)$$

and

$$V(\langle a_i\rangle_{i=1}^n) = \int_{\Omega^{\mathcal{A}}} u\left(\sum_{i=1}^n a_i(\omega^{a_i})\right) d\pi(\vec{\omega}).$$

By increasing the dimension of uncertainty, the DM acts as if she is an expected utility maximizer on the larger state space $\Omega^{\mathcal{A}}$. Every $\vec{\omega} \in \Omega^{\mathcal{A}}$ determines a joint realization of the outcomes of the corresponding actions, so all additional uncertainty corresponds to the perception of correlations. In Example 1, the objective state space is the temperature τ, regardless of whether it is expressed in Celsius or Fahrenheit, but in a basic correlation representation, the DM's state space is instead a vector of temperatures, one for each action. The action b_C is assigned to one "copy" of the set of temperatures, and b_F to another. Then, the utility of the profile $\langle b_C, b_F\rangle$ is, setting $u(0) = 0$,

$$u(200)\,\pi\left(\tau^{b_C} < 30, \tau^{b_F} \geq 86\right) + u(100)\,\pi\left(\tau^{b_C} < 30, \tau^{b_F} < 86 \text{ or } \tau^{b_C} \geq 30, \tau^{b_F} \geq 86\right),$$

which can, of course, be different from $u(100)$. In particular, this representation allows the DM to attach positive probability to events such as "b_C yields 0 and b_F yields 0" that cannot occur if all uncertainty is captured by Ω.

The next theorem shows that, under our axioms, such representation always exists.

Theorem 1. *The preference $\succsim$ satisfies Weak Order, Continuity, Independence, and Weak Monotonicity if and only if $\succsim$ has a basic correlation representation. When there exist $p, q \in \mathcal{F}$ with $p \succ q$, u is unique up to a positive affine transformation.*

We now outline the proof. Observe that our axioms imply the standard expected utility axioms when restricted to lotteries over outcomes, so, for standard lotteries, the preference $\succsim$ has an expected utility representation with utility index u. Recall that a state $\vec{\omega} \in \Omega^{\mathcal{A}}$ assigns to each action a its "own" state $\omega^a \in \Omega$ and outcome $a(\omega^a)$. In the key step, we map each lottery p over profiles to a vector f_p where for each $\vec{\omega} \in \Omega^{\mathcal{A}}$, $f_p(\vec{\omega})$ equals the expected utility of lottery p according to u when action a returns the outcome $a(\omega^a)$. We use Weak Monotonicity to show that f_p is sufficient for the ranking of p: $f_p \geq f_q$ implies $p \succsim q$. Intuitively, if $f_p \geq f_q$, then for each plausible realization the expected utility of p is as large as the expected utility of q, and thus $p \succsim q$. This allows us to define a function F on these vectors such that $F(f_p) \geq F(f_q)$ if and only if $p \succsim q$ by setting $F(f_p)$ to be the utility of a lottery over X that exhibits indifference with p. Independence then yields that F is linear, so standard results ensure that F can be written as an integral with respect to a probability measure.

Remark 3. Adaptations of our axioms are necessary, but not sufficient, when we define $\succsim$ over $\mathcal{F}$ instead of $\Delta\mathcal{F}$. However Independence, which establishes a linear and thus separable representation, is not defined in the absence of lotteries. The formulation of alternative axioms needed to restore the lost properties remains an open question.

4.2. **Probabilistic Correlation Representation and Uniqueness.** Theorem 1 captures the minimal behavioral assumptions needed to represent the DM's perception of correlations. Our next step is to identify the DM's beliefs about the correlation as well as the coarsest space on which these beliefs can be expressed. The choice of $\Omega^{\mathcal{A}}$ for the basic correlation representation is far from parsimonious. To illustrate this point, consider Example 1 with an additional action b_C^o that pays \$100 if τ is at least 30 degrees Celsius (\$0 otherwise). Obviously, $\langle b_C, b_C^o \rangle$ is a constant act that pays \$100, and since both bets use Celsius temperatures, it is reasonable to presume that the DM perceives it so. Nevertheless, in our basic representation, the action b_C^o is also assigned its own copy of the set of temperatures.

For the remainder of the paper, we consider a *probabilistic correlation representation (PCR)* of the preference, that is equivalent to the basic representation but has a more frugal state space. In addition to being easy to apply and interpret in many situations, it allows us to provide a tighter characterization of the parameters of the model such as the DM's beliefs. A PCR envisages the DM as implementing following procedure. She first groups together certain actions, the correlations among which she understands as per Definition 1. We call such a set of actions an *understanding class* and the set of all understanding classes a *correlation cover*. She then forms beliefs within and across classes, which we model as a probability measure on a product state space indexed by the understanding classes. All the actions in the same class are assigned to the same copy of Ω, so the possible joint realizations within a class are identical to the objective ones. Consequently, any profile of the actions in the same class reduces to a single action. Of course, this procedure is only a representational device. The resulting classes are endogenously revealed from her choices; indeed, the DM is not necessarily aware which of the joint realizations are understood by her are "objective" and which are "perceived".

Formally, the representation of a PCR is as follows. A correlation cover $\mathcal{U}$ is a collection of subsets of $\mathcal{A}$ such that $\mathcal{U}$ covers $\mathcal{A}$ and no $C \in \mathcal{U}$ contains a distinct $C' \in \mathcal{U}$. Beliefs are defined on $\Omega^{\mathcal{U}} = \prod_{C \in \mathcal{U}} \Omega$, with the C-coordinate denoted by Ω^C, endowed with the product σ-algebra $\Sigma_{\mathcal{U}} = \otimes_{C \in \mathcal{U}} \Sigma_C$ where for each $C \in \mathcal{U}$, Σ_C is the coarsest σ-algebra by which every $a \in C$ is measurable. Given a state $\vec{\omega} \in \Omega^{\mathcal{U}}$, ω^C denotes the component of $\vec{\omega}$ assigned to C.

Definition 4. The preference $\succsim$ has a *probabilistic correlation representation* $(\mathcal{U}, \pi, u)$ for a correlation cover $\mathcal{U}$, a probability measure π over $\Sigma_{\mathcal{U}}$, and a utility index $u : X \to \mathbb{R}$ if it is represented by

$$U(p) = \sum_{p(\langle a_i \rangle) > 0} p(\langle a_i \rangle_{i=1}^n) V(\langle a_i \rangle_{i=1}^n)$$

where

$$V(\langle a_j \rangle_{j=1}^n) = \int_{\Omega^{\mathcal{U}}} u\left(\sum_{j=1}^n a_j(\omega^{C_j})\right) d\pi(\vec{\omega})$$

for any $\langle a_j \rangle_{j=1}^n$ and $C_1, ..., C_n \in \mathcal{U}$ with $a_j \in C_j$ for $j = 1, ..., n$.

While the PCR may seem a rather ad hoc procedure for implementing choice, it is easily seen to be equivalent to the basic correlation representation. Since the PCR places no restrictions on the understanding classes, any basic representation can be rewritten as a PCR, and vice versa.[6] To illustrate the PCR using Example 1, suppose that $\mathcal{U} = \{B_C, B_F\}$, with B_C and B_F interpreted as the bets evaluated in terms of Celsius and of Fahrenheit, respectively, so $b_C, b_C^o \in B_C$ and $b_F \in B_F$. Thus each $\vec{\tau} \in \Omega^{\{B_C, B_F\}}$ can be thought of as a pair of temperatures, one in Celsius and the other in Fahrenheit. As with the basic correlation representation, a DM for whom $100 \succ \langle b_C, b_F \rangle$ must attach positive probability to $\tau^{B_C} > 30°C$ and $\tau^{B_F} \leq 86°F$.

The notion of PCR is crucial for the unique identification of beliefs because many actions can be allocated to the same dimension. When understanding classes contain suitably diverse actions, one can tease out DM's perception of correlation; this is impossible in the basic representation because each action has its own dimension. Hence, we restrict attention to *rich* PCRs: $(\mathcal{U}, \pi, u)$ is rich if every $C \in \mathcal{U}$ is a rich set. Rich PCRs arise in many natural applications; see Section 4.3.

[6] The basic representation is a PCR where each understanding class is a singleton, and the converse follows from extending the probability measure from $\Omega^{\mathcal{U}}$ to $\Omega^{\mathcal{A}}$. See online appendix for details.

Rich PCRs also permit the characterization of the correlation covers that are consistent with the DM's preference. Such a characterization is critical for the study of comparative notions of understanding, such as an increase or a decrease in the misperception for a DM or heterogeneity in misperception. The basic representation provides one extreme in terms of dimensionality: every action has its own understanding class. A PCR allows us to determine the other extreme by asking if and when a coarsest grouping of actions into understanding classes exists. Such a coarsest correlation cover pins down rather unequivocally the DM's misperceptions of joint realizations.

Definition 5. A collection $\mathcal{U}$ of subsets of $\mathcal{A}$ is the *coarsest correlation cover* for the preference $\succsim$ if there is a rich PCR $(\mathcal{U}, \pi, u)$ of $\succsim$, and for any rich PCR $(\mathcal{U}', \pi', u')$ of $\succsim$ and any $B' \in \mathcal{U}'$, $B' \subseteq B$ for some $B \in \mathcal{U}$.

For any understanding class in a correlation cover $\mathcal{U}'$, there is a larger class in the coarsest cover. Therefore, any accurate perception captured by $\mathcal{U}'$ is also captured by the coarsest cover. Obviously, the coarsest correlation cover, should it exist, is unique.

We can now present our main identification result, which relies on the utility index u being continuous, as implied by a standard axiom (Grandmont, 1972), and Non-Singularity.

Theorem 2. *The preference $\succsim$ has a basic correlation representation with continuous u and Non-Singularity holds if and only if it has a rich PCR $(\mathcal{U}, \pi, u)$. Furthermore, there exists a coarsest correlation cover, and π is unique if u is not a polynomial.*

Proof. See Appendix A.2. □

Theorem 2 shows that when u is continuous, Non-Singularity and Axioms 1-4 hold if and only if a rich PCR exists, and that a rich PCR exists if and only if a coarsest correlation cover exists. Since the coarsest cover is unique, a minimal state space can be meaningfully defined and, unless the utility is polynomial, beliefs on this space are uniquely identified. When the utility is polynomial, uniqueness typically fails: for instance, a risk-neutral DM's beliefs about the expected return of each asset suffice to determine her behavior, so only her marginal beliefs are unique. In general, the curvature of u determines precisely which beliefs affect the agent's behavior, and for many standard utilities such as CARA or CRRA, beliefs are unique. In the online appendix, we give a behavioral characterization of a polynomial utility index and identify the set of all beliefs consistent with a given rich PCR.

We note here that a coarsest correlation cover need not be a partition. In fact, a coarsest correlation cover is a partition only if it has exactly one understand class and the DM understands the relation among all actions. To see why, observe that each understanding class can always be enlarged to include all constant actions; thus, if the coarsest correlation cover is a partition, it can only have one cell. More interestingly, consider once more Example 1. Suppose that the DM knows that $0°C = 32°F$ but is unsure about the scaling factor; that is, she understands the connection between any actions that only depend on whether or not the temperature is below freezing. Any such action can belong to both understanding classes, regardless of being expressed in Celsius or Fahrenheit.

We close this section by addressing some interpretive issues. Our framework implicitly assumes that the description of an asset is included in its formalization. In particular, a DM can perceive different framings of the same object as distinct assets – say two differently worded insurance contracts that are otherwise identical. Nevertheless, the descriptions of the individual assets must be sufficient for her perception of the returns of a profile. In a PCR, any additional information the DM acquires about the returns of a profile must be consistently incorporated into her beliefs about the joint returns of the underlying assets. For instance, let us return to a DM who strictly prefers an S&P 500 index-tracking fund to a portfolio of the stocks of the S&P 500. If we inform the DM that the portfolio and the index fund are actually identical, we would expect her to become indifferent. Our model can accommodate such indifference only by adjusting her beliefs about the correlations. Of course, this is a demanding requirement in the presence of misperception of the correlations among the underlying stocks.

The endogenous association of each action with a class and the unique determination of a coarsest set of classes are key advantages of our approach. One could, of course, begin by specifying a set of classes $\mathcal{U}$ and considering all acts on $\Omega^{\mathcal{U}}$. However, this would require setting the relevant dimensions of uncertainty and their association with alternatives exogenously. Furthermore, it would require that DM expresses preferences over some possibly nonsensical acts that are inconsistent with the primitive profiles.

4.3. **Applications of the PCR.** The following examples illustrates some natural applications of our representation.

4.3.1. *Framing.* Each action consists of a (Savage) act a and a frame $\mathbf{f} \in \mathbb{F}$, such as Celsius or Fahrenheit. The DM understands the connection between any acts framed in the same way. We can model this as a PCR where the correlation cover consists of the sets $B_{\mathbf{f}} = \{(a, \mathbf{f}) : a \in \mathcal{H}\}$, where $\mathcal{H}$ is the set of all acts.

4.3.2. *Imperfect inference in incomplete information games.* The PCR framework can be easily adapted to strategic environments (see Ellis and Piccione (2016)). Here, each action represents a behavioral strategy of a given player. This allows the formulation of a broad solution concept for players who misperceive the opponent's strategies in a systematic way and generalizes existing approaches such as Eyster and Rabin (2005) or Jehiel and Koessler (2008).

4.3.3. *Asset pricing.* An asset market consists of a set $\mathcal{A}^o$ of assets and the derivatives thereof. We model a derivative as a pair (γ, a^o), where γ is a function from X to itself and a^o is an asset in $\mathcal{A}^o$, that yields $\gamma(x)$ when a^o yields x. If the DM understands the set of all derivatives that depend on the same underlying asset, then she has a PCR when $\mathcal{A} = \mathcal{G} \times \mathcal{A}^o$, for $\mathcal{G}$ equal to all functions from X to X, and the correlation cover consists of the sets $B_{a^o} = \{(\gamma, a^o) : \gamma \in \mathcal{G}\}$.

4.3.4. *Source preference.* Each action is associated with a source S_i from a set $\mathcal{S}$. Each S_i is a sub-σ-algebra of Σ and corresponds to a set of actions B_i expressed in terms of the source. The correlation cover consists of all sets B_i, so the DM reduces any profile whose contents depend on the same source to one act but fails to do so when it depends on more than one source.

5. Discussion

We conclude by considering some special cases and by discussing some of the implications of our model in the context of portfolio choice.

5.1. **Special Cases.** In this subsection, we consider two special cases of particular interest. For simplicity of exposition, we maintain throughout that σ-algebra of each understanding class is the power set. It is easy to adapt Theorem 2 to show existence of such a representation by strengthening appropriately our definition of a rich set in the Non-Singularity assumption.

Weak Monotonicity may be too permissive in some circumstances. In particular, it allows the DM to perceive differently the distribution of two actions with identical mappings from Ω to X. This is undesirable if the DM evaluates the distribution of individual actions consistently and misperceives only their correlation. For instance, a ratings agency may accurately evaluate the chances of any given asset defaulting but misjudge the likelihood of joint defaults. The axiom below yields that beliefs over the outcomes of individual actions are consistent with the objective state space.

Axiom 5 (Simple Monotonicity)**.** If $a(\omega) = b(\omega)$ for all $\omega \in \Omega$, then $\langle a \rangle \sim \langle b \rangle$.

Simple Monotonicity implies that the DM's belief about the distribution of Ω does not depend on the action being evaluated. Formally, the preference $\succsim$ has a representation in the following class. Define π_C as the marginal over the copy of Ω assigned to class C.

Definition 6. A rich PCR $(\mathcal{U}, \pi, u)$ has *consistent marginal beliefs* if $\pi_{C_1}(\omega) = \pi_{C_2}(\omega)$ for all $C_1, C_2 \in \mathcal{U}$ and all $\omega \in \Omega$.

This specification obtains in most cited models of imperfect inference.

Proposition 1. *Let the preference $\succsim$ have a rich PCR $(\mathcal{U}, \pi, u)$ with non-constant u. Then, $\succsim$ satisfies Simple Monotonicity if and only if its PCR has consistent marginal beliefs.*

Proof. For sufficiency, suppose $\succsim$ has a rich PCR $(\mathcal{U}, \pi, u)$ and satisfies Simple Monotonicity. Pick $x, y \in X$ with $u(x) > u(y)$. By richness, for any $\omega \in \Omega$ and $C_1, C_2 \in \mathcal{U}$, there exists $a_i \in C_i$ that yields x at state ω and y otherwise for $i = 1, 2$. By Simple Monotonicity, $\langle a_1 \rangle \sim \langle a_2 \rangle$, so $[u(x) - u(y)]\pi_{C_1}(\omega) + u(y) = [u(x) - u(y)]\pi_{C_2}(\omega) + u(y)$ and thus $\pi_{C_1}(\omega) = \pi_{C_1}(\omega)$. Necessity is trivial. □

Unsurprisingly, our model collapses to the standard expected utility model if and only if the preference $\succsim$ satisfies the typical Monotonicity condition. For completeness, we state and prove this formally.

Axiom 6 (Monotonicity)**.** For any profiles $\langle a_i \rangle_{i=1}^n$ and $\langle b_i \rangle_{i=1}^m$,
if $\sum_{i=1}^n a_i(\omega) \succsim \sum_{i=1}^m b_i(\omega)$ for all $\omega \in \Omega$, then $\langle a_i \rangle_{i=1}^n \succsim \langle b_i \rangle_{i=1}^m$.

Proposition 2. *Suppose $\succsim$ has a rich PCR $(\mathcal{U}, \pi, u)$. The preference $\succsim$ satisfies Monotonicity if and only if its coarsest correlation cover equals $\{\mathcal{A}\}$.*

Proof. Suppose $\succsim$ has a rich PCR $(\mathcal{U}, \pi, u)$. Pick any $C \in \mathcal{U}$ and for any profile $\langle a_j \rangle_{j=1}^n$ choose $a \in C$ satisfying $a(\omega) = \sum_{j=1}^n a_i(\omega)$ for all ω. Then, $\langle a \rangle \sim \langle a_j \rangle_{j=1}^n$ by Monotonicity, which implies that

$$V(\langle a_j \rangle_{j=1}^n) = \int_\Omega u\left(\sum_{j=1}^n a_j(\omega)\right) d\pi_C(\omega).$$

Thus, $\succsim$ has a rich PCR $(\{\mathcal{A}\}, \pi_C, u)$, and $\{\mathcal{A}\}$ is the coarsest correlation cover. □

5.2. **Implications.** A DM with a fixed risk attitude simultaneously undervalues certain profiles while overvaluing others. For a very simple example, consider a strictly risk-averse trader who has a rich PCR with consistent marginal beliefs. Fix assets a, b, c so that $a(\omega) = b(\omega) = -c(\omega) = 1$ for all $\omega \in E$, $a(\omega) = b(\omega) = -c(\omega) = -1$ for all $\omega \notin E$, and suppose that $\langle b, c\rangle \sim 0$. This indifference reveals that b and c are in the same understanding class. It is easy to see that whenever this trader misperceives the joint realizations of $\langle a, b\rangle$, she overvalues $\langle a, b\rangle$ and undervalues $\langle a, c\rangle$ relative to a standard trader with the same beliefs: for $\langle a, b\rangle$ risk is smoothed out since a return equal to one is perceived as possible, whereas $\langle a, c\rangle$ is not perceived as riskless.

Independence requires that the DM is unsophisticated about her misperception. To illustrate this, consider the same trader and actions a, b, c as above, and suppose again that b and c are in the same understanding class while a is not. Thus, she misperceives the relationship between a and b as well as between a and c. A sophisticated trader may recognize her own lack of understanding. As in the ambiguity aversion literature, she may express the preference

$$\frac{1}{2}\langle a, b\rangle + \frac{1}{2}\langle a, c\rangle \succ \langle a, b\rangle \sim \langle a, c\rangle$$

because $\frac{1}{2}\langle a, b\rangle + \frac{1}{2}\langle a, c\rangle$ is "safer" than either alternative in that it offers a 50-50 lottery regardless of the correlation across classes; for instance, the DM studied by Levy and Razin (2016) would express such a preference. Obviously, the Independence axiom fails to hold. A full study of such behavior is left to future work.

A trader in our model typically perceives markets as incomplete, even when they are objectively complete. This can lead the agent to exhibit a flight to safety. Consider a simple asset market with one trader who divides fixed wealth w between three assets $\{a_1, a_2, a_S\}$, whose returns are governed by two states $S = \{s_1, s_2\}$ according to the return matrix

$$R = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

for columns a_1, a_2, a_S and rows s_1, s_2. This maps naturally into our formal setting, with a typical action corresponding to buying x shares of security a_i for $x \in \mathbb{R}$.

The above matrix has full rank, so markets are objectively complete. Indeed, a standard trader finds a_S to be redundant and perceives no arbitrage if and only if the price of a_S equals the sum of the prices of a_1 and a_2; she must express indifference between buying exactly x shares of a_S and buying x shares of a_1 together with x shares of a_2. Now assume the trader's preference has a PCR with consistent marginal beliefs and two understanding classes, one for buying shares of a_1 and the other for buying shares of a_2. The PCR trader *perceives* the return matrix to be

$$\tilde{R} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

in states (from top to bottom) (s_1, s_1), (s_1, s_2), (s_2, s_1), and (s_2, s_2). Our strictly-risk trader never purchases strictly positive quantities of both a_1 and a_2 if her belief has full support and the price of a_S equals the sum of the prices of a_1 and a_2.

Finally, we illustrate how incentives interact with correlation misperception by considering how tranching alters the perceived expected return of a CDO. If the CDO is untranched, then its expected return equals the sum of that of the underlying assets. Any two traders that agree on the expected value of each component asset also agree on the expected value of the untranched CDO, even if they disagree about the correlation between the assets. Their evaluations diverge, however, when a CDO is divided into tranches. Suppose, for instance, that returns are allocated to two tranches and that the senior tranche has a claim on the first y dollars of return while the junior receives the return of the CDO in excess of y. The expected returns of the junior and senior tranches are calculated using the utility indexes $u^J(x) = \max\{x - y, 0\}$ and $u^S(x) = \min\{x, y\}$. Since neither is a polynomial, Theorem 2 implies that all aspects of the joint distribution can in principle affect the valuation of the tranches: misperception of correlation can lead to an inaccurate assessment, even when each of the underlying assets is evaluated correctly.

To demonstrate interesting additional implications and perform relevant comparative statics, consider a trader with a PCR $(\{C_i\}_{i=1}^N, u, \pi^\chi)$ where π^χ satisfies

$$\pi^\chi(\omega^{C_1}, ..., \omega^{C_N}) = \chi q(\bigcap_{i=1}^N \{\omega^{C_i}\}) + (1 - \chi) \prod_{i=1}^N q(\{\omega^{C_i}\})$$

for some probability measure q over Ω. Such a probability measure has numerous interpretational advantages. For instance, the trader satisfies Monotonicity if and only if $\chi = 1$, the distance between π^1 and π^χ (as measured by Kullback-Leibler divergence of π^1 relative to π^χ) decreases with χ, and when actions a and b are perfectly correlated, the trader perceives their correlation coefficient to be χ. Also, this PCR has consistent marginal beliefs and q can be interpreted as the objective distribution on Ω.

Suppose the trader is risk-neutral and considers purchasing a CDO. To make the example stark, assume that the CDO is a profile $\langle a_1^n, ..., a_n^n \rangle$, where each a_i^n is a $\frac{1}{n}$

share of an asset a_i, $a_i^n \in C_i$ for each i and n, and a_i is "objectively" identical to a_j: each $a_i^n(\omega) = \frac{1}{n}a(\omega) \geq 0$ for all ω and some fixed $a \in \mathcal{A}$. As noted above, this trader correctly evaluates an untranched $\langle a_1^n, ..., a_n^n\rangle$ as exactly $E_q[a]$. However, suppose the CDO is split into a senior and a junior tranche with a return of y claimed by the senior tranche, with returns evaluated using u^J and u^S. It is easy to show that this trader undervalues the junior tranche and overvalues the senior tranche and that such misvaluations are monotonic in n and χ. Thus, the senior tranche can be sold at a profit, with the junior tranche kept on the books. Furthermore, an investment bank could "repackage" the junior tranche to create a CDO-squared, or a CDO made up of CDOs (Coval et al., 2009). This trader would again overvalue the senior tranche of the synthetic CDO. However, a second trader who understands the correlation correctly could make arbitrage profits by shorting the senior tranche and going long on the junior (even without repackaging). Indeed, Lewis (2010) reports the story of a Morgan Stanley trader adopting the opposite trade strategy and losing over \$9 billion.

Appendix A. Proofs not in main text

The following notation is used throughout. We denote the complement of a set E by $\overline{E}$. We sometimes denote an element of $\mathcal{F}$ by F.

A.1. **Proof of Theorem 1.** Necessity is trivial.

Assume for the remainder that $\succsim$ satisfies Weak Order, Continuity, Independence, and Weak Monotonicity. Herstein and Milnor (1953) implies that when restricted to the set of finite lotteries over X, $\succsim$ has an expected utility representation with utility index u normalized such that $u(0) = 0$. The key step is to show that we can map each lottery over action profiles into a (utility valued) act on the state space $\Omega^{\mathcal{A}}$. For any $p \in \Delta\mathcal{F}$, define the mapping $f_p : \Omega^{\mathcal{A}} \to \mathbb{R}$ by

$$f_p(\vec{\omega}) = \sum_{p(\langle a_i\rangle_{i=1}^n)>0} p(\langle a_i\rangle_{i=1}^n)u(\sum_{i=1}^n a_i(\omega^{a_i}))$$

for every $\vec{\omega} \in \Omega^{\mathcal{A}}$, where ω^{a_i} is the component of $\vec{\omega}$ corresponding to action a_i.

Lemma 1. *If $f_p \geq f_q$, then $p \succsim q$.*

Proof. Fix an arbitrary plausible realization $(x^a)_{a\in B}$ of $B = \mathcal{B}(p) \bigcup \mathcal{B}(q)$. By definition, there exists $\omega^a \in \Omega^a$ such that $x^a = a(\omega^a)$. Then note that

$$\left(p(\langle a_i\rangle_{i=1}^n), \sum_{i=1}^n x^{a_i}\right) \succsim \left(q(\langle b_i\rangle_{i=1}^m), \sum_{i=1}^m x^{b_i}\right)$$

if and only if

$$\left(p(\langle a_i\rangle_{i=1}^n), \sum_{i=1}^n a_i(\omega^{a_i})\right) \succsim \left(q(\langle b_i\rangle_{i=1}^m), \sum_{i=1}^m b_i(\omega^{b_i})\right)$$

if and only if

$$\sum_{p(\langle a_i\rangle)>0} p(\langle a_i\rangle)u(\sum_{i=1}^n a_i(\omega^{a_i})) \geq \sum_{q(\langle b_i\rangle)>0} q(\langle b_i\rangle)u(\sum_{i=1}^m b_i(\omega^{b_i}))$$

by the above. By $f_p \geq f_q$, the last inequality is true. Since (x^a) was chosen arbitrarily, p plausibly dominates q. By Weak Monotonicity, $p \succsim q$. □

Define $W = \{f_p : p \in \Delta(\mathcal{F})\}$, noting that W is convex. For ϕ in W, define $I(\phi) = \int u(x)dr$ for some $p \in \Delta(\mathcal{F})$ s.t. $f_p = \phi$ and a lottery r over X satisfying $r \sim p$. Such an r exists for every p by Weak Monotonicity, Completeness, and Continuity, so I is well-defined. Moreover, Independence and Weak Monotonicity imply that I is a positive linear functional, i.e. $x \geq 0 \implies I(x) \geq 0$. Obviously, $I(f_p) \geq I(f_q)$ if and only if $p \succsim q$.

Lemma 2. *I has a positive linear extension F to the smallest subspace W^* that contains W.*

Proof. Define $W^* = \{\lambda_1 x_1 - \lambda_2 x_2 : x_1, x_2 \in W, \lambda_1, \lambda_2 \in \mathbb{R}_+\}$. W^* is clearly a subspace and contains W. Let W' be any other subspace containing W. Pick any $y \in W^*$. Then $y = \lambda_1 x_1 - \lambda_2 x_2$, and since $x_1, x_2 \in W \subset W'$, $y \in W'$; hence $W^* \subseteq W'$. Suppose that $\lambda_1 x_1 - \lambda_2 x_2 = y$ and $y \in W$. Then

$$\frac{\lambda_1}{1+\lambda_1+\lambda_2}x_1 = \frac{\lambda_2}{1+\lambda_1+\lambda_2}x_2 + \frac{1}{1+\lambda_1+\lambda_2}y.$$

Since $x_1, x_2, y, 0 \in W$, so are the LHS and RHS above. Linearity of I on W gives that $I(y) = \lambda_1 I(x_1) - \lambda_2 I(x_2)$. So the function $F = y \mapsto \lambda_1 I(x_1) - \lambda_2 I(x_2)$ whenever $y = \lambda_1 x_1 - \lambda_2 x_2$ is well-defined and extends I. Linearity of F follows from linearity of I. To see that F is a positive linear functional, fix $\phi \in W^*$ with $\phi \geq 0$. Then $\phi = \lambda_1 x_1 - \lambda_2 x_2$; if $\lambda_1 = \lambda_2 = 0$, $\phi = 0$ so $F(\phi) = F(0) = 0$. Otherwise,

$$\frac{\lambda_1}{\lambda_1+\lambda_2}x_1 \geq \frac{\lambda_2}{\lambda_1+\lambda_2}x_2$$

and since both the LHS and RHS are in U, $F(\frac{\lambda_1}{\lambda_1+\lambda_2}x_1) \geq F(\frac{\lambda_2}{\lambda_1+\lambda_2}x_2)$. The remainder follows from linearity of F. □

For any $J \subseteq \mathcal{A}$, define $\Sigma_J = \otimes_{a\in J}\sigma(a)$, the product $\sigma-$algebra on Ω^J and $B_0(\Sigma_{\mathcal{A}})$ the set of simple $\Sigma_{\mathcal{A}}$-measurable functions. Note that the set W^* is a vector subspace of $B_0(\Sigma_{\mathcal{A}})$.

Lemma 3. *There is a positive linear extension $\hat{F}$ of F to all of $B_0(\Sigma_{\mathcal{A}})$ such that*

$$\hat{F}(\phi) = \int \phi d\pi^o,$$

for a finitely additive probability measure π^o on $\Sigma_{\mathcal{A}}$.

Proof. The function F is linear on W^* and $x \geq 0$ implies $F(x) \geq 0$. Pick any $y \in B_0(\Sigma_{\mathcal{A}})$. Since y is bounded, let z be an upper bound for y. z is a constant so $z \in W^*$. Hence W^* majorizes $B_0(\Sigma_{\mathcal{A}})$. By Theorem 8.32 of Aliprantis and Border (2006), F extends to a positive linear function on $B_0(\Sigma_{\mathcal{A}})$. By Theorem 14.4 of Aliprantis and Border (2006), there is a finitely additive signed measure of bounded variation, $\pi^o : \Sigma_{\mathcal{A}} \to R$, such that

$$\hat{F}(\phi) = \int \phi d\pi^o.$$

To see $\pi^o(\Omega^{\mathcal{A}}) = 1$, let ϕ be such that $\phi(\vec{\omega}) = 1$, $\hat{F}(\phi) = 1 = \pi^o(\Omega^{\mathcal{A}})1$. To see $\pi^o(E) \geq 0$ positive for any $E \in \Sigma_{\mathcal{A}}$, consider χ_E with $\chi_E(\omega) = 1$ for $\omega \in E$ and 0 otherwise, and since $\chi_E \geq 0$, $0 = \hat{F}(0) \leq \hat{F}(\chi_E) = \pi^o(E)$. Consequently, π^o is a finitely additive probability measure. □

To construct a countably additive probability, for every finite $J \subseteq \mathcal{A}$ define a set function π_J on (Ω^J, Σ_J) using the formula

$$\pi_J(E) = \pi^o(E \times \Omega^{\mathcal{A} \setminus J})$$

for every $E \in \Sigma_J$. Each π_J inherits finite additivity from π^o; in fact, since Σ_J has a finite number of members, π_J is countably additive and so a probability measure. By construction the family $\{\pi_J\}$ is Kolmogorov consistent. As a finite set, each Σ_J is a compact class, and trivially

$$\pi_J(E) = \sup\{\pi_J(E') : E \supseteq E' \in \Sigma_J\}.$$

By Kolmogorov's extension theorem (Theorem 15.26, Aliprantis and Border (2006)), there is a unique, countably additive $\pi : \Sigma_{\mathcal{A}} \to [0, 1]$ that extends each π_J. For any p, there is a finite $J_p \subset \mathcal{A}$ such that f_p is Σ_{J_p} measurable. Letting $\hat{f}_p$ be the natural projection of f_p onto Ω_{J_p},

$$\int_{\Omega^{\mathcal{A}}} f_p d\pi^o = \int_{\Omega_{J_p}} \hat{f}_p d\pi_{J_p} = \int_{\Omega^{\mathcal{A}}} f_p d\pi.$$

Therefore the function $U : \Delta\mathcal{F} \to \mathbb{R}$ defined by

$$U(p) = \int f_p d\pi$$

represents the DM's preference. To conclude, rewrite $U(p)$ as

$$\int_{\Omega^{\mathcal{A}}} f_p d\pi = \int_{\Omega^{\mathcal{A}}} \sum_{p(\langle a_i \rangle)>0} p(\langle a_i \rangle_{i=1}^n) u(\sum_{i=1}^n a_i(\omega^{a_i}))d\pi$$
$$= \sum_{p(\langle a_i \rangle)>0} p(\langle a_i \rangle_{i=1}^n) \int_{\Omega^{\mathcal{A}}} u(\sum_{i=1}^n a_i(\omega^{a_i}))d\pi,$$

the desired representation, which completes the proof.

A.2. **Proof of Theorem 2.** We split the proof of Theorem 2 into three propositions. Supposing that $\succsim$ has a basic correlation representation with continuous u, we apply Propositions 3 and 4 to yield that $\succsim$ has a rich PCR if and only if Non-Singularity holds. The existence of a unique coarsest correlation cover follows from Proposition 3. Uniqueness of beliefs is follows from Proposition 5.

We say that the preference $\succsim$ has a *rich and understood* correlation cover if there exists a correlation cover $\mathcal{U}$ such that every B in $\mathcal{U}$ is rich and understood.

Proposition 3. *The preference $\succsim$ satisfies Non-Singularity if and only if there exists a unique rich and understood correlation cover $\mathcal{U}$ such that any rich understood set C is contained in a set $B \in \mathcal{U}$.*

Proof of Proposition 3. Necessity is obvious. For sufficiency, let $\mathcal{U}$ be the set of all $\subseteq$-maximal, rich and understood subsets. We show this is non-empty via Zorn's Lemma. Fix any chain $\{B_t\}_{t\in T}$ of rich understood subsets of $\mathcal{A}$. We claim that $B^* = \bigcup_{t\in T} B_t$ is rich and understood and thus an upper bound by set inclusion. The set B^* is understood because for any $p, q \in \Delta\mathcal{F}$, the set of actions included both in B^* and in some profile in their support is finite. Thus, it is also contained in some understood B_t and therefore considering only B^*-synchronous plausible realizations suffice for preference between arbitrary p and q. Richness follows since, if a, b in B^*, $a, b \in B_t$ for some t so for any $f : \Omega \to X$ that is $\sigma(a, b)$-measurable, there exists $c \in B_t \subseteq B^*$ with $c(\omega) = f(\omega)$ for any $\omega \in \Omega$. By Zorn's lemma, there exists at least one maximal element. By Non-Singularity, each a belongs to at least one set $B \in \mathcal{U}$. The claim then follows from $\subseteq$-maximality. □

Proposition 4. *The preference $\succsim$ has a basic representation with continuous u and a rich and understood correlation cover $\mathcal{U}$ if and only it has a rich PCR $(\mathcal{U}, \pi, u)$ with continuous u.*

Proof of Proposition 4. Necessity is straightforward. For sufficiency, let $\mathcal{U}$ be a rich and understood correlation cover. Suppose that the preference relation $\succsim$ has a basic representation $(\{a\}_{a\in\mathcal{A}}, \pi_0, u)$ with continuous u. We normalize so $u(0) = 0$. We show there exists π so that $\succsim$ has rich PCR $(\mathcal{U}, \pi, u)$; the claim is trivial if u is constant, so that case is ignored.

Suppose first that u is linear. Then with no loss of generality,

$$V(\langle a_i\rangle_{i=1}^n) = \sum_{i=1}^n \int_{\Omega^{\mathcal{A}}} a_i(\omega) d\pi^i = \sum_{i=1}^n V(a_i)$$

where π^i is the marginal probability on Ω^a corresponding action a_i under π_0. Using standard arguments, it is easy to verify that $\succsim$ restricted to any $B \in \mathcal{U}$ has an affine representation V_B where $V_B(a_i) = \int a_i d\pi^B$ for some probability measure π^B. It also clearly has an affine representation $V(a_i)$. By uniqueness of affine representations, $V_B(a_i) = V(a_i)$. Picking any $C_1, ..., C_n \in \mathcal{U}$ s.t. $a_i \in C_i$,

$$V(\langle a_i\rangle_{i=1}^n) = \sum_{B\in\{C_1,...,C_n\}} \int \left(\sum_{C_i=B} a_i\right) d\pi^B.$$

Defining cross-class beliefs arbitrarily, for instance independently, delivers the result.

Suppose now that u is not linear, so there exist $x, y \in X$ such that

$$u(x + y) \neq u(x) + u(y).$$

The proof proceeds as follows. First, we prove Lemma 4 showing that π_0 assigns zero probability to any collection of "small-stakes bets" in the same understanding class yielding "misaligned" outcomes. Second, for any profile, we construct a profile of such bets indifferent to it. Finally, we use these bet profiles to apply the arguments of Theorem 1 with $\Omega^{\mathcal{U}}$ replacing $\Omega^{\mathcal{A}}$.

Throughout this case, we write Na for N copies of the action a, where N is a positive integer. For each $B \in \mathcal{U}$, let $\{E_B^k\}_{k=1}^{K_B}$ be the finest partition of Ω for which every action in B is measurable, and for $x \in X$ choose an action $\beta_x^{B,k} \in B$ so that $\beta_x^{B,k}(\omega)$ equals x if $\omega \in E^k$ and 0 otherwise and define the corresponding event

$$\mathcal{E}^{B,k,x} = \{\vec{\omega} \in \Omega^{\mathcal{A}} : \omega^{\beta_x^{B,k}} \in E_B^k\}.$$

Note such actions exist because B is rich. Let Θ_ε be an open interval of size ε around 0 that excludes 0. We first prove two preliminary lemmas.

Lemma 4. *If there exist* $x, y \in X$ *such that* $u(x+y) \neq u(x)+u(y)$*, then there exists* $\varepsilon > 0$ *such that for every* $B \in \mathcal{U}$*,*

$$\pi(\mathcal{E}^{B,i,x'} \cap \mathcal{E}^{B,j,y'}) = 0 \tag{1}$$

$$\pi_0\left(\mathcal{E}^{B,i,x'} \bigcap \mathcal{E}^{B,i,y'}\right) = \pi_0\left(\mathcal{E}^{B,i,y'}\right) = \pi_0(\mathcal{E}^{B,i,x'}) \tag{2}$$

$$\text{and } \sum_{k=1}^{K_B} \pi_0\left(\mathcal{E}^{B,k,x_k}\right) = 1 \tag{3}$$

for distinct $i, j \in \{1, ..., K_B\}$ *and* $x', y', x_1, ..., x_{K_B} \in \Theta_\varepsilon$*.*

In words, Equation (1) implies that the DM believes it impossible that bets on distinct E_i and E_j in the same class pay off jointly; Equation (2) implies that if one bet on E_i pays off, then all bets on E_i in the same class pay off; and Equation (3) implies that from a set of bets on all events in a partition and contained in the same class, at least one bet pays off. In sum, within the same understanding class, all the bets on one and only one of the elements of its finest partition pay off jointly.

Proof of Lemma 4. Pick $x', y' \in X$ and fix $B \in \mathcal{U}$. To save notation, we omit the dependence on B, i.e we write β_x^k instead of $\beta_x^{B,k}$, throughout the proof of this lemma. Since B is understood,

$$\frac{1}{2}\langle N\beta_{x'}^i, z_0\rangle + \frac{1}{2}\langle M\beta_{y'}^j, z_0\rangle \sim \frac{1}{2}\langle N\beta_{x'}^i, M\beta_{y'}^j, z_0\rangle + \frac{1}{2}z_0$$

for any positive integers N, M and $z_0 \in X$. Then, setting $\mathcal{E} = \mathcal{E}^{i,x'} \bigcap \mathcal{E}^{j,y'}$, we have

$$\begin{aligned}
&\pi_0(\mathcal{E}^{i,x'})[u(Nx'+z_0) - u(z_0)] + \pi_0(\mathcal{E}^{j,y'})[u(My'+z_0) - u(z_0)] + u(z_0) \\
=&[\pi_0(\mathcal{E}^{i,x'}) - \pi_0(\mathcal{E})][u(Nx'+z_0) - u(z_0)] + [\pi_0(\mathcal{E}^{j,y'}) - \pi_0(\mathcal{E})][u(My'+z_0) - u(z_0)] \\
&\quad + \pi_0(\mathcal{E})[u(Nx'+My'+z_0) - u(z_0)] + u(z_0)
\end{aligned}$$

and thus

$$\pi_0(\mathcal{E})\Big[u(Nx'+My'+z_0) + u(z_0) - u(Nx'+z_0) - u(My'+z_0)\Big] = 0. \tag{4}$$

If we can find non negative integers $N, M \in \mathbb{N}$ and $z_0 \in X$ such that the term in brackets is not-zero, then (1) must hold.

Without loss of generality, either $x, y > 0$, $x, y < 0$, or $x > 0 > y$ and either $x', y' > 0$, $x', y' < 0$, or $x' > 0 > y'$. We have four cases:

(1) If $x', y' > 0$ and $x, y > 0$; or $x', y' < 0$ and $x, y < 0$; or $x > 0 > y$ and $x' > 0 > y'$, then choose $z_0 = 0$.
(2) If $x', y' > 0$ and $x, y < 0$; or $x', y' > 0$ and $x, y < 0$, then choose $z_0 = x + y$.

(3) If $x', y' > 0$ and $x > 0 > y$; or $x, y > 0$ and $x' > 0 > y'$, then choose $z_0 = y$.
(4) If $x', y' < 0$ and $x > 0 > y$; or $x, y < 0$ and $x' > 0 > y'$, then choose $z_0 = x$.

Since u is continuous, it is easy to verify by applying the appropriate case for z_0 that there exist $\varepsilon > 0$ such that $x', y' \in \Theta_\varepsilon$ implies that the absolute value of the term in brackets in (4) is sufficiently close to $|u(x+y) - u(x) - u(y)| > 0$ for some positive integers N and M. Conclude (1) holds.

To see (2), fix non-zero $x', y' \in \Theta_\varepsilon$. Let $b \in B$ be a bet yielding x' on $\overline{E_i}$ and 0 otherwise and define

$$\mathcal{E}^b = \{\vec{\omega} \in \Omega^{\mathcal{A}} : \omega^b \in \overline{E^i}\}.$$

By arguments analogous to the above,

$$\pi_0\left(\mathcal{E}^{i,x'} \bigcap \mathcal{E}^b\right) = \pi_0\left(\mathcal{E}^{i,y'} \bigcap \mathcal{E}^b\right) = 0.$$

Since

$$(\frac{1}{2}, \langle N\beta^i_{x'}, z\rangle; \frac{1}{2}, \langle Nb, z\rangle) \sim (\frac{1}{2}, \langle Nx', z\rangle; \frac{1}{2}, \langle z\rangle)$$

by picking N such that $u(z + Nx') \neq u(z)$ for some $z \in \mathbb{R}$, we have

$$\left[\pi_0(\mathcal{E}^{i,x'}) + \pi_0(\mathcal{E}^b)\right]\left(u\left(Nx' + z\right) - u\left(z\right)\right) = u\left(Nx' + z\right) - u\left(z\right)$$

and thus

$$\pi_0(\mathcal{E}^{i,x'}) + \pi_0(\mathcal{E}^b) = 1.$$

Such N and z must exist by continuity as long as u is not constant by choosing a possibly smaller value for ε in Θ_ε. Plugging these into the inclusion-exclusion formula gives that

$$1 \geq \pi_0\left(\mathcal{E}^{i,x'} \bigcup \mathcal{E}^{i,y'} \bigcup \mathcal{E}_b\right) = 1 + \pi_0(\mathcal{E}^{i,y'}) - \pi_0\left(\mathcal{E}^{i,x'} \bigcap \mathcal{E}^{i,y'}\right)$$

and thus $\pi_0(\mathcal{E}^{i,y'}) = \pi_0\left(\mathcal{E}^{i,x'} \cap \mathcal{E}^{i,y'}\right)$. A symmetric argument with b' defined using y' instead of x' yields (2).

Consider any $x_1, ..., x_K \in \Theta_\varepsilon$, and choose $y \in \Theta_\varepsilon$ and positive integer N so that $u(Ny) \neq 0$. Since B is understood we have that

$$\left(\frac{1}{K}, \langle N\beta^k_y\rangle\right)_{k=1}^K \sim \left(\frac{1}{K}, \langle Ny\rangle; \frac{K-1}{K}, 0\right).$$

By (2) and the representation,

$$u\left(Ny\right) = \sum_{k=1}^{K} \pi_0\left(\mathcal{E}^{k,y}\right) u\left(Ny\right) = \sum_{k=1}^{K} \pi_0\left(\mathcal{E}^{k,x_k}\right) u\left(Ny\right)$$

which gives (3). □

Consider a lottery r. For each profile $F = \langle a_j \rangle_{j=1}^n$ in the support select a vector $C^F = \left(C_1^F, ..., C_n^F\right)$, which we call a *profile allocation*, such that $C_j^F \in \mathcal{U}$ and $a_j \in C_j^F$, $j = 1, ..., n$. Construct a map $\hat{f}_r : \Omega^{\mathcal{U}} \to \mathbb{R}$ such that for every $\vec{\omega} \in \Omega^{\mathcal{U}}$,

$$\hat{f}_r(\vec{\omega}) = \sum_{r(F)>0} r(F) u \left(\sum_{j=1}^n a_j(\omega^{C_j^F}) \right).$$

The vector of profile allocations $\left(C^F\right)_{r(F)>0}$ assigns each action in each profile that has a positive probability to an understanding class to which it belongs. Several allocations may be associated with the same r and thus several $\hat{f}_r$'s are generated for it; since the DM is indifferent between r and itself, to save notation, we omit the dependence on the profile allocations of the $\hat{f}_r$'s generated for the same lottery. The remainder of the proof follows from the arguments in Theorem 1 if we show that, for any such maps, if for some choice of profile allocations $\hat{f}_p \geq \hat{f}_q$ then $p \succsim q$.

Lemma 5. *Given any* $\varepsilon > 0$ *and profile* $F = \langle a_i \rangle_{i=1}^n$ *with allocation* $\left(C_1^F, ..., C_n^F\right)$, *there exist* $\beta_1, ..., \beta_T \in \mathcal{A}$, $B_1, ..., B_T \in \mathcal{U}$ *and* $N_1, ..., N_T \in \mathbb{N}_+$ *such that:*
(i) $\langle a_i \rangle_{i=1}^n \sim \langle N_j \beta_j \rangle_{j=1}^T$;
(ii) for any B_j, $j = 1, ..., T$, *there exists* $C_i^F = B_j$ *for some* $i = 1, ..., n$;
(iii) for any $j = 1, ..., T$, $\beta_j = \beta_x^{B_j,k} \in B_j$ *for some* $k \in \{1, ..., K_{B_j}\}$ *and* $x \in \Theta_\varepsilon$;
(iv) For any C_t^F, $t = 1, ..., n$, *and all* $\omega \in \Omega$,

$$\sum_{\{j: B_j = C_t^F\}} N_j \beta_j(\omega) = \sum_{\{i: C_i^F = C_t^F\}} a_i(\omega).$$

Proof. Statements (ii)-(iv) follow from the richness of each C_i^F. To see (i), note that because the preference $\succsim$ understands C_1^F and of statement (ii), DM is indifferent between F and the profile obtained by replacing the actions in the set $\{a_i : C_i^F = C_1^F\}$ with $\{N_j \beta_j\}_{\beta_j \in C_1^F}$. Statement (i) then follows from successive replacements as above and applying Weak Order, since each C_i^F is understood. □

Thus, the actions assigned to the understanding class B_i by the profile allocation can be replaced with N_j copies of each bet β_j in B_i while maintaining indifference.

Now, fix ε as per Lemma 4. Pick arbitrary p, q satisfying $\hat{f}_p \geq \hat{f}_q$ for the vectors of profile allocations $\left(C^F\right)_{p(F)>0}$ and $(C'^F)_{q(F)>0}$. Choose actions $\beta_1, ...\beta_T$, understanding classes $B_1, ..., B_T$, and positive integers $N_1, ..., N_T$ such that for a partition $\{J_F\}_{p(F)>0}$ of $\{1, ..., T\}$, $\{\beta_i : i \in J_F\}$, $\{B_i : i \in J_F\}$, and $\{N_i : i \in J_F\}$ are as in Lemma 5 for ε, F, and C^F. Similarly, choose actions $\beta_{T+1}, ...\beta_{T'}$, understanding classes $B_{T+1}, ..., B_{T'}$, and positive integers $N_{T+1}, ..., N_{T'}$ such that for a partition $\{J'_F\}_{q(F)>0}$ of $\{T+1, ..., T'\}$, $\{\beta_i : i \in J'_F\}$, $\{B_i : i \in J'_F\}$, and $\{N_i : i \in J'_F\}$ are as in Lemma 5 for ε, F, and C'^{F}. Replacing each profile in the support of p or q with corresponding profile of $N_j$copies of the bets β_j, yields lotteries p' and q'. Note $p' \sim p$ and $q' \sim q$ by Independence and Weak Order, so $p \succsim q$ if and only if $p' \succsim q'$.

Suppose without loss of generality that all the bets $\{\beta_j\}_{j=1}^{T'}$ are distinct and let $\mathcal{E}_j$ be the event $\mathcal{E}^{B,k,x}$ that corresponds to β_j. The maps $f_{p'}$ and $f_{q'}$ constructed in Theorem 1 are measurable by cylinders of the form

$$\mathcal{D} = \bigcap_{j=1}^{T'} L_j$$

where each L_j is either $\mathcal{E}_j$ or $\overline{\mathcal{E}_j}$. Fix any such $\mathcal{D}$ where $\pi_0(\mathcal{D}) > 0$ and $\vec{\omega} \in \mathcal{D}$. The difference in expected utility of lottery p' and q' at the state $\vec{\omega} \in \mathcal{D}$ is

$$\Gamma\left(\mathcal{D}\right) = \sum_{p(F)>0} p(F)u\left(\sum_{j\in J_F} N_j\beta_j\left(\omega^{\beta_j}\right)\right) - \sum_{q(F)>0} q(F)u\left(\sum_{j\in J'_F} N_j\beta_j\left(\omega^{\beta_j}\right)\right).$$

Since $\mathcal{D}$ has positive probability, by Lemma 4 there exists $\omega^B \in \Omega$ such that $\beta_j(\omega^B) = \beta_j(\omega^{\beta_j})$ for all j with $B_j = B$. Picking any $\vec{\tau} \in \Omega^{\mathcal{U}}$ such that $\tau^B = \omega^B$ whenever $B = B_j$ for some $j \in \{1, ..., T'\}$,

$$\Gamma(\mathcal{D}) = \hat{f}_p(\vec{\tau}) - \hat{f}_q(\vec{\tau}) \geq 0.$$

Since $\mathcal{D}$ was arbitrary, $U(p') - U(q') = \sum_{\mathcal{D}} \Gamma(\mathcal{D})\pi_0(\mathcal{D}) \geq 0$ and $p' \succsim q'$, which in turn implies $p \succsim q$. Repeating the remaining steps of Theorem 1 completes the sufficiency proof. □

Proposition 5. *If $\succsim$ has rich PCRs $(\mathcal{U}, \pi, u)$ and $(\mathcal{U}, \mu, u)$ where u is continuous but not a polynomial, then $\pi = \mu$.*

Proof. Suppose that $(\mathcal{U}, \pi, u)$ and $(\mathcal{U}, \mu, u)$ both represent the preference $\succsim$, and that u is continuous but not a polynomial. Let V_π and V_μ be the respective utility indexes. Say that an event $\mathcal{E} \in \Sigma_{\mathcal{U}}$ is a rectangle for $\{C_1, ..., C_n\}$ if there are $a_i \in C_i$ and $E_i \in \sigma\left(a_i\right)$ such that

$$\mathcal{E} \equiv \bigcap_{i=1}^{n} \{\vec{\omega} : \omega^{C_i} \in E_i\}.$$

The set of all rectangles is a π-system that generates the domain of π and μ, so if $\pi(\mathcal{E}) = \mu(\mathcal{E})$ whenever $\mathcal{E}$ is a rectangle, then $\pi = \mu$ by Caratheodory. We show this by induction, relying on the following lemma.

Lemma 6. *If $\mathcal{E}$ is a rectangle for $\{C_1, ..., C_n\}$, then*

$$\pi(\bigcap_{j\le n}\mathcal{E}_j \bigcap_{j\ge n+1}\overline{\mathcal{E}_j}) = \sum_{i=0}^{N-n-1}(-1)^i\pi(\bigcap_{j\le n+i}\mathcal{E}_j \bigcap_{j\ge n+2+i}\overline{\mathcal{E}_j}) + (-1)^{N-n}\pi(\bigcap_j \mathcal{E}_j).$$

where $\mathcal{E}_i = \{\vec{\omega} : \omega^{C_i} \in E_i\}$ are such that $\mathcal{E} = \bigcap_{i=1}^n \mathcal{E}_i$.

Proof. The claim follows by recursive substitutions, noting that

$$\pi(\bigcap_{j\le n+i}\mathcal{E}_j \bigcap_{j\ge n+1+i}\overline{\mathcal{E}_j})$$

equals

$$\pi(\bigcap_{j\le n+i}\mathcal{E}_j \bigcap_{j\ge n+2+i}\overline{\mathcal{E}_j}) - \pi(\bigcap_{j\le n+1+i}\mathcal{E}_j \bigcap_{j\ge n+2+i}\overline{\mathcal{E}_j})$$

and $\pi(\bigcap_{i\le N-1}\mathcal{E}_i \bigcap \overline{\mathcal{E}_N}) = \pi(\bigcap_{i\le N-1}\mathcal{E}_i) - \pi(\bigcap_i \mathcal{E}_i)$. □

We claim that if $\mathcal{E}$ is a rectangle for B, then $\pi(\mathcal{E}) = \mu(\mathcal{E})$. Proceed by induction on $\#B$. The case of $\#B = 1$ is standard since both PCRs are rich. Suppose that $\pi(\mathcal{E}') = \mu(\mathcal{E}')$ whenever $\mathcal{E}'$ is a rectangle for B with $\#B \le N-1$. Let $\mathcal{E}$ be an arbitrary rectangle for $\{C_1, ..., C_N\} \subseteq \mathcal{U}$, generated by $E_1, ..., E_N$ where $E_i \in \sigma(a_i')$ for some $a_i' \in C_i$.

Define the function

$$S_N(x_1, x_2, ..., x_N) = \sum_{Q\subseteq\{1,...,N\}}(-1)^{[N-\#Q]}u(\sum_{i\in Q}x_i). \tag{5}$$

If u is continuous, then Fréchet (1909) shows that $S_N(\vec{x}) = 0$ for all $\vec{x}$ if and only if u is a polynomial with degree less than N; see Almira and Lopez-Moreno (2007) for a proof. Thus, there exists $x_1, ..., x_N$ such that $S_N(x_1, x_2, ..., x_N) \neq 0$.

Consider the profile $\langle a_i\rangle_{i=1}^N$ where $a_i \in C_i$ and $a_i(\omega)$ equals x_i if $\omega \in E_i$ and equals 0 otherwise. Define

$$\mathcal{E}_i = \{\vec{\omega} : \omega^{C_i} \in E_i\}.$$

Note that

$$\begin{aligned}
V_\pi(\langle a_i\rangle_{i=1}^N) &= \sum_{Q\subseteq\{1,...,N\}}\pi(\bigcap_{i\in Q}\mathcal{E}_i \bigcap_{j\notin Q}\overline{\mathcal{E}_j})u(\sum_{i\in Q}x_i) \\
&= \sum_{Q\subseteq\{1,...,N\}}[\mathcal{K}(Q,N) + (-1)^{[N-\#Q]}\pi(\mathcal{E})]u(\sum_{i\in Q}x_i) \\
&= K + S_N(x_1, ..., x_n)\,\pi(\mathcal{E})
\end{aligned}$$

where $\mathcal{K}(Q,N)$ and K are weighted sums of rectangles for B''s with less than N members. Such a decomposition exists by Lemma 6. Since μ agrees with π on these rectangles,

$$V_\mu(\langle a_i\rangle_{i=1}^N) = K + S_N(x_1, ..., x_n)\mu(\mathcal{E}).$$

There exists a lottery q such $q \sim F$. Hence

$$V_\mu(\langle a_i\rangle_{i=1}^N) = \sum_{q(y)>0} q(y)u(y) = V_\pi(\langle a_i\rangle_{i=1}^N),$$

and since $S_N(x_1, ..., x_n) \neq 0$, $\mu(E) = \pi(E)$. □

Appendix B. Online Only Appendix

Proposition 6. *The preference $\succsim$ has basic correlation representation if and only if it has a PCR.*

Proof. It is easy to see that if $\succsim$ has a basic representation, it has a PCR with $\mathcal{U} = \{\{a\} : a \in \mathcal{A}\}$. Suppose $\succsim$ has a PCR $(\mathcal{U}, \pi, u)$. For every $a \in \mathcal{A}$, choose $C_a \in \mathcal{U}$ with $a \in C_a$. Pick any $B = \{a_1, ..., a_n\} \subset \mathcal{A}$. Define

$$\pi_B(\{\vec{\tau} \in \Omega^B : \tau_i \in E_i \ \forall i\}) = \pi(\{\vec{\omega} \in \Omega^{\mathcal{U}} : \omega^{C_{a_i}} \in E_i \forall \ i\})$$

where $E_i \in \sigma(a_i)$ for $i = 1, ..., n$. This π_B is clearly a measure defined on the π-system that generates $\otimes_{i=1}^n \sigma(a_i)$ and so can be uniquely extended to it. Moreover, the collection $\{\pi_B\}$ is Kolmogorov consistent and so by Kolmogorov's extension theorem, we can define π_0 on Σ_A to agree with every π_B. Thus $\succsim$ has a basic correlation representation with probability π_0 and utility u. □

For a PCR $(\mathcal{U}, \pi, u)$ and finite $B \subseteq \mathcal{U}$, let π_B denote the marginal distribution over the copies of Ω assigned to understanding classes in B. Note that the utility of any profile consisting of n actions is determined by some π_B with $\#B \leq n$.

Theorem 3. *If $\succsim$ has a rich PCR $(\mathcal{U}, \pi, u)$ and u is a polynomial of degree N, then it also has a PCR $(\mathcal{U}, \mu, u)$ if and only if $\mu_B = \pi_B$ for any $B \subseteq \mathcal{U}$ with $\#B \leq N$.*

Recall that $S_N(x_1, x_2, ..., x_N) = \sum_{Q \subseteq \{1,...,N\}} (-1)^{[N-\#Q]} u(\sum_{i \in Q} x_i)$. From our observation in the proof of Theorem 2, if u is continuous, then $S_N(x_1, x_2, ..., x_N) = 0$ for all $x_1, ..., x_N$ if and only if u is a polynomial of degree $N - 1$. Therefore, the result follows from the below Proposition.

Proposition 7. *If the preference $\succsim$ has a rich PCR $(\mathcal{U}, \pi, u)$, and*

$$N^* = \inf\{N : S_N(\vec{x}) = 0 \text{ for all } \vec{x}\},$$

then the PCR $(\mathcal{U}, \mu, u)$ also represents $\succsim$ if and only if $\mu(E) = \pi(E)$ for every rectangle for B with $\#B < N^*$.

From primitives, $S_N(x_1, ..., x_N) = 0$ for all $x_1, ..., x_n$ if and only if $p_N^E \sim p_N^O$ where

$$p_N^O = \left(2^{-(N-1)}, \sum_{x \in Q} x\right)_{\#Q \ odd} \text{ and } p_N^E = \left(2^{-(N-1)}, \sum_{x \in Q} x\right)_{\#Q \ even}$$

and Q ranges over all subsets (including $\emptyset$) of $\{x_1, ..., x_N\}$. When $x_i > 0$ for each i, a result in Eeckhoudt et al. (2009) implies p_N^O N-order stochastically dominates p_N^E.

Proof. Sufficiency follows from exactly the same arguments used in Thoerem 2. To see necessity, suppose that $S_N(\vec{x}) = 0$ for all $\vec{x}$ and that π agrees with μ on any rectangle for B when $\#B < N - 1$. Consider any profile $\langle a_i \rangle_{i=1}^m$, and assume WLOG that each a_i belongs to a distinct understanding class C_i; we show that

$$V_\pi(\langle a_i \rangle_{i=1}^m) = V_\mu(\langle a_i \rangle_{i=1}^m).$$

This is trivially true if $m < N$. The claim is proved if we show that, when $m \geq N$, we can replace each $V_\pi(\langle a_i \rangle_{i=1}^m)$ and $V_\mu(\langle a_i \rangle_{i=1}^m)$ with the (possibly negatively) weighted sum of the utilities of "sub-profiles" of $\langle a_i \rangle_{i=1}^m$ with at most $N - 1$ elements. Rearranging the equation $S_N(x_1, ..., x_N) = 0$,

$$u(\sum_{i=1}^N x_i) = - \sum_{Q \subseteq \{1,...,N\}, \#Q < N} (-1)^{[N - \#Q]} u(\sum_{i \in Q} x_i). \tag{6}$$

for any $x_1, ..., x_N$. Now,

$$V_\pi(\langle a_i \rangle_{i=1}^m) = \int u \left(\sum_{i=1}^m a_i(\omega^{C_i}) \right) d\pi,$$

so by (6) where $x_i = a_i(\omega^{C_i})$, $i = 1, ..., N - 1$, and $x_N = \sum_{i=N}^m a_i(\omega^{C_i})$, each term

$$u \left(\sum_{i=1}^m a_i(\omega^{C_i}) \right) = u \left(\sum_{i=1}^{N-1} a_i(\omega^{C_i}) + [\sum_{i=N}^m a_{C_i}(\omega^{C_i})] \right)$$

can be written as the sum of utilities where each argument contains the sum of at most $m - 1$ terms. We can repeat this procedure until the arguments of each $u(\cdot)$ contain the sum of at most $N - 1$ terms. Naturally, the exact same procedure can be applied to V_μ. This establishes the result. □

References

Ahn, David S. and Haluk Ergin (2010), "Framing contingencies." *Econometrica*, 78, pp. 655–695.

Al-Najjar, Nabil I., Ramon Casadesus-Masanell, and Emre Ozdenoren (2003), "Probabilistic representation of complexity." *Journal of Economic Theory*, 111, 49–87.

Aliprantis, Charalambos D. and Kim C. Border (2006), *Infinite dimensional analysis: A hitchhiker's guide (Third ed.)*. Springer, Berlin.

Almira, Jose Maria and Antonio Jesus Lopez-Moreno (2007), "On solutions of the fréchet functional equation." *Math. Anal. Appl.*, 332, 1119–1133.

Anscombe, F. J. and R. J. Aumann (1963), "A definition of subjective probability." *The Annals of Mathematical Statistics*, 34, pp. 199–205.

Battigalli, Pierpaolo, Simone Cerreia Vioglio, Fabio Maccheroni, and Massimo Marinacci (2013), "Mixed extensions of decision problems under uncertainty." *Mimeo*.

Brunnermeier, Markus K. (2009), "Deciphering the liquidity and credit crunch 2007-2008." *Journal of Economic Perspectives*, 23, 77–100.

Chew, Soo Hong and Jacob S. Sagi (2008), "Small worlds: Modeling attitudes toward sources of uncertainty." *Journal of Economic Theory*, 139, 1 – 24.

Coval, Joshua, Jakub Jurek, and Erik Stafford (2009), "The economics of structured finance." *Journal of Economic Perspectives-Volume*, 23, 3–25.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini (2001), "Representing preferences with a. unique subjective state space." *Econometrica*, 69, 891–934.

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003), "Persuasion bias, social influence, and unidimensional opinions." *Quarterly Journal of Economics*.

Eeckhoudt, Louis, Harris Schlesinger, and Ilia Tsetlin (2009), "Apportioning of risks via stochastic dominance." *Journal of Economic Theory*, 144, 994–1003.

Ellis, Andrew and Michele Piccione (2016), "Complexity, correlation and choice." *mimeo*.

Enke, B. and F. Zimmerman (2013), "Correlation neglect in belief formation." *Mimeo*.

Esponda, Ignacio (2008), "Behavioral equilibrium in economies with adverse selection." *The American Economic Review*, 98, pp. 1269–1291.

Eyster, Erik and Michele Piccione (2013), "An approach to asset pricing under incomplete and diverse perceptions." *Econometrica*, 81, pp. 1483–1506.

Eyster, Erik and Matthew Rabin (2005), "Cursed equilibrium." *Econometrica*, 73, pp. 1623–1672.

Eyster, Erik and Georg Weizsäcker (2010), "Correlation neglect in financial decision making." *Mimeo*.

Fishburn, Peter C. (1970), *Utility theory for decision making*. Wiley.

Fréchet, M. (1909), "Une définition fonctionnelle des polynomes." *Nouv Ann.*, 9, 145–162.

Grandmont, Jean-Michel (1972), "Continuity properties of von Neumann-Morgenstern utility." *Journal of Economic Theory*, 4, 45–57.

Gul, Faruk and Wolfgang Pesendorfer (2001), "Temptation and self-control." *Econometrica*, 69, 1403–1435.

Gul, Faruk and Wolfgang Pesendorfer (2015), "Hurwicz expected utility and subjective sources." *Journal of Economic Theory*, 159, Part A, 465 – 488.

Hellwig, Martin (2009), "Systemic risk in the financial sector: An analysis of the subprime-mortgage financial crisis." *De Economist*, 157, 129–208.

Herstein, I. N. and John Milnor (1953), "An axiomatic approach to measurable utility." *Econometrica*, 21, pp. 291–297.

Jehiel, Philippe (2005), "Analogy-based expectation equilibrium." *Journal of Economic Theory*, 123, 81 – 104.

Jehiel, Philippe and Frederic Koessler (2008), "Revisiting games of incomplete information with analogy-based expectations." *Games and Economic Behavior*, 62, 533–557.

Kochov, Asen (2015), "A behavioral definition of unforeseen contingencies." *Mimeo*.

Kreps, David M. (1979), "A representation theorem for "preference for flexibility"." *Econometrica*, 47, 565–577.

Kreps, David M. (1988), *The Foundations of Statistics*. Westview.

Levy, Gilat and Ronny Razin (2015a), "Correlation neglect, voting behaviour and information aggregation." *American Economic Review*, Forthcoming.

Levy, Gilat and Ronny Razin (2015b), "Does polarization of opinions lead to polarization of platforms? the case of correlation neglect." *Quarterly Journal of Political Science*, 10, 321–355.

Levy, Gilat and Ronny Razin (2016), "Correlation capacity, ambiguity and group shifts." *mimeo*.

Lewis, Michael (2010), *The Big Short*. Penguin.

Lipman, Barton L. (1999), "Decision theory without logical omniscience: Toward an axiomatic framework for bounded rationality." *The Review of Economic Studies*, 66, pp. 339–361.

Ortoleva, Pietro and Erik Snowberg (2015), "Overconfidence in political behavior." *American Economic Review*, 105, 504–35.

Piccione, Michele and Ariel Rubinstein (2003), "Modeling the economic interaction of agents with diverse abilities to recognize equilibrium patterns." *Journal of the European Economic Association*, 1, 212–223.

Rubinstein, Ariel and Yuval Salant (2015), ""They do what I do": Positive correlation in ex-post beliefs." *Mimeo*.

Saito, Kota (2013), "Social preferences under risk: Equality of opportunity vs. equality of outcome." *American Economic Review*, 103, 3084–3101.

Salant, Yuval and Ariel Rubinstein (2008), "(a, f): Choice with frames." *The Review of Economic Studies*, 75, pp. 1287–1296.

Savage, Leonard J. (1954), *The Foundations of Statistics*. Dover.

Seo, Kyoungwon (2009), "Ambiguity and second-order belief." *Econometrica*, 77, pp. 1575–1605.

Spiegler, Rani (2015), "Bayesian networks and boundedly rational expectations." *Mimeo*.