Correlation Preference

Yi-Chun Chen[*] Soo Hong Chew[†] Xinhan Zhang[‡]

October 21 2024

Abstract

Correlatedness between lotteries arising from social situations is ubiquitous. We propose a general utility representation for binary choice between pairs of possibly correlated lotteries rather than adopt the common practice of treating them as being independent. We axiomatize two classes of preference corresponding respectively to the correlation counterparts of the *independence* axiom for EU and to two axioms in non-EU characterizing weighted utility: *betweenness* and *projective independence*. The resulting correlation expected utility and correlation weighted utility can be discriminated using recent evidence of correlation-sensitive Allais behavior which supports the latter model rather than the former.

[*]Department of Economics and Risk Management Institute, National University of Singapore.
[†]China Center for Behavioral Economics and Finance, Southwestern University of Finance and Economics, and Department of Economics, National University of Singapore.
[‡]IGIER, Università Bocconi.

"It seems that the essential point is, and this is of general bearing, that, if conceptually we imagine a choice being made between two alternatives, we cannot exclude any probability distribution over those two choices as a possible alternative. The precise shape of a formulation of rationality which takes the last point into account or the consequences of such a reformulation on the theory of choice in general or the theory of social choice in particular cannot be foreseen; ..." – *Arrow (1951, pp. 20)*

1 Introduction

The Condorcet (1785) paradox has inspired a voluminous follow-up literature in which intransitive choice has a central role. It has been most prominent in the study of social choice since Arrow (1951) who seeks to address the question of how to aggregate individual preferences into a social decision. He starts by outlining the fundamental conditions that a rational collective decision-making process should satisfy, including transitivity of individual preferences. However, he also acknowledges that "*we could as well build up our economic theory on other assumptions as to the structure of choice functions if the fact seemed to call for it*". Indeed, the recent social choice literature appears to be gravitating towards questioning the desirability of transitivity for members of a society in a preference aggregation setting (see, e.g., Brandl et al. (2016); Brandl and Brandt (2020)). By contrast, in game theory, participants' identities and associated transitive preferences, often assumed to be EU, are modeled explicitly. Aumann's (1974) correlated equilibrium enriches transitive players with the ability to perceive strategy recommendations potentially *correlated* with each other through opponents' strategies.[1]

Motivated by the Allais (1953) paradox which inspires much of the development of transitive non-EU models, Fishburn (1982) initiates the nontransitive direction with a binary choice model based on a *skew-symmetric bilinear* (SSB) kernel ψ defined on pairs of lotteries (p, q) such that p is preferred to q if $\psi(p, q) \geqslant 0$.[2] SSB utility can be viewed as an EU model applied to $\psi(x, y)$ interpreted as utility of receiving

[1] Aumann (1987) recasts correlated equilibrium as EU maximization in line with Savage (1954).

[2] Skew-symmetry means $\psi(p, q) = -\psi(q, p)$; bilinearity requires $\psi(\alpha p + \beta q, r) = \alpha\psi(p, r) + \beta\psi(q, r)$ and $\psi(r, \alpha p + \beta q) = \alpha\psi(r, p) + \beta\psi(r, q)$.

outcome x while foregoing another outcome y, i.e., $\psi(p,q) = \sum_{x,y\in X} p(x)q(y)\psi(x,y)$.

The SSB utility has been applied extensively to social choice theory where individual and collective preferences are not necessarily transitive.[3] Of particular interest is Brandl et al.'s (2016) demonstration that the only probabilistic social choice function satisfying two appealing properties – population-consistency and composition-consistency – is what Fishburn (1984) calls the *maximal lottery*, which is guaranteed to exist under SSB utility.[4] In another line of study, Brandl and Brandt (2020) restore Harsanyi's utilitarian aggregation under Arrovian axioms within a specific SSB preference subdomain, the *pairwise comparison* preference, where the SSB kernel $\psi(x,y)$ only takes one of the three values $\{0, 1, -1\}$.[5] Here, Brandl and Brandt draw attention to their nontransitive direction via the quote from Arrow which we have adopted as epigraph for our paper.

Inspired by the same quote, we pursue a complementary direction to broaden the class of admissible binary choice models towards what Arrow suggests, "..., *if conceptually we imagine a choice being made between two alternatives, we cannot exclude any probability distribution over those two choices as a possible alternative*". Fishburn (1988) poses a binary choice question between two even-chance bets—might the choice between these two bets depend on whether one or two coins are used (to revisit in Section 6). In the context of game theory, SSB utility resembles the payoff in a zero-sum game where a row player chooses a mixed strategy p over actions x and the column player chooses another mixed strategy q over actions y. SSB has recently motivated Lanzani (2022) to axiomatize a correlation-sensitive binary choice model encompassing regret theory (Bell, 1982; Loomes and Sugden, 1982) and salience theory (Bordalo et al., 2012) by expanding its domain from lottery pairs (p,q) taken from $\Delta X \times \Delta X$ to *joint densities* π taken from $\Delta(X \times X)$ where $\Delta(X)$ denotes the set of lotteries defined on outcome set X.[6]

Beginning with a general utility model for binary choice under risk which can accommodate correlation sensitivity without requiring transitivity nor complete-

[3]See, for instance, Kreweras (1965), Fishburn (1984), Laffond et al. (1993), Rivest and Shen (2010), Brandl et al. (2016), Brandl and Brandt (2020).

[4]Population-consistency regards consistency with respect to a variable electorate, while composition-consistency requires composition-consistency (Brandl et al., 2016). See their paper for the formal definitions and a discussion on the development of these axioms.

[5]p is pairwise-comparison preferred to q if, when a pair of alternatives (x,y) is randomly selected according to p and q respectively, p generates the winner with higher probability in expectation.

[6]Fishburn (1989) accomplishes this in a Savagean setting leading to a skew-symmetric additive utility. We will revisit this issue in Section 6.

ness, we extend Lanzani's (2022) model by introducing correlation sensitivity to Samuelson's (1952) four-lottery version of the independence axiom: when each of a pair of lotteries p and p' is preferred to a corresponding lottery in another pair, q and q', a probability mixture $\alpha p + (1-\alpha)p'$ between p and p' is also preferred to the corresponding probability mixture $\alpha q + (1-\alpha)q'$ between q and q' for $\alpha \in (0, 1)$. This yields a characterization, in Section 4, of *correlation expected utility* (CEU) in which the expectation of the utility $\phi(x, y)$ of receiving x while foregoing y is taken over the joint density π (see Table 3 in the next section for illustration). In other words, receiving the row marginal lottery is preferred to receiving the column marginal under their join density π if $\mathbb{E}_\pi \phi \geqslant 0$. Notably, CEU reduces to Lanzani's model when the bivariate utility ϕ is skew symmetric. Without skew symmetry, CEU includes additionally an extension of the *expectations-based reference dependent* (ERD) model to the setting of correlation preference.

The figure below depicting a pair of lotteries with the same three equally likely outcomes bears some resemblance to the Condorcet paradox.

A: 100	A: 50	A: 0
B: 50	B: 0	B: 100

Figure 1: Same-marginal lotteries with three outcomes.

A strict preference between the two lotteries reveals an instance of correlation sensitivity. Observe that under SSB or any transitive utility model, lotteries A and B, being viewed as "identical", would be treated the same. In a latent class analysis, Loewenfeld and Zheng (2023) find that most subjects' choices are consonant with correlation insensitivity while finding a category at 37.4% favoring A over B without producing a category for the opposite choice behavior (see subsection 3.1 for further details and analysis). While CEU can exhibit non-indifference between the two lotteries, its restriction to salience theory and regret theory (under standard assumptions) through skew symmetry implies a preference for B over A. Our more general CEU can accommodate both forms of correlation sensitivity given that it encompasses skew symmetry as well as asymmetry.

The Allais (1953) paradoxes, especially the common-consequence paradox (see description in subsection 3.2) have provided much of the impetus for the development of axiomatic non-EU models under transitivity in the two decades before the millennium.[7] At the same time, it has been reported that the rate of Allais behavior tends to be significantly lower when choices are presented in a state-act framework involving maximal correlations.[8] In this regard, the table below displays the results of two representative within-subject studies by Bruhin et al. (2022) followed by Loewenfeld and Zheng (2024).[9] Together displayed there are the results from a more complete within-subject experiment by Frydman and Mormann (2018) who report a change in subjects' propensity to exhibit Allais behavior when the corresponding lottery pairs are correlated at different levels depending on a single probability parameter. Subjects are asked to make binary choices from two pairs of lotteries $L_1^z = (\$25, 33\%;\ z, 66\%;\ 0, 1\%)$ versus $L_2^z = (\$24, 34\%;\ z, 66\%)$, $z \in \{0, \$24\}$, under zero, intermediate, and maximal correlation (see details relating to Table 6 in subsection 3.2). We refer to this dependence of Allais behavior on the level of correlation as the *extended* Allais paradox.

Degree of correlation	Independent	Intermediate	Maximal
Frydman and Mormann (2018)	48%	36%	15%
Bruhin et al. (2022)	48%	-	20%
Loewenfeld and Zheng (2024)	62%	-	18%

Table 1: Extended Allais paradox

Observe that neither SSB nor transitive non-EU can account for the *difference* in rates under different correlations for the same pair of marginals in the extended Allais paradox. Despite its greater flexibility, CEU also cannot account for the extended Allais paradox given that it satisfies Savage's Postulate 2 (P2) in the interdependent state-by-state setting as described by Savage (1954, p.103; see discussion in subsection 3.2). To accommodate the extended Allais paradox, we formulate a *correlation betweenness* axiom in Section 5 to axiomatize a non-CEU preference, called *correlation betweenness utility* (CBU), which reduces to betweenness utility (Dekel, 1986; Chew, 1989). In conjunction with an alternative weakening of correlation independence, called *correlation projective independence*, CBU specializes further to a *correlation weighted utility* (CWU) representation, $\mathbb{E}_\pi \phi + \psi(p, q)$, as

[7] E.g., rank-dependent utility (Quiggin, 1982), weighted utility (Chew, 1983), (cumulative) prospective theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), quadratic utility (Chew et al., 1991), etc.

[8] See Table 6 for an example involving different levels of correlation.

[9] The literature documents a wide range of proportions of Allais behavior with reported proportions ranging from 33-36% in Starmer (1992), 20% in Bordalo et al. (2012), 14-36% in Ostermair (2022), 40% in Humphrey and Kruse (2023).

the sum of a correlation-insensitive SSB kernel ψ defined on the marginals p and q of π and a correlation-sensitive CEU kernel defined directly on π.

<table>
<tr><th colspan="2" rowspan="2">Representing $p \geq^{\pi} q$</th><th colspan="2">Correlation Sensitivity</th></tr>
<tr><th>Insensitive</th><th>Sensitive</th></tr>
<tr><td rowspan="2">Linearity of representation in (joint) probability</td><td>Linear</td><td>EU
$\mathbb{E}_p u - \mathbb{E}_q u$</td><td>CEU
$\mathbb{E}_\pi \phi$</td></tr>
<tr><td>Bilinear</td><td>SSB/WU
$\psi(p,q)$</td><td>CWU
$\mathbb{E}_\pi \phi + \psi(p,q)$</td></tr>
</table>

Table 2: Correlation sensitivity and linearity of representation.

Table 2 above summarizes how different correlation generalizations of EU, previewed in the next section, can account for the experimental evidence relating to same-marginal lotteries (SML) and Allais behavior in the literature. Under correlation insensitivity, while SSB and weighted utility (WU) can both exhibit Allais behavior, the rates of violations do not depend on the degree of correlation. Under correlation sensitivity, CEU can exhibit SML preference but it is incompatible with the incidence of Allais behavior in the maximally correlated common-consequence problem. CWU is the only model that can both exhibit SML preference and account for the extended Allais paradox. The specific arguments are discussed in Section 3.

In the penultimate Section 6, we incorporate correlation sensitivity to the two main directions of generalization of SEU: the nontransitive approach of Fishburn (1989) which retains Savage's Postulate 2 (P2) versus the transitive approach of *probabilistic sophistication* (PS) due to Machina and Schmeidler (1992) without P2. Through a weakening of transitivity by restricting it to correlation-insensitive comparisons only, SSB serves an intermediary role between transitive PS and our more general correlation-sensitive PS, which reduces to transitive PS once binary choice over Savagean acts is assumed to be transitive. We further extend correlation PS from the Savagean grand world to the small worlds setting of Chew and Sagi (2008) in order to model the potential dependence of correlation PS on the source of uncertainty. This approach also provides an efficient account for the Epstein and Halevy (2019) finding of incremental aversion to correlation ambiguity.

We conclude in Section 7 with discussions on applying our findings to the topics of inertia via correlation PS without completeness, followed by rank dependence based on Yaari's (1987) dual utility, before returning to our motivating topic of social choice.

2 Preview of Main Findings

In binary choice under risk, it is customary to denote a (nonstrict) preference for a lottery p over another lottery q using the notation $p \succeq q$. In this regard, the domain of choice is implicitly $\Delta(X) \times \Delta(X)$ where X refers to a real outcome set and the set of (finite support) lotteries defined on X is denoted by $\Delta(X)$. The preference $p \succeq q$ is often written equivalently as $(p, q) \in R$ where the subset $R \subset \Delta(X) \times \Delta(X)$, called *preference set*, comprises all ordered pairs of lotteries such that the first element is preferred to the second element.

To accommodate the possibility of binary preference being sensitive to potential correlation between lotteries, we would need to expand the domain from $\Delta(X) \times \Delta(X)$ to $\Delta(X \times X)$ comprising joint densities defined on $X \times X$. A joint density π illustrated in the table below can be written as $\sum_{i=1}^{m} \sum_{j=1}^{n} \pi(x_i, y_j)\delta(x_i, y_j)$ where $\delta(x, y)$ denotes the degenerate comparison of receiving x against receiving y for sure.

π	y_1	$\cdots$	y_n	
x_1	$\pi(x_1, y_1)$	$\cdots$	$\pi(x_1, y_n)$	$p(x_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_m	$\pi(x_m, y_1)$	$\cdots$	$\pi(x_m, y_n)$	$p(x_m)$
	$q(y_1)$	$\cdots$	$q(y_n)$	

Table 3: A joint density π with row marginal $p(x)$ and column marginal $q(y)$.

Denote by $p \succeq^{\pi} q$ the decision maker's (nonstrict) preference for "p over q *under joint density* π" with p and q being the row and column marginals of π. In line with the practice of identifying $\succeq$ with its preference set $R \subset \Delta(X) \times \Delta(X)$, it is convenient to refer to $\succeq^{\pi}$ equivalently in terms of the corresponding preference set $\Pi \subset \Delta(X \times X)$ comprising all joint densities π for which the row marginal p given by π_1 is weakly preferred to the corresponding column marginal q given by π_2. We adopt the convention $\pi \in \Pi$ if DM weakly prefers receiving its row marginal $\pi_1 = \sum_{i=1}^{m} p(x_i)\delta(x_i)$ to receiving its column marginal $\pi_2 = \sum_{j=1}^{n} q(y_j)\delta(y_j)$, and refer to $\{\succeq^{\pi}\}_{\pi \in \Pi}$ and Π interchangeably as DM's *correlation preference*.

Denote by π^{T} the transpose of π, $\hat{\Pi} = \{\pi : \pi \in \Pi,\ \pi^{\mathrm{T}} \notin \Pi\}$ as the *strict preference* set, $\tilde{\Pi} = \Pi \backslash \hat{\Pi}$ as the *indifference set*, and $\check{\Pi}$ as the complement of Π. Notice the distinct role of *indifference set* $\tilde{\Pi}$ from that of *indifference curves (lines)* in the two settings: under transitivity, indifference curves *partition* the entire domain, while in the correlation setting, points in $\tilde{\Pi}$ do not entertain such an equivalence relation. The table below displays the relation between two sets of notations.

p vs. q	$p \succeq^\pi q$	$p \succ^\pi q$	$p \sim^\pi q$	$p \not\succeq^\pi q$
$\pi \in \cdot$	Π	$\hat{\Pi}$	$\tilde{\Pi}$	$\check{\Pi}$

Table 4: Notations for correlation preference.

2.1 Correlation Utility Representation

We offer a general representation for binary choice which can accommodate correlation between lotteries. The following continuity axiom is maintained throughout the paper.

Axiom 0 (Continuity). Π *is closed relative to* $\Delta(X \times X)$ *under the topology of pointwise convergence.*

Definition 0 (Correlation Utility Representation). *The correlation preference* Π *is represented by a correlation utility function* $U : \Delta(X \times X) \to \mathbb{R}$ *if* U *is continuous and for* $\pi \in \Delta(X \times X)$, $\pi \in \Pi \iff U(\pi) \geqslant 0$.

It turns out that continuity is the only axiom needed to obtain a correlation utility representation.

Proposition 0. *The correlation preference* Π *admits a correlation utility representation if and only if* Π *is continuous.*

The "only if" part of the proposition is immediate, whereas the "if part" follows by, for example, setting $U(\pi) = \inf_{\pi' \in \Pi} d(\pi, \pi')$ where d is the distance defined on $\Delta(X \times X)$ for the topology of pointwise convergence. Notice that the representation is not unique: any *continuous* function that assigns nonnegative values to Π and negative values to its complement suffices.

2.2 Correlation Insensitivity and Transitivity

We first introduce a notion called correlation insensitivity, under which a correlation preference bears close relation to the classical preference under risk where completeness and transitivity are standard.

Let $\Gamma(p, q) = \{\pi \in \Delta(X \times X) : (\pi_1, \pi_2) = (p, q)\}$. The *correlation-insensitive preference* binary relation $\succeq^*$ induced by Π is defined as $p \succeq^* q \iff \Gamma(p, q) \subset \Pi$:

DM weakly prefers p to q no matter how they are correlated; equivalently, $p \succeq^\pi q$ for all $\pi \in \Gamma(p,q)$. The weak preference $\succeq^*$ naturally induces its strict and indifferent parts defined as $p \succ^* q \iff p \succeq^* q$ but not $q \succeq^* p$, and $p \sim^* q \iff p \succeq^* q$ and $q \succeq^* p$.

Axiom 1 (Correlation Insensitivity, CI)**.** *For all $(p,q) \in \Delta X \times \Delta X$, $p \succeq^\pi q$ for some π implies $p \succeq^* q$: equivalently, either $p \succeq^* q$, or $\Gamma(p,q) \cap \Pi = \varnothing$.*

A correlation-insensitive decision maker always compares a pair of lotteries p and q regardless of their correlation, i.e., preference between p and q under one correlation pins down the preference between them under any other correlations. Correlation insensitivity naturally leads to the following definition which is intermediate between correlation insensitivity and full transitivity.[10]

Axiom 2 (CI-Transitivity)**.** *For all $p,q,r \in \Delta X$, $p \succeq^* q$ and $q \succeq^* r$ implies $p \succeq^* r$.*

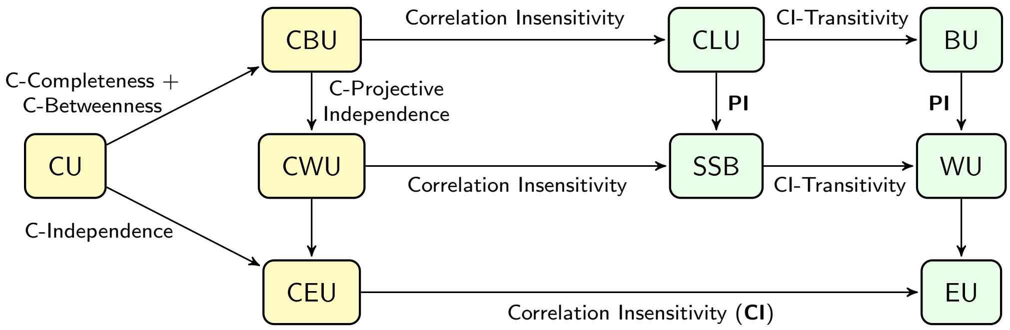
As a corollary of Proposition 0, for correlation-insensitive preference Π (and the induced $\succeq^*$), there exists a continuous $\psi : \Delta X \times \Delta X \to \mathbb{R}$ such that $p \succeq^* q \iff \psi(p,q) \geqslant 0$ with SSB utility as a special case where ψ is both skew-symmetric and bilinear in (p,q).[11] The following weaker notion of correlation insensitivity involving *same-marginal lotteries* (SML) (Loewenfeld and Zheng, 2023) will be further discussed in Section 3.

Axiom 1* (SML Correlation Insensitivity)**.** *For all $p \in \Delta(X)$, $\Gamma(p,p) \subset \Pi$.*

If DM is SML correlation-insensitive, $p \succeq^\pi p$ for any π with identical marginals (p,p). In contrast, SML correlation *sensitivity* can be viewed as a *pure* correlation preference that does not involve a difference in marginals. SML correlation insensitivity appears weaker than full correlation insensitivity. As it turns out, the two are equivalent under correlation completeness for CEU and CBU. Figure 2 illustrates the interrelations between the transitive models on the right and their correlation counterparts on the left. The next two subsections will preview CEU and CWU before they are formally discussed in sections 4 and 5 after section 3 on experimental evidence.

[10] Lanzani (2022) proposes the following transitivity: for $p,q,r \in \Delta X$ and $\pi \in \Gamma(p,q)$, $\pi' \in \Gamma(q,r)$, $p \succeq^\pi q$ and $q \succeq^{\pi'} r \Rightarrow p \succeq^* r$. It implies CI and hence the reduction from CEU to EU with correlation independence; we defer the details to Online Appendix A.

[11] Observe that the correlation-insensitive preference is complete and CI-transitive if and only if it admits a classical transitive utility, i.e., there exists a continuous $u : \Delta X \to \mathbb{R}$ such that $p \succeq^* q \iff u(p) \geqslant u(q)$ (see, e.g., Debreu (1964)).



Correlation Utility (CU)			
Correlation Expected Utility	CEU	CLU	Conditional Linear Utility
Correlation Betweenness Utility	CBU	BU	Betweenness Utility
Correlation Weighted Utility	CWU	SSB	Skew-Symmetric Bilinear Utility

Figure 2: Relation among axiomatic models of correlation preference under risk.

2.3 Correlation Expected Utility

We state below the definition of the correlation counterpart of expected utility.

Definition 1 (Correlation Expected Utility)**.** Π *admits a correlation expected utility (CEU) representation if there exists a function* $\phi : X \times X \to \mathbb{R}$ *with* $\phi(x, x) = 0$ $\forall x \in X$ *such that* $\forall \pi \in \Delta(X \times X), \pi \in \Pi \iff \mathbb{E}_\pi \phi \geqslant 0$.

Notice that linearity of the CEU representation implies continuity of the correlation preference Π and its representation. The kernel $\phi(x, y)$ reflects the utility of receiving x while foregoing y, In the setting of decision making under risk, it is useful to require monotonicity and transitivity of the underlying *certainty ordering*, i.e., $\phi(x, y) > 0$ for $x > y$ and that $\phi(x, y)$ increases in x and decreases in y.

We say the CEU preference is *symmetric* if ϕ is skew-symmetric, i.e., $\phi(x, y) = -\phi(y, x)$, as DM experiences the same level of utility change whether receiving x while foregoing y or vice versa. Symmetry of CEU preference implies $\mathbb{E}_\pi \phi > 0 \iff \pi \in \hat{\Pi}$,[12] and $\mathbb{E}_\pi \phi = 0 \iff \pi \in \tilde{\Pi}$. Asymmetry of CEU preference could come from various sources, e.g., when there is asymmetric regret in counterfactual comparison, reference dependence, and when the column outcome y is an aggregate statistic either through taking the *average* (Bordalo et al., 2022) or the *maximum*

[12] As $\mathbb{E}_{\pi^{\mathrm{T}}} \phi = -\mathbb{E}_\pi \phi < 0$, so $\pi^{\mathrm{T}} \in \check{\Pi}$.

over other candidate lotteries (Quiggin, 1994).[13]

Besides EU where $\phi^{EU}(x, y) = u(x) - u(y)$, the symmetric CEU class includes the following important utility models.

- Regret Theory (RT) (Bell, 1982; Loomes and Sugden, 1982): ϕ^{RT} is often assumed to satisfy *regret aversion* $\phi(x, y) > \phi(x, z) + \phi(z, y)$ for all $x > z > y$.
- Salience Theory (ST) (Bordalo et al., 2012): $\phi^{ST}(x, y) = \sigma(x, y)(x - y)$ with the *salience function* σ satisfying a number of properties in their Definition 1.
- The correlation-sensitive representation in Lanzani (2022) corresponds to the symmetric CEU model.

More generally, CEU includes an asymmetric model proposed by Lanzani (2022) who adapts the expectations-based reference dependent model (Kőszegi and Rabin, 2007) for binary choice. This is discussed in subsection 4.4 on asymmetric CEU.

Correlation insensitivity serves a bridging role linking CEU to EU without directly invoking transitivity. According to Theorem 2, EU is equivalent to CEU under correlation insensitivity as well equivalent to symmetric CEU under SML correlation insensitivity. We may view these results as providing an alternative avenue to axiomatize the classical expected utility theorem.[14]

Given that SSB and symmetric CEU both involve a skew-symmetric kernel, it may be tempting to view SSB as resulting from restricting the latter domain to $\Delta X \times \Delta X$. In this regard, Lanzani (2022) invokes Theorem 1 in Fishburn (1982) to conclude that the skew-symmetric kernel for a symmetric CEU is uniquely pinned down by comparisons across independent lotteries. However, this suggestion implying correlation insensitivity is not compatible with the equivalence between correlation-insensitive CEU and EU discussed in the preceding paragraph, neither is it compatible with Fishburn's (1983) demonstration of SSB reducing to WU rather than EU under transitivity.

[13] In this way, we can also study the problem of choice from multiple lotteries. Another way to formulate this problem would be to allow for multiple kernels ϕ_k measuring when each of the lotteries p_k is selected from the menu.

[14] We can induce a correlation preference $\Pi^{\succeq}$ from a correlation-insensitive preference $\succeq$. The EU axioms of *ordering* and *independence* (e.g., Mas-Colell et al. (1995)) would imply their correlation counterparts. Consequently, Theorem 1 yields a CEU representation which further giving rise to EU under Theorem 2.

2.4 Correlation Weighted Utility

The Allais (1953) paradox has provided the primary impetus for the development of transitive non-EU models. The betweenness direction in this literature has started with Chew's (1983) weighted utility (WU) which is shown in Fishburn (1983) to be the transitive counterpart to the nontransitive SSB model discussed in the Introduction. We formally state both specifications below.

Definition 2 (WU)**.** Π *admits a WU representation if there exist* $u : X \to \mathbb{R}$ *and weight function* $w : \Delta X \to \mathbb{R}^+$*, such that* $p \succeq^\pi q \iff \mathbb{E}_{w[p]}u \geqslant \mathbb{E}_{w[q]}u$*, where* $w[p](x) = \sum \frac{w(x)p(x)}{\sum w(x')p(x')}$ *where* $w(x) := w(\delta_x)$ *in an abuse of notation.*

Notice that the above WU representation admits a Radon-Nikodym derivative interpretation for the weight function $w(x)$ (Chew, 1983).

Definition 3 (SSB Utility)**.** Π *admits an SSB representation if there is a skew-symmetric bilinear* $\psi : \Delta X \times \Delta X \to \mathbb{R}$ *such that* $p \succeq^\pi q \iff \psi(p, q) \geqslant 0$*.*

SSB has a central role in our account of the *extended* Allais paradox through a characteristic property of WU called *ratio consistency* (to discuss further in Appendix B).[15] Observe that SSB is correlation-insensitive and reduces to WU under CI-transitivity. This is suggestive of the following *correlation weighted utility* (CWU) model proposed formally in Section 5.

Definition 4 (CWU)**.** Π *admits a CWU representation on* $\Delta' \subset \Delta(X \times X)$ *if there is a skew-symmetric kernel* $\phi : X \times X \to \mathbb{R}$*, and a skew-symmetric bilinear kernel* $\psi : \Delta X \times \Delta X \to \mathbb{R}$ *such that for* $\pi \in \Delta'$*,* $\pi \in \Pi \iff \mathbb{E}_\pi \phi + \psi(\pi_1, \pi_2) \geqslant 0$*.*

Notice that when $\phi = 0$, CWU reduces to the SSB utility. In subsection 5.2, we define CWU representation on the whole domain $\Delta(X \times X)$ that allows for two different ϕ in two subdomains Δ', one of which has $\phi = 0$ and thus reduces to an SSB representation. For a correlation preference Π that admits a CWU representation, it is correlation-insensitive if and only if it admits an SSB representation.

A CWU decision maker combines a correlation-sensitive symmetric CEU kernel with a correlation-insensitive SSB kernel defined on pairs of marginals. CEU can exhibit correlation-sensitive Allais behavior except in the state-by-state case involving maximal correlation. At the same time, SSB encompassing WU can only

[15]Nakamura (1990) points out that SSB satisfies the ratio consistency property: for p, q, r with $p \sim q$ and $p \nsim r, q \nsim r$, $\beta_i p + (1 - \beta_i)r \sim \gamma_i q + (1 - \gamma_i)r$ for $i = 1, 2$ implies $\frac{\gamma_1/1-\gamma_1}{\beta_1/1-\beta_1} = \frac{\gamma_2/1-\gamma_2}{\beta_2/1-\beta_2}$.

exhibit correlation-insensitive Allais behavior. In subsection 3.2, we demonstrate formally how CWU can fully account for the extended Allais paradox in exhibiting a declining rate of Allais behavior from the case of independent lotteries to the state-by-state case of maximal correlation.

In terms of parametric form, we may adopt the *constant absolute risk aversion* specification of WU in Chew and Tan (2005),[16] serving as a transitive SSB kernel $\psi(x, y) = w(x)w(y)[u(x) - u(y)]$ with $w(x) = e^{\rho x}$, $u(x) = -e^{-\lambda x}$ and delivering correlation-insensitive Allais behavior. This can be combined with the ST kernel $\phi(x, y) = \sigma(x, y)(x - y)$ where $\sigma(x, y) = \frac{|x-y|}{|x|+|y|+\theta}$ in Bordalo et al. (2012) exhibiting correlation-sensitive Allais behavior except at maximal correlation. The relative weight between the correlation-sensitive and insensitive parts will determine the degree of correlation sensitivity and the corresponding degree of Allais behavior.

3 Experimental Evidence

We provide details on how our correlation preference approach can account for the experimental findings discussed in the Introduction. We begin with the finding on pure correlation preference using SML. This is followed by evidence based on the independent and correlated versions of Allais' common-consequence problems.

3.1 Same-Marginal Lotteries

One implication of CEU is tested by Loewenfeld and Zheng (2023) who examine a key assumption in the special cases of RT and ST: $\phi(h, l) > \phi(h, m) + \phi(m, l)$ for all $l < m < h$. When skew-symmetry of ϕ is assumed, this is equivalent to $\phi(l, h) + \phi(h, m) + \phi(m, l) < 0$. DM is said to be *increasingly* (resp. *decreasingly*, *constantly*) *sensitive to outcome differences*, denoted ISOD (resp. DSOD, CSOD) if $\phi(l, h) + \phi(h, m) + \phi(m, l) <$ (resp. $>, =$) 0 for $l < m < h$. Notice that transitive non-EU is CSOD. We can verify that RT and ST are both ISOD while asymmetric CEU under ERD is DSOD.

[16]The Arrow-Pratt measure of local risk attitude $A(y)$ at wealth level y for WU is given by $-u''/u' + 2w'/w$ (Chew, 1983) while the corresponding Arrow-Pratt measure for CWU is given by $A(y) = -\frac{\bar{\phi}_{22}(y,y)+\psi_{22}(y,y)}{\bar{\phi}_2(y,y)+\psi_2(y,y)}$, where $\bar{\phi}_2(x, y) = \frac{1}{2}\phi_2(x, y^+) + \frac{1}{2}\phi_2(x, y^-)$ is the average of left and right derivatives, and similarly for $\bar{\phi}_{22}$; notation ϕ_2 means the partial derivative of the second marginal of ϕ, and similarly for others. The ϕ term turns out to equal zero so that $A(y) = -\frac{\psi_{22}(y,y)}{\psi_2(y,y)} = \lambda - 2\rho$ which is the same as in Chew and Tan (2005).

From prior papers which have tended to report *indirect* evidence supporting ISOD in conjunction with RT and ST, Loewenfeld and Zheng (2023) have distilled an SML test where subjects are asked to choose between two *interdependent* lotteries with the same marginal (see Figure 1). The test is designed to discriminate *directly* between ISOD and DSOD corresponding respectively to subjects choosing Lottery A or Lottery B in Figure 1. Notice that the explicitly presented state-by-state comparisons in the SML test may make the *counterfactual* associated with an outcome forgone in the lottery not chosen more salient.

In one lab and two online experiments involving 919 participants where Loewenfeld and Zheng (2023) collected more than 18,000 observations, the observed choices in the aggregate provide evidence implying DSOD for 38% of the participants,[17] while most participants exhibit behavior consistent with CSOD. Moreover, both aggregate and individual-level results strongly reject ISOD given that a latent class analysis does not produce a corresponding category of such participants. Observe that the Loewenfeld-Zheng finding is compatible with our nonsymmetric CEU which encompasses both symmetric and asymmetric CEU.

3.2 Extended Allais Paradox

We trace the extended Allais paradox described in the Introduction to Savage's (1954, p.102) exposition of how he corrects his error in violating EU by choosing Gamble 1 in Situation 1 and Gamble 4 in Situation 2 in the two binary choice situations below:

Situation 1. *Gamble 1*: \$500,000 with probability 1, versus *Gamble 2*: \$2,500,000 with 10% probability, \$500,000 with 89%, 0 with 1%, and

Situation 2. *Gamble 3*: \$500,000 with 11% probability, 0 with 89%, versus *Gamble 4*: \$2,500,000 with 10% probability, 0 with 90%.

Given the common perception of gambles 3 and 4 as being independent lotteries, the above may be referred to as the *independent* version of the Allais common-consequence problem.

In correcting his self confessed choice error, Savage first transforms the two choice situations into the seemingly equivalent *correlated* form illustrated in Table

[17] Around 17% exhibit DSOD behavior even if it violates FOSD (when a premium is added to the outcomes of lottery B) while another 21% display DSOD when there is no FOSD.

	Ticket Number Drawn		
	1	*2-11*	*12-100*
Gamble 1	\$500,000	\$500,000	\$500,000
Gamble 2	0	\$2,500,000	\$500,000
Gamble 3	\$500,000	\$500,000	0
Gamble 4	0	\$2,500,000	0

Table 5: Savage's presentation of Allais paradox.

5.[18] He then describes a thought process involving the application of a contingent reasoning principle, called the sure-thing principle (STP).[19] Observing that it would not matter, in either situation, which gamble is chosen, if one of the tickets numbered from 12 through 100 is drawn, he focuses on the complementary event of the ticket drawn being numbered between 1 and 11, in which case Situations 1 and 2 are exactly parallel. He writes, *"The subsidiary decision depends in both situations on whether I would trade an outright gift of \$*$500,000$ *for a 10-to-1 chance of winning win \$*$2,500,000$*, a conclusion that I think has a claim to universality, or objectivity. Consulting my purely personal taste, I find that I would prefer the gift of \$*$500,000$ *and, accordingly, that I prefer Gamble 1 to Gamble 2 and (contrary to my initial reaction) Gamble 3 to Gamble 4."* And yet, Savage writes that he still feels an intuitive attraction to his favoring Gamble 4 over Gamble 3 when they are independent after applying STP and arrive at the opposite preference when the two lotteries are correlated, pointing to the potential instance of correlation sensitivity.

In Frydman and Mormann's (2018) investigation of correlation sensitivity in the Allais common-consequence problem where correlation is described with a single parameter β (Table 6), they report a declining rate of Allais behavior from "independent" (uncorrelated) to "intermediate" to "maximally correlated" (Table 1), based on two pairs of correlated lotteries $L_1^z = (25, 33\%;\ z, 66\%;\ 0, 1\%)$ versus $L_2^z = (24, 34\%;\ z, 66\%)$, $z \in \{0, 24\}$ from Kahneman and Tversky (1979). This extended Allais paradox is incompatible with CEU which would imply no violation, as well as any transitive non-EU preference implying uniform rates of violation.

We can view a CWU decision maker (see Definition 4) as evaluating joint densities linearly, but with different thresholds $\psi(p, q)$ when the marginals (p, q) change, thereby accounting for the *extended* Allais paradox. Consider $\phi(x, y)$ given by

[18] Notice that the marginals of gambles 2, 3, and 4 are the same regardless of whether they appear in the independent or the correlated versions.

[19] "If the person would not prefer f to g, either knowing that the event B obtained, or knowing that the event $\sim B$ obtained, then he does not prefer f to g." (Savage, 1954, p.21)

$\sigma(x,y)(x-y)$ where the single-dip salience function $\sigma(x,y) = \frac{|x-y|}{|x|+|y|+\theta}$ as in Bordalo et al. (2012), and $\psi(x,y)$ be the WU kernel in subsection 2.4. When the marginals are unchanged, we have that

$$\partial U(\pi^0)/\partial\beta = -0.66 \times 24(2\theta^2 + 123\theta + 1200)/(\theta + 24)(\theta + 25)(\theta + 49) < 0,$$

so that the chance of choosing L_1^0, and hence the rate of exhibiting Allais behavior, decreases with correlation β. Meanwhile, a non-degenerate ψ kernel would allow for a non-zero rate when there is maximal correlation $\beta = 1$.

Joint prob.	0.66β	$0.66 - 0.66\beta$	$0.67 - 0.66\beta$	$0.66\beta - 0.33$
L_1^0	0	25	0	25
L_2^0	0	0	24	24

$\pi_\perp$	0	24
0	44.22%	22.78%
24		
25	21.78%	11.22%

π_{int}	0	24
0	65%	2%
24		
25	1%	32%

π_{max}	0	24
0	66%	1%
24		
25		33%

Table 6: Joint densities used in Frydman and Mormann (2018) when $z = 0$, with zero (stochastically independent), intermediate, maximal correlation.

4 Correlation Expected Utility

Our development of *correlation expected utility* (CEU) generalizes Lanzani's (2022) correlation-sensitive representation by weakening his completeness and independence axioms adapted to the correlation setting. This delivers a more flexible representation in satisfying what we call *pointwise reflexivity*, thereby accommodating the incidence of a form of preference incompleteness called *inertia*. We arrive at Lanzani's representation with a skew symmetric kernel as a corollary through an alternative route after imposing his completeness and independence axioms.

We begin with a fundamental axiom about choice involving degenerate lotteries which underpins correlation preference in general.

Axiom 3 (Pointwise Reflexivity). *$\forall x \in X$, $\delta(x,x)$ lies on the boundary of* Π.

Pointwise reflexivity requires that DM weakly prefers the degenerate lottery x to itself, but a slight perturbation away from it *could* change DM's preference. Mathematically, this is saying for any $x \in X$, $\delta(x,x) \in \Pi$ and for any of its neighborhood N, $N \bigcap \breve{\Pi} \neq \varnothing$.

We state a maintained assumption of *non-triviality* for Π throughout the paper: Π is continuous and its complement $\check{\Pi}$ as well as interior Π° relative to $\Delta(X \times X)$ are not empty. The latter holds, for instance, when Π satisfies *certainty ordering* as in subsection 2.3 and further that there exist $x > y$ and a neighborhood $\mathcal{N}$ of $\delta(x, y)$ such that $\mathcal{N} \subset \Pi$. We also assume X being compact.

4.1 Correlation Independence Axiom

Consider Samuelson's (1952) and subsequently Fishburn's (1975) four-lottery version of the independence axiom: $p \sim (\succ)\ q$ and $p' \sim (\succ)\ q'$ imply $\alpha p + (1-\alpha)p' \sim (\succ)\alpha q + (1-\alpha)q'$ for $\alpha \in (0,1)$.[20] Below is the correlation counterpart to the four-lottery independence axiom.

Axiom 4 (Independence)**.** *For any $\alpha \in (0,1)$,*

(i) $\pi, \pi' \in \Pi \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \Pi$;

(ii) $\pi, \pi' \in \check{\Pi} \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \check{\Pi}$.

The above may be stated, equivalently, in a form that resembles closely the original four-lottery independence axioms: $p \succeq^{\pi} (\not\preceq^{\pi})\ q$ and $p' \succeq^{\pi'} (\not\preceq^{\pi'})\ q'$ implies $p_\alpha \succeq^{\pi_\alpha} (\not\preceq^{\pi_\alpha})\ q_\alpha$, where $\pi_\alpha := \alpha\pi + (1-\alpha)\pi'$, and similarly for p_α and q_α.

Without completeness, correlation independence includes a separate statement on $\check{\Pi}$. Notice that it is silent on behavior within the strict preference and indifference sets while it implies that Π and $\check{\Pi}$ are both convex, so that $\tilde{\Pi}$ serves to *separate* Π and $\check{\Pi}$ in the simplex. Observe that correlation independence allows for *imprecise preference discriminability* (Fishburn, 1982; Nakamura, 1990) or, more broadly, *inertia* (Bewley, 1986): when $\pi \in \hat{\Pi}$ and $\pi' \in \tilde{\Pi}$, it is possible that their convex combination still belongs to the indifference set.[21]

We axiomatize below a general CEU representation that allows for incompleteness as well as inertia in decisions without requiring (skew) symmetry of the kernel.

Theorem 1 (Axiomatization of CEU)**.** Π *satisfies pointwise reflexivity and independence, if and only if it admits a CEU representation.*

Proof. The demonstration of sufficiency is immediate. To prove necessity, no-

[20]In a footnote, Samuelson acknowledges that the above four-lottery version of his independence axiom "differs trivially from the Paris version" where three lotteries are used, under transitivity.

[21]This is also related to the notion of *utility of gambling* (page 28 of von Neumann and Morgenstern (1944)).

tice that since both Π and $\check{\Pi}$ are convex sets, they can be weakly separated by Hahn-Banach separation theorem (see, e.g., Corollary 5.62 of Aliprantis and Border (2006)). That is, there exists some $\phi : X \times X \to \mathbb{R}$ and $c \in \mathbb{R}$ such that $\mathbb{E}_\pi \phi \geqslant (\leqslant)\ c$ if $\pi \in \Pi$ ($\check{\Pi}$). Since each π is a joint density, we may set $c = 0$.[22]

Suppose there exists a $\pi' \in \check{\Pi}$ such that $\mathbb{E}_{\pi'}\phi = 0$ $(*)$. Pick a $\pi \in \Pi^\circ$, the non-empty interior of Π (relative to the simplex). It must be that $\mathbb{E}_\pi \phi > 0$ $(**)$.[23] By continuity, there exists a $\beta \in (0,1)$ such that $\pi'' = \beta\pi + (1-\beta)\pi' \in \check{\Pi}$, so that $\mathbb{E}_{\pi''}\phi \leqslant 0$, which contradicts $(*)$ and $(**)$. Hence, $\mathbb{E}_{\pi'}\phi < 0$ for all $\pi' \in \check{\Pi}$. Consequently, $\mathbb{E}_\pi \phi \geqslant 0 \iff \pi \in \Pi$. Pointwise reflexivity gives $\phi(x,x) = 0$. $\square$

Observe that the CEU representation is unique up to multiplication by a positive constant, as ϕ now represents the hyperplane $\mathbb{E}_\pi \phi = 0$ between half-spaces. Here, convexity plays a simple but pivotal role. Notice that for CEU, the preference Π (and ϕ) are determined by the indifference set $\tilde{\Pi}$ given the convexity brought by the correlation independence axiom, which can be further pinned down from examining $\tilde{\Pi}|_{Indep}$, the indifference set within the joint densities for (stochastically) independent marginals, as it forms a spanning set given the linear structure. Alternatively, one may view ϕ and Π as determined by $\Pi|_{Indep}$ first through spanning,[24] and further by $\tilde{\Pi}|_{Indep}$ given convexity on the marginals; this is also observed by Lanzani (2022, footnote 9). These two ways of double reduction will no longer be available when we axiomatize CWU in the next section.

4.2 Correlation Strong Independence Axiom

We first strengthen pointwise reflexivity to correlation completeness.

Axiom 3* (Completeness) For all $\pi \in \Delta(X \times X)$, $\pi \in \check{\Pi} \Rightarrow \pi^{\mathrm{T}} \in \Pi$.

Observe that if Π is represented by a correlation utility function U, the correlation preference is complete if for every π, either $U(\pi) \geqslant 0$ or $U(\pi^{\mathrm{T}}) \geqslant 0$. Under correlation completeness, each symmetric π (in the sense of $\pi = \pi^{\mathrm{T}}$) belongs to $\tilde{\Pi}$. We further show in Online Appendix A that $\check{\Pi} = \hat{\Pi}^{\mathrm{T}} := \{\pi^{\mathrm{T}} : \pi \in \hat{\Pi}\}$.

We further refine correlation independence along the line what is proposed in

[22] Otherwise, let $\phi' = \phi - c$, then the corresponding threshold $c' = c - c = 0$.

[23] Otherwise, $\mathbb{E}_\pi \phi = 0$ for all $\pi \in \Pi$ by continuity. Then, the cone $Cone\,\Pi = \{\alpha\pi : \alpha \geqslant 0, \pi \in \Pi\}$ which has nonempty interior, is contained in the hyperplane $\phi \cdot \bar{\pi} = 0$, implying $\phi \equiv 0$.

[24] In fact, for a (stochastically) independent $\pi \in \Gamma(p,q)$, $\mathbb{E}_\pi \phi = \sum_{x,y} p(x)q(y)\phi(x,y)$ coincides with the SSB utility for marginals (p,q) with the SSB kernel ψ properly induced from ϕ.

Lanzani (2022) requiring that $p \sim^{\pi} q$ and $p' \sim^{\pi'} q'$ imply $p_\alpha \sim^{\pi_\alpha} q_\alpha$, while $p \succ^{\pi}$ ($\not\succ^{\pi}$) q and $p' \succeq^{\pi'}$ ($\not\succeq^{\pi'}$) q' imply $p_\alpha \succ^{\pi_\alpha}$ ($\not\succeq^{\pi_\alpha}$) q_α.

Axiom 4* (Strong Independence)**.** *For any* $\alpha \in (0,1)$*,*

(i) $\pi, \pi' \in \tilde{\Pi} \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \tilde{\Pi}$*;*

(ii) $\pi \in \hat{\Pi}, \pi' \in \Pi \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \hat{\Pi}$*;*

(iii) $\pi \in \check{\Pi} \bigcup \tilde{\Pi}, \pi' \in \check{\Pi} \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \check{\Pi}$.

Corollary 1 (Axiomatization of Symmetric CEU)**.** Π *satisfies completeness and strong independence, if and only if it admits a symmetric CEU representation.*

Proof. To see how the above corollary follows from Theorem 1, notice that the axioms already imply the existence of a CEU representation ϕ. That is, $\mathbb{E}_\pi\phi < 0 \Leftrightarrow \pi \in \check{\Pi}$. It remains to show the skew symmetry of ϕ. By correlation strong independence, for any π, $\mathbb{E}_\pi\phi > 0 \Rightarrow \pi \in \hat{\Pi}$. Now suppose $\phi(x,y) + \phi(y,x) \neq 0$ for some (x,y). Then, for $\pi = \frac{1}{2}\delta_{(x,y)} + \frac{1}{2}\delta_{(y,x)}$, $\mathbb{E}_\pi\phi$ is either positive or negative. Notice that we have $\pi^{\mathrm{T}} = \pi$ here. Then, the former case implies both $\pi, \pi^{\mathrm{T}} \in \hat{\Pi}$, contradicting the definition of $\hat{\Pi}$, while the latter implies both $\pi, \pi^{\mathrm{T}} \in \check{\Pi}$, contradicting completeness. We conclude that $\phi(x,y) + \phi(y,x) = 0$ $\forall (x,y)$. The opposite direction is immediate. □

Here our proof utilizes a correlated lottery $\pi = \frac{1}{2}\delta_{(x,y)} + \frac{1}{2}\delta_{(y,x)}$ and is substantially simpler than the proof of skew symmetry in the SSB representation of Fishburn (1982) which only makes use of independent lotteries. As mentioned earlier, the *symmetric* CEU representation shares a similar functional form with Fishburn's other utility model, nontransitive SEU (Fishburn, 1989). We will elaborate on their close connection in subsection 6.2.

4.3 Correlation Insensitivity and Expected Utility

While it is clear that CEU reduces to EU under transitivity, we explore the extent to which correlation sensitivity may play a similar role. As exposited in Section 2, CEU (and symmetric CEU) reduces to EU via correlation insensitivity (SML correlation insensitivity. Here, we demonstrate how we arrive at this finding through ideas from *optimal transport* theory (OT).

Recall from subsection 2.2 that the correlation-insensitive preference relation $\succeq^*$ is defined as $p \succeq^* q \iff \Gamma(p,q) \subset \Pi$, which is equivalent to

$\min_{\pi\in\Gamma(p,q)} \sum_{x,y} \phi(x,y)\pi(x,y) \geqslant 0$. From the Kantorovich (1942) duality, we arrive at the following equation

$$\min_{\pi\in\Gamma(p,q)} \mathbb{E}_\pi \phi = \max_{\substack{u,v \\ u(x)-v(y)\leqslant\phi(x,y)}} \mathbb{E}_p u - \mathbb{E}_q v.$$

This plays a valuable role in linking CEU to EU through correlation insensitivity without directly involving transitivity.

Theorem 2 (Correlation Insensitivity and Expected Utility)**.** *Given* Π*, the following are equivalent:*

(i) Π *admits a CEU representation and exhibits correlation insensitivity;*

(ii) Π *admits a symmetric CEU representation and exhibits SML correlation insensitivity;*

(iii) Π *admits a EU representation.*

We provide a sketch of the the proof of Theorem 2 leaving its formal proof to Appendix A. To see that (i) implies (iii), We can invoke the Kantorovich duality. For a CEU representation ϕ, consider the following minimax problem on $X' \times X'$ with finite support:

$$\min_{(p,q)\in\mathcal{D}} \max_{\substack{u,v \\ u(x)-v(y)\leqslant\phi(x,y)}} \sum_x u(x)p(x) - \sum_y v(y)q(y),$$

where $\mathcal{D} = \{(p,q) : p \succeq^* q\}$. By the equivalent definition, correlation insensitivity and pointwise reflexivity implies the problem has value 0. By correlation independence and correlation insensitivity, $\mathcal{D}$ is convex; moreover, it is compact by continuity. We can restrict the dual variables (u,v) to a bounded set due to properties of the Kantorovich duality, so that it is also compact and convex. By minimax theorem, we can swap min and max; hence, $\exists(\hat{u},\hat{v})$ such that $\hat{u}(x)-\hat{v}(y) \leqslant \phi(x,y)$ and $\sum \hat{u}(x)p(x) - \sum \hat{v}(y)q(y) \geqslant 0$ holds for all $(p,q) \in \mathcal{D}$. Pointwise reflexivity then implies $\hat{v} = \hat{u}$ through the dual constraint. Finally, we apply the Kantorovich duality once again, but in the opposite direction, to show that $\sum \hat{u}(x)p(x) - \sum \hat{u}(y)q(y) \geqslant 0$ implies $p \succeq^* q$. This completes the proof of (i) $\Rightarrow$ (iii) on $X' \times X'$. Uniqueness of ϕ then guarantees the extension to the whole space.

Although we can apply a similar duality argument to show that (iii) follows from (ii), we invoke a different one here that highlights the respective roles of SML correlation insensitivity and skew-symmetry of ϕ. Notice that the two together imply $\sum \phi(x,y)\pi(x,y) = 0$ for all π with SML (so that $\pi \in \tilde{\Pi}$). Then, we use a perturbation argument to show that ϕ must have zero *synergy* in the sense of Anderson and Smith (2024), namely $\phi(x_1, y_1)+\phi(x_2, y_2)-\phi(x_1, y_2)-\phi(x_2, y_1) = 0$.[25] It follows that ϕ is additively separable and thus reduces to EU through skew-symmetry.

Three versions of linear duality play the primary role in this demonstration. *Convexification* and then *de-convexification* are used to link the two theories through the separation theorem, the Kantorovich duality, and the minimax theorem. This alternative route involves the strategy of convexification in a way similar to the essence of Aumann's (1974) argument for extending Nash equilibrium to correlated equilibrium. The idea resembles the proof of Hart and Schmeidler (1989) on the existence of correlated equilibrium, revealing explicitly the advantage of using preference sets. The idea of convexification (*without* introducing correlation) relates to the proof by Fishburn (1975), who makes use of the cone $C = Cone\{p - q : p \geq q\}$. A similar technique also appears in Dubra et al. (2004) for an incomplete preference under EU (see subsection 7.1 for further discussion).

Together, theorems 1 and 2 imply that it is not possible to have non-EU preferences that satisfy correlation independence and are fully determined by comparing marginals alone. In Theorem 2, points (i) and (ii) strengthen Lanzani's Proposition 1 in two ways: point (i) relaxes symmetry, while point (ii) limits the examination of correlation insensitivity to SML. Both correlation insensitivity and SML correlation insensitivity are weaker than the classical notion of transitivity. Indeed, correlation insensitivity only concerns a pair of lotteries without referring to a third lottery, while SML correlation insensitivity only involves one lottery. While Theorem 2 shows that correlation insensitivity does imply transitivity of CEU preference, this will not be the case for correlation preferences satisfying a betweenness property which will developed in the next section.

4.4 Asymmetric Correlation Expected Utility

One may wonder whether SML correlation insensitivity in Theorem 2 suffices to guarantee the reduction to EU without requiring symmetry. That this is not the

[25] We discuss in Online Appendix E how OT helps identify correlation sensitivity for CEU through identifying *extremal correlations (joint densities)*, and more broadly, how OT relates to other strands of literature such as matching, no-regret-learning, and information design.

case is shown below using Lanzani's (2022) adaptation of the *expectations-based reference dependent* (ERD) model (Kőszegi and Rabin (2007) for binary choice:

$$\phi^{ERD}(x, y) = (\lambda \cdot \mathbf{1}_{x>y} + \mathbf{1}_{x \leqslant y})(x - y), \ \lambda > 1.$$

Notice that $\phi^{ERD}(x, x) = 0$, but it is not skew-symmetric when $\lambda > 1$. In fact, as observed by Lanzani (2022), ERD violates correlation strong independence. Observe that the function $\phi^{ERD}(x, y)$ is submodular. Then, when $q = p$, to minimize the OT primal, we should choose the *perfectly positively* correlated $\pi^{perfect} = \sum p(x)\delta(x, x)$. This gives the primal objective value 0, i.e., $\min_{\pi \in \Gamma(p,p)} \sum \phi(x, y)\pi(x, y) = 0$, so DM is in fact SML correlation-insensitive. In contrast, as long as p is not degenerate, we have $\max_{\pi \in \Gamma(p,p)} \sum \phi(x, y)\pi(x, y) > 0$ under a *perfectly negative* correlation. Then, for q that shifts a small probability from the best outcome to the worst outcome of p, we must have $\min_{\pi \in \Gamma(q,p)} \sum \phi(x, y)\pi(x, y) < 0$ while $\max_{\pi \in \Gamma(q,p)} \sum \phi(x, y)\pi(x, y)$ remains positive. Thus, SML correlation insensitivity is strictly weaker than correlation insensitivity when the correlation preference is not symmetric.

As it turns out, SML correlation insensitivity implies a generalized reference dependence representation through an adaptation from our proof of Theorem 2.[26]

Definition 5 (Generalized ERD)**.** *The preference* Π *admits a generalized expectations-based reference dependence (generalized ERD) representation if it can be represented by a* ϕ *in the form of* $\phi(x, y) = u(x) - u(y) + m(x, y)$*, where* $m(x, y) \geqslant 0 \ \forall (x, y)$*, and* $m(x, x) = 0 \ \forall x$*.*

Proposition 1. *Suppose* Π *admits an asymmetric CEU representation. The following are equivalent:*

(i) Π *exhibits SML correlation insensitivity;*

(ii) Π *has a generalized ERD representation.*

For a non-skew-symmetric ϕ, the corresponding preference set Π could either be incomplete or violate strong independence. In his Example 3, Lanzani argues that ERD representation violates the latter: when perturbing slightly away from a $\pi \in \tilde{\Pi}$ by mixing with a $\pi' \in \hat{\Pi}$, DM's preference could remain unchanged. This difference between the two independence axioms and their induced representations is akin to

[26] As in the proof of (i) implying (iii), we can perform a similar exercise on the set $\bar{\mathcal{D}} = \{(p, p) : p \succeq^* p\}$, which is equal to the compact and convex set $\{(p, p) : p \in \Delta X\}$ by SML correlation insensitivity. We would obtain $(\bar{u}, \bar{v})$ such that $\bar{u}(x) - \bar{v}(y) \leqslant \phi(x, y)$ and $\sum \bar{u}(x)p(x) - \sum \bar{v}(y)p(y) \geqslant 0$ holds for all p, and further $\bar{v} = \bar{u}$. Let $m(x, y) = \phi(x, y) - (\bar{u}(x) - \bar{u}(y))$. Then, dual constraint implies $m \geqslant 0$ while $m(x, x) = 0$ as $\phi(x, x) = 0$.

that which is reflected in Fishburn's SSB utility versus Nakamura's (1990) non-SSB utility. In his model, Nakamura also relaxes the constraint on combining pairs of lotteries across strict preference and indifferent sets, thereby resulting in a non-skew-symmetric representation. He relates his utility model to a generalized SSB utility with a *threshold function*, which bears some resemblance with the $m(x, y)$ term in our generalized ERD model.

Incompleteness may also arise from behavioral and psychological factors such as inertia, status quo bias,[27] and reference point effect.[28] One way to incorporate reference point involving correlation is discussed in Cerreia-Vioglio et al. (2024). They propose a *cautious utility* model where DM evaluates each lottery p with respect to a stochastic reference given by a fixed lottery r which may be correlated with p. The utility of p is the expectation of the utility of *difference* between the realized and reference outcomes over the *joint* probability, giving rise to a complete and transitive model.

5 Correlation Weighted Utility

As discussed in the Introduction, the CEU model maintaining linearity in the joint density cannot account for the extended Allais paradox. This motivates the development of non-CEU model here by formulating the correlation counterparts of two weakenings of the independence axiom, namely betweenness (subsection 5.1) and projective independence (subsection 5.3). The resulting axiomatizations of correlation betweenness utility and correlation weighted utility are discussed in the sequel.

5.1 Correlation Betweenness Axiom

The statement of our correlation betweenness axiom below bears some resemblance to the betweenness property in Chew (1989) requiring probability mixtures of a pair of indifferent lotteries to also be indifferent to each of them. The resulting planar indifference surfaces each divides ΔX into convex better-than and worse-than sets.

[27]See, e.g., Bewley (1986); Masatlioglu and Ok (2005); Sagi (2006).

[28]Consider the ERD model with representation function: $\phi^{ERD}(x, y) = (\lambda \cdot \mathbf{1}_{x>y} + \mathbf{1}_{x\leqslant y})(x - y)$, $\lambda > 1$. One feature of the preference is that due to the submodularity of ϕ^{ERD}, it involves "too much" completeness in the sense that there exists π such that both π and $\pi^{\mathrm{T}} \in \Pi$. This also reflects the effect of reference point. On the other hand, if we instead require $\lambda < 1$, then completeness no longer holds while strong independence remains satisfied.

Axiom 5 (Betweenness)**.** *For any $\pi, \pi' \in \Delta(X \times X)$ such that $\pi_1 = \pi'_1$, and $\alpha \in (0,1)$,*

(i) $\pi, \pi' \in \tilde{\Pi} \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \tilde{\Pi}$*;*

(ii) $\pi \in \hat{\Pi}, \pi' \in \Pi \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \hat{\Pi}$*;*

(iii) $\pi \in \check{\Pi} \bigcup \tilde{\Pi}, \pi' \in \check{\Pi} \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \check{\Pi}$.

For comparison, correlation independence implies that both Π and $\check{\Pi}$ are convex by strengthening the convexity on $\Delta_p = \{\pi \in \Delta(X \times X) : \pi_1 = p\}$ across different marginals.

Definition 6 (Correlation Betweenness Utility)**.** Π *admits a correlation betweenness utility (CBU) representation if for each $p \in \Delta X$, there exists a ϕ_p continuous in p such that $\pi \in \Pi \iff \mathbb{E}_\pi \phi_{\pi_1} \geqslant 0$, and $\pi \in \tilde{\Pi} \iff \mathbb{E}_\pi \phi_{\pi_1} = 0$.*

The CBU representation generalizes the following *conditional linear utility* (CLU) representation due to Fishburn (1982): for each $p \in \Delta X$, there exists a linear $v_p : \Delta X \to \mathbb{R}$ such that $p \succ q$ (resp. $\sim$) $\iff$ $v_p(q) < 0$ (resp. $=$). Our CBU kernel is also sensitive to the row marginal p, while further allowing for correlation sensitivity. The nontransitive CLU plays a key role in Fishburn's (1982, Lemma 3) proof of the SSB representation.

Theorem 3 (Axiomatization of CBU)**.** Π *satisfies correlation betweenness if and only if it admits a CBU representation.*

The proof is similar to the one for Theorem 1 (proofs for this section are provided in Appendix B) with the separation argument now applied to the subdomain Δ_p for each $p \in \Delta X$. Each ϕ_p concerns only those π whose first marginals equal p. Viewed as a separating hyperplane, it intersects the subspace (subset) Δ_p at the indifference set, while it is unconstrained in the direction orthogonal to this subspace. In the sequel, denote $\Delta_P = \bigcup_{p \in P} \Delta_p$ for $P \subset \Delta X$.

5.2 Correlation Completeness and Correlation Sensitivity

Assume that the outcome space X is finite with $|X| = N$. We show now that a correlation-insensitive CBU reduces, under CI-transitivity, to the class of *betweenness utility* due to Dekel (1986) and Chew (1989). To see this, suppose DM is correlation-insensitive. By applying a duality argument similar to that in the proof of Theorem 2, we obtain the following two results.

Definition 7 (Conditional Linear Utility (Fishburn, 1982)). *The preference* Π *admits a conditional linear utility (CLU) representation if there exists a linear* v_p *for each* $p \in \Delta X$ *such that* $p \succ^\pi q$ *(resp.* $\sim^\pi$*) if and only if* $v_p(q) < 0$ *(resp.* $=$*).*

Proposition 2. *A CBU representation is correlation-insensitive and complete if and only if it admits a CLU representation.*

While correlation insensitivity implies transitivity for CEU by Theorem 2, this is not the case for CBU according to the above proposition. As illustrated in Figure 2, CLU reduces further to BU under our weaker CI-transitivity requirement.

Definition 8 (Betweenness Utility). Π *admits a BU representation if there exist implicit utility function* V *and utility function* $u : X \times \mathbb{R} \to \mathbb{R}$ *defined as* $\sum_x u(x, V(p))p(x) = V(p)$ *such that* $p \succ^\pi q$ *(resp.* $\sim^\pi$*) if and only if* $V(p) > V(q)$ *(resp.* $=$*).*

Proposition 3. *A CLU representation admits a BU representation if and only if it is CI-transitive.*

From the two propositions, we arrive at the following.

Theorem 4 (Correlation Insensitivity and Betweenness Utility). *CBU is correlation-insensitive, complete, and CI-transitive if and only if it is BU.*

We first demonstrate how the correlation completeness axiom suffices to guarantee a CWU representation for a CBU preference on a correlation-sensitive subdomain of $\Delta(X \times X)$ through the following definition of a local notion of correlation (in-)sensitivity.

Definition 9 (Local Strict Correlation Insensitivity). Π *is locally strictly correlation-insensitive on* (p, q)*, if either* $\Gamma(p, q) \subset \hat{\Pi}$ *or* $\Gamma(p, q) \subset \check{\Pi}$.

With the above definition, we arrive at the following partition of $\Delta X \times \Delta X$ consisting of three subsets:

(i) $\mathcal{D}^I$ consists of all pairs (p, q) satisfying the above definition.

(ii) $\mathcal{D}^{Ind}$ includes all pairs (p, q) such that $\Gamma(p, q) \in \tilde{\Pi}$.

(iii) $\mathcal{D}^S$ consists of the rest at which DM is *locally correlation-sensitive.*

For a subset $\mathcal{Q} \subset \Delta X \times \Delta X$, we say there exists a *generalized skew-symmetric* (GSS) representation $\{\phi_p\}$ in $\mathcal{Q}$ if for every $(p, q) \in Q$, $\mathbb{E}_\pi \phi_p + \mathbb{E}_{\pi^\intercal} \phi_q = 0$ for all $\pi \in \Gamma(p, q)$. This generalizes the original notion of skew-symmetry where $\phi_p = \phi$ for all $p \in \Delta X$, which relates closely to correlation completeness in CEU. The proof of the following proposition applies a theorem of the alternative in linear programming.

Proposition 4 (Generalized Skew Symmetry). *A CBU preference satisfying completeness admits a generalized skew-symmetric representation in* $\mathcal{D}^S \bigcup \mathcal{D}^{Ind}$.

We next introduce a partition of ΔX using the following equivalence result between correlation insensitivity and SML correlation insensitivity. This echoes a similar equivalence result obtained for symmetric CEU preference in Theorem 2.

Proposition 5 (SML Correlation Insensitivity). *For a CBU preference satisfying completeness, correlation insensitivity is equivalent to SML correlation insensitivity.*

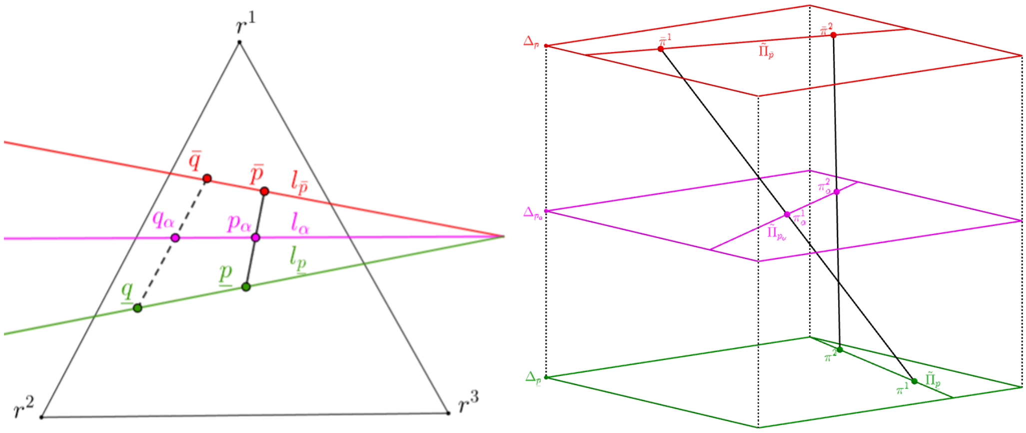
By Proposition 5, we can partition $\Delta X = \mathcal{P}^I \bigcup \mathcal{P}^S$, where $\mathcal{P}^I$ consists of all those p that are SML correlation-insensitive and $\mathcal{P}^S$ of those being SML correlation-sensitive. We further decompose $\mathcal{P}^I$ and $\mathcal{P}^S$ into *connected components*. Notice that by continuity, each $p \in \mathcal{P}^S$ has a neighborhood that is also SML correlation-sensitive, and hence $\mathcal{P}^S$ is open and so is each of its connected component. $\mathcal{P}^I$ is thus closed (and compact).

5.3 Projective Independence Axiom

Notice that if DM admits a CWU representation on a subdomain, then Π satisfies a correlation counterpart of the *projective independence* axiom (Chew et al., 1994): if $\bar{p} \sim \bar{q}$, $\underline{p} \sim \underline{q}$, and there is an $\alpha \in (0,1)$ such that $\alpha\bar{p} + (1-\alpha)\underline{p} \sim \alpha\bar{q} + (1-\alpha)\underline{q}$, then for all $\beta \in (0,1)$, $\beta\bar{p} + (1-\beta)\underline{p} \sim \beta\bar{q} + (1-\beta)\underline{q}$ (see Figure 3 below), motivated by Samuelson's (1952) 4-lottery version of the independence axiom. Together with betweenness, it implies Fishburn's SSB utility, reducing further to WU under CI-transitivity as previewed in subsection 2.4. Below is a statement of projective independence extended to the correlation setting.

Axiom 6 (Projective Independence). *For any* $\pi, \pi' \in \tilde{\Pi}$, *if* $\exists \alpha \in (0,1)$ *such that* $\alpha\pi + (1-\alpha)\pi' \in \tilde{\Pi}$, *then* $\beta\pi + (1-\beta)\pi' \in \tilde{\Pi}$ *for all* $\beta \in (0,1)$.

We next show that correlation projective independence (see also Figure 3) specifies a uniform CWU representation across different connected components $P^S \subset \mathcal{P}^S$, and further dictates the bilinear representation on the entire domain.



Left: With transitivity, for each $\tilde{p} \in \{p, p', p_\alpha\}$, each color represents an indifference $l_{\tilde{p}}$ in a 2-dimensional probability simplex of $\{r^1, r^2, r^3\}$. Projective independence implies once p_α and q_α are on a same indifference curve for one $\alpha \in (0,1)$, then this applies for all $\alpha \in (0,1)$, and hence all l_p are projective.

Right: Each color represents a subdomain $\Delta_{\tilde{p}}$ where solid line segments are the respective indifference sets $\tilde{\Pi}|_{\tilde{p}}$ separating $\hat{\Pi}|_{\tilde{p}}$ and $\check{\Pi}|_{\tilde{p}}$. CPI implies that once $\pi_\alpha \in \tilde{\Pi}|_{p_\alpha}$ for one $\alpha \in (0,1)$, then this applies for all other α, and hence the black line segment is in the indifference set $\tilde{\Pi}$.

Figure 3: Projective independence and its correlation counterpart.

Assumption R. For any point $p \in \mathcal{P}^I$, there exists a sequence in the interior of $\mathcal{P}^I$ that converges to p.

The above regularity assumption requires $\mathcal{P}^I$ to equal the closure of its interior $int\mathcal{P}^I$.[29] Specifically, if $\mathcal{P}^I$ is non-empty, then so is its interior.

Proposition 6. *Under Assumption R, if the CBU preference Π satisfies completeness and projective independence, then it admits a CWU representation on the whole domain: a non-degenerate CWU on $\mathcal{P}^S$ and an SSB utility on $\mathcal{P}^I$.*

The dichotomous representation is obviously necessary and sufficient for the axioms within each subdomain. When comparing two marginals from different subdomains, correlation completeness further requires the following *cross-domain completeness*: for π with $\pi_1 \in \mathcal{P}^S$ and $\pi_2 \in \mathcal{P}^I$, $\phi \cdot \pi + \psi(\pi_1, \pi_2) > 0 \Rightarrow V(\pi_2, \pi_1) < 0$, and similarly $\phi \cdot \pi + \psi(\pi_1, \pi_2) <$ (resp. $=$)$0 \Rightarrow V(\pi_2, \pi_1) >$ (resp. $=$)0. Observe that in those cases, the ψ term locally *dominates* the ϕ term, so that the overall preference appears *locally insensitive* at (π_1, π_2). With this observation, we arrive at the following characterization theorem from Proposition 6.

Theorem 5 (CWU representation)**.** *Under Assumption R, Π satisfies between-*

[29]Both relative to the affine subspace of the joint densities.

ness, completeness, and projective independence if and only if it admits a CWU representation (ϕ, ψ) *on* $\mathcal{P}^S$ *and an SSB utility representation* V *on* $\mathcal{P}^I$ *satisfying cross-domain completeness.*

One may have noticed the seeming discontinuity in representation across the boundary of two subdomains. If we forgo bilinearity and allow for a re-scaling on $\phi_p = \phi + \psi$, then we can smoothly paste the two representation at each π on the boundary as $\gamma_\pi(\phi \cdot \pi + \psi(\pi_1, \pi_2)) = V(\pi_1, \pi_2)$, with the weight λ_π being continuous in π and positive as required by correlation completeness across domains. We then continuously extend γ_π to the whole domain to obtain a global continuous $U(\pi)$ representation.

When DM is correlation-insensitive, the CWU representation reduces to SSB utility $\psi'(p, q)$, and further to WU with the kernel $\psi'(x, y) = w(x)w(y)[v(x) - v(y)]$ under CI-transitivity given that (uncorrelated) axioms such as betweenness and projective independence are satisfied, which is summarized in the proposition below following the earlier results by Fishburn (1982, 1983). This can also be viewed as the limit of a sequence of CWU in which the CEU kernels ϕ converging to a null kernel $\phi = 0$. Specifically, an SSB utility can be viewed as a correlation-insensitive preference Π satisfying correlation completeness, correlation betweenness, and correlation projective independence.

Proposition 7 (SSB and WU)**.** *For* Π *admitting a CWU representation on the whole domain, it satisfies correlation insensitivity if and only if it is SSB. Moreover, SSB utility is CI-transitive if and only if it admits a WU representation.*

SSB, WU, and Symmetric CEU. When developing his nontransitive SEU, Fishburn (1989) states that "*SSB theory is both a parent and a child of*" nontransitive SEU as, in particular, the skew-symmetric representation of SSB utility "*emerges as the specialization*" of the nontransitive SEU representation "*for pairs of stochastically independent acts*". Subsequently, Lanzani (2022) points out that a CEU model "*coincides*" with the SSB utility model "*when restricted to the comparison between independent lotteries*", and writes, "*Theorem 1 provides an alternative set of axioms for the SSB model*". From the perspective of shrinking the domain to the set of independent joint densities, strong independence is not applicable since a convex combination of two arbitrary *independent* joint densities may not be independent. Instead, should we wish to apply strong independence to the whole simplex $\Delta(X \times X)$, we would need to impose correlation insensitivity so that the "restriction" is well defined. In this regard, our Theorem 2 establishes that a correlation-insensitive CEU model, without invoking transitivity, reduces to EU, not SSB utility. Moreover, we show that any SSB model can be recast as a correlation-insensitive CWU model and vice versa.[30]

6 Correlation Preference in a Savagean Setting

The two main directions of research following Savage (1954 – S54) are illustrated in Figure 4 below. The direction to the left (Fishburn, 1989 – F89) drops transitivity while maintaining completeness and Savage's P2. This yields his skew-symmetric additive representation corresponding to *non-transitive subjective expected utility* (NTSEU). Machina and Schmeidler (1992—MS92) initiate the *transitive* probabilistic sophistication (TPS) direction, depicted by the right arrow, which maintains transitivity and completeness without imposing P2. This direction encompasses a rich range of non-EU models, including WU, BU, quadratic utility, and rank-dependent utility. It is noteworthy that this direction also includes the transitive but incomplete multi-utility representation of Dubra et al. (2004) which will be discussed at the start of the next section. Subsuming both directions is our generalized definition of *probabilistic sophistication* (PS), at the bottom of the figure, first suggested in MS92 (p. 770) and further developed in the present paper (CCZ24).

From TPS, the bottom-right arrow points towards PS without requiring completeness nor transitivity. This direction allows for preference to be *locally* CI-transitive in some subdomains: for instance, a CWU preference could be transitive on its CI subdomain $\mathcal{P}^I$ with a WU representation. In the absence of correlation sensitivity, PS strengthens to correlation-insensitive PS (CIPS), exemplified by SSB which reduces, under (CI-) transitivity, to WU exemplifying TPS. Meanwhile, PS can also be achieved from the bottom-left direction generalizing nontransitive SEU

[30]Recently, Chambers et al. (2024) study coherent distorted beliefs that commute with conditioning: DM obtains the same decision weights regardless of distortion preceding information updating or updating preceding distortion. They show that the only suitably continuous coherent distorted beliefs that induce expected utility preferences are those that are weighted per the WU representation. Interestingly, if we apply such idea to the domain of joint densities (i.e., belief as a joint density $\pi \in \Delta(X \times X)$), then the same analysis would imply that CEU already captures coherency: any coherent distortion can be "absorbed" into the utility term, so that equivalently there is no need of distorting beliefs.

by dropping P2. This direction can accommodate more general preferences, such as those corresponding to CBU and CWU, given that our PS axiomatization (CCZ24) does not rely on P2.[31]

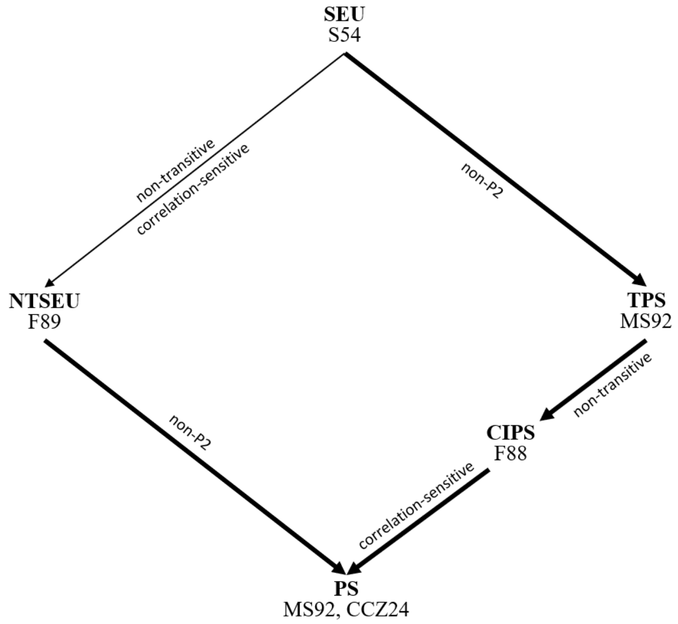


Figure 4: Parallelisms involving transitivity, P2, and correlation sensitivity

6.1 Framing Effects and Consequentialism

In his development of nontransitive SEU, Fishburn (1988) offers a binary choice principle as basis for a rigorous discussion of *framing effects* in Section 2.

Distribution Reduction (DR). "*For comparative purposes of preference and choice in risky decisions, it suffices to characterize each alternative in terms of its probability distribution over potential outcomes.*"

DR may be viewed as a *consequentialist* decision making principle in the sense that choice behavior across situations ought to remain the same should DM perceive the underlying alternatives as being consequentially the same. Fishburn then illustrates DR with two situations each involving a choice between the "same" pair of even-chance bets—$(1000, 1/2; 0, 1/2)$ versus $(1200, 1/2; -80, 1/2)$—where the only difference concerns whether one or two coins are used.

SI. "*A fair coin is to be flipped. Under* a_1*, you win \$1,000 if a head appears and*

[31] Just as the translation for CEU and nontransitive SEU in the online appendix, we can perform a similar exercise to arrive successively at CBU and then CWU in the Savagean setting by applying correlation betweenness followed by adding correlation projective independence and systematically translate the axioms in the risk setting back to those in the Savagean setting.

get \$0 if a tail appears; under a_2 you win \$1,200 if a head appears and lose \$80 if a tail appears."

SII. "*Two fair coins are to be flipped. Under a_1 you win \$1,000 if the first coin lands heads and get \$0 otherwise; under a_2 you win \$1,200 if the second coin lands tails and lose \$80 otherwise.*"

According to Fishburn, choice behavior in the two situations – interdependent **SI** and independent **SII** – ought to be the same under DR, i.e., does not depend on whether a single coin is tossed (**SI**) or two coins are used (**SII**). Consequently, an observed "inconsistency" in choice across the two situations would constitute an error should the two situations be perceived as being *consequentially* the same. This implication applies directly to the discussion in subsection 3.2 of how Savage (1954, p.103) recasts gambles 3 and 4 of the *independent* Allais paradox into the state-by-state comparison displayed in Table 5. Viewing the two comparisons as being consequentially the same, Savage confesses that he made an *error* in choosing Gamble 4 over Gamble 3 when they are *causally* independent given that he favors Gamble 3 over Gamble 4 in Table 6, which is in line with his preference for Gamble 1 over Gamble 2 after applying STP twice, tantamount to applying P2.

The discussion here points to different takes for when DR is violated. The one exemplified by the 1952 Allais-Savage exchange may be attributed to framing effects leading to Savage's correction. Fishburn (1988) offers another take based on his sense that the independence axiom, applied to causally *independent* lotteries as in **SII**, does not have the same degree of normativity as its P2 counterpart applied to lotteries implemented in an *interdependent* state-by-state manner as in **SI**. In other words, a comparison between two lotteries may be perceived as being consequentially distinct depending on whether a single coin or two coins are used. Fishburn makes use of this perspective to motivate his development of nontransitive SEU, which is related to his prior axiomatization of SSB.

Notice that DR could also be stated in terms of the certainty equivalent of a lottery depending exclusively on its outcomes and probabilities.[32] This would, however, rule out the possibility of intransitive preference such as SSB which is consonant with Fishburn's statement above. The present paper embraces a broader consequentialism principle than DR: *binary choice situations are consequentially the*

[32]This is implicitly the case with MS92, Grant (1995), and Chew and Sagi (2006).

same if the underlying lotteries have the same joint densities.

6.2 Correlation Probabilistic Sophistication without Completeness nor Transitivity

To go beyond nontransitive SEU as well as TPS, we adapt Chew and Sagi's (2006) characterization of TPS without relying on continuity nor monotonicity for the preference ordering. Formally, S and X refer to the sets of states and outcomes. The set of acts $F \subset X^S$ consists of mappings from states to a *finite* subset of outcomes. For a preference relation on F, we abuse the notation and let Π be the set of acts (f, g) such that DM weakly prefers f over g. Notice that Π plays a similar role as the preference set Π in the preceding exposition of correlation preference in a risk setting. As before, we also write $f \succeq g$ for $(f, g) \in \Pi$ when there is no confusion in notation.

Definition 10 (Probabilistic Sophistication). *DM exhibits (correlation) probabilistic sophistication if there exists a probability measure μ on S such that for any two pairs of acts (f, g) and (f', g') inducing the same joint density under μ, $f \succeq g \iff f' \succeq g'$.*

DM is said to exhibit PS with respect to μ if such a μ exists.

We offer the following definition as the correlation counterpart to event exchangeability in Chew and Sagi (2006). For outcomes $x, x' \in X$, acts $f \in F$, and disjoint events $E, E' \subset S$, let $xEx'E'f$ denote the act giving outcomes x, x' on event E, E' respectively, and identical to f on the rest.

Definition 11 (Event Exchangeability). *Two disjoint events E and E' are exchangeable, denoted $E \approx E'$, if $xEx'E'f \succeq yEy'E'g \iff x'ExE'f \succeq y'EyE'g$ for all $x, y, x', y' \in X$ and acts f, g.*

Like its non-correlated counterpart, event exchangeability can be viewed as a notion of equal likelihood: if $E \approx E'$, DM does not change her preference over f and g if the outcomes on E and E' are swapped. This yields the definition of *comparative likelihood* $\succeq^C$ and the corresponding statements of the three axioms in Chew and Sagi (2006): (C) Completeness of $\succeq^C$, (A) Event Archimedean Property, and (N) Event Nonsatiation, in the same ways as defined by Chew and Sagi (2006).

Definition 12 (Comparative Likelihood)**.** *For any events E, E', $E \succeq^C E'$ whenever there exists a sub-event $e \subset E \backslash E'$ such that $e \approx E' \backslash E$.*

Like event exchangeability, $\succeq^C$ represents an at-least-as-likely relation.

Axiom 7. *Axioms C, A, and N of Chew and Sagi (2006):*

- *Completeness (C): given any disjoint pair of events E and E', either $E \succeq^C E'$ or $E' \succeq^C E$.*
- *Event Archimedean Property (A): any sequence of pairwise disjoint and non-null events $\{e_i\}_{i=0}$ such that $e_i \approx e_{i+1}$ for every i is necessarily finite.*
- *Event Non-satiation (N): for any pairwise disjoint E, E', A, if $E \approx E'$ and A is non-null, then there exists no sub-event $e \subset E'$ such that $e \approx E \bigcup A$.*

We arrive at a characterization of PS without providing the proof, adapted from Chew and Sagi (2006), which is straightforward and somewhat repetitious.

Theorem 6. *Axioms C, A, and N are satisfied if and only if (i) there exists a unique, solvable, and finitely additive agreeing probability measure μ for $\succeq^C$, (ii) any two events have the same measure if and only if they are event exchangeable, and (iii) DM exhibits PS with respect to μ.*

Compared to the original proof, notice that the equivalence relation in Definition 10 assumes the role of transitive indifference played by Chew and Sagi (2006) in the derivation of an agreeing probability measure.

According to the proposition below (see proof in Appendix C), imposing transitivity would obliterate any influence of correlation sensitivity for binary choice in the Savagean setting of a single source of uncertainty. The next subsection adapts the small worlds perspective of Chew and Sagi (2008) which enables different degrees of correlation sensitivity, including TPS, in different small worlds.

Proposition 8 (PS and TPS)**.** *For Π satisfying completeness and PS, if it is transitive, then it satisfies TPS.*

Notice that correlation insensitivity serves an intermediate role between PS and TPS as reflected in the relation between SSB and WU in Figure 2. Under PS, Lanzani's (2022) symmetric CEU in the risk setting is traceable to F89 in a Savagean framework. In this regard, Fishburn imposes an axiom S1* with some flavor of STP, serving as a precursor to Lanzani's (2022) correlation strong independence.

The relevant definitions and statements are formally stated in Online Appendix C. Given the key role of completeness in F89, Fishburn's arguments do not extend to general, non-symmetric CEU.

6.3 Small Worlds

Under the Savagean perspective, DM always faces correlated lotteries given that all uncertainties arise from a single grand world. Yet, this does not matter given the maintained assumption of transitivity in the preference ordering. In this subsection, we explore the implications of having multiple sources of uncertainty by adapting the axiomatization of small worlds TPS in Chew and Sagi (2008).[33]

Building on the approach taken in the preceding subsection, we adapt the definition of event exchangeability to the small worlds setting. This yields a definition of small worlds comparative likelihood on which to apply the axioms of Completeness, Archimedean, and Non-satiation. This process can then deliver a source-dependent PS in which events are duly partitioned into different collections (small worlds); a bet belongs to the small world with respect to which it is measurable. Depending on the small world associated with the specific binary choice situation, DM possesses varying degrees of correlation sensitivity, ranging from being transitive and fully correlation-insensitive to being SML correlation-sensitive. For across-small-worlds comparisons, DM can rely on the certainty equivalents of the bets assessed within their respective small worlds.

Correlation Ambiguity. In Epstein and Halevy's (2019) thought experiment concerning the two-urn Ellsberg (1961) problem, there are two urns numbered 1 and 2 each containing two balls which are either red or black without additional information about the compositions. One ball is drawn from each urn simultaneously, so that the set of possible outcomes is $\mathcal{S} = \{R_1B_2, B_1R_2, R_1R_2, B_1B_2\}$. Before the balls are drawn, an individual chooses between a pair of bets on the colors of the two balls, each giving the same prize x: for instance, bet *Same* corresponds to the event of balls being in the same color, $\{R_1R_2, B_1B_2\}$, and bet *Diff* corresponds to $\{R_1B_2, B_1R_2\}$; bet R_1 is based on the event of urn 1 draw being red, $\{R_1R_2, R_1B_2\}$, and B_1 is based on the opposite event of $\{B_1R_2, B_1B_2\}$. Consider the choice pattern: $R_1 \succ$ *Same*, and $B_1 \succ$ *Diff*. As observed in Epstein and Halevy (2019), there

[33] See Abdellaoui et al. (2011), Armantier and Treich (2016), Chew et al. (2023).

is no probability P on $\mathcal{S}$ such that $P(R_1) > P(\mathit{Same})$, $P(B_1) > P(\mathit{Diff})$, while maintaining $P(R_1) + P(B_1) = 1 = P(\mathit{Same}) + P(\mathit{Diff})$. They offer an account of this choice behavior in terms of DM being more averse to betting on correlation ambiguity (*Same* and *Diff*) than within-urn ambiguity bets R_1 and B_1.

Correlation Source Preference. We first observe that the above choice behavior is not compatible with CEU nor CWU under PS in the Savagean grand world setting. Suppose DM has CEU ϕ and attaches the following joint probability to the states. Then, preferences $R_1 \succ \mathit{Same}$ is equivalent to $\mathbb{E}_\pi \phi \geqslant 0$ and $\mathbb{E}_{\pi^\top} \phi < 0$ for π in Table 7, i.e., $p_{rb}\phi(x,0) + p_{bb}\phi(0,x) \geqslant 0$ and $p_{bb}\phi(x,0) + p_{rb}\phi(0,x) < 0$. Likewise $B_1 \succ \mathit{Diff}$ means $p_{bb}\phi(x,0) + p_{rb}\phi(0,x) \geqslant 0$ and $p_{rb}\phi(x,0) + p_{bb}\phi(0,x) < 0$, which is inconsistent with the previous two inequalities. The argument for CWU is similar.

	R_2	B_2
R_1	p_{rr}	p_{rb}
B_1	p_{br}	p_{bb}

	x	0
x	$\pi(x,x) = p_{rr}$	$\pi(x,0) = p_{rb}$
0	$\pi(0,x) = p_{bb}$	$\pi(0,0) = p_{br}$

Table 7: Joint probability π for comparing bets R_1 and *Same*.

Observe that the events R_1 and B_1 in the single-source small world and the events *Same* and *Diff* in the mixed-source small world are respectively exchangeable. We apply small worlds PS to account for the presence of incremental aversion towards correlation ambiguity via a *source-dependent CEU* representation using different kernels ϕ^1 and ϕ^2 for the two small worlds. In the single-source small world, the CE for R_1, denoted c^{R_1}, is based on $\frac{1}{2}\phi^1(x, c^{R_1}) + \frac{1}{2}\phi^1(0, c^{R_1}) = 0$, reducing to $\phi^1(x, c^{R_1}) = \phi^1(c^{R_1}, 0)$, and correspondingly $\phi^1(x, c^{B_1}) = \phi^1(c^{B_1}, 0)$. This also applies to c^{Same} and c^{Diff}, i.e., $\phi^2(x,c) = \phi^2(c,0)$, given the exchangeability between *Same* and *Diff*. To account for Epstein and Halevy's (2019) finding of incremental aversion to correlation ambiguity, we can identify conditions on ϕ^1 and ϕ^2 such that $c^{R_1} > c^{Same}$ and $c^{B_1} > c^{Diff}$. One possibility is to apply small worlds TPS, e.g., assume that DM is SEU in both small worlds with a more concave utility function for mixed-source than for same-source.

In the international finance literature, the equity home bias puzzle (French and Poterba, 1991) and its variants based solely on domestic US equity markets (Coval and Moskowitz, 1999; Huberman, 2001) provide indirect empirical support for the incidence of source preference. Small worlds PS may also apply to the differentiation between investment and insurance discussed in Armantier et al. (2023), exemplifying a more correlation-based reference-dependence across gain- versus loss-oriented choice situations. Labor markets, often involving matching, could provide a rich source of small worlds given the inherent heterogeneity in how diverse individuals relate to jobs and other occupational arrangements.

7 Discussions

7.1 Inertia

Recall that subsection 4.4 links to the idea of *inertia* which relates to the phenomenon of status quo bias. Here, we discuss another approach to modeling inertia through a non-CEU correlation utility based on Dubra et al. (2004). Consider the following correlation multi-linear representation based on a weakening of correlation independence.[34]

Definition 13. Π *admits a correlation multi-utility representation if there exists a closed family* $\{\phi : \phi \in \Phi\}$, *such that* $\pi \in \Pi \iff \inf_{\phi\in\Phi} \mathbb{E}_\pi \phi \geqslant 0$.

The family Φ corresponds to DM's *indecisiveness* in the sense that she compares each correlated pair with multiple ϕ, and the determining ϕ may be different at different π. This intuition is reflected in the following convexity axiom.

Convexity. For any $\pi, \pi' \in \Pi$ and $\alpha \in (0, 1)$, $\alpha\pi + (1-\alpha)\pi' \in \Pi$.

Observe that a correlation preference is convex if its utility representation is quasi-concave. Applying a supporting hyperplane argument immediately yields the following characterization of correlation multi-utility in terms of convexity.

Proposition 9. Π *satisfies convexity if and only if it admits a correlation multi-utility representation.*

A natural family of Φ corresponds to the family of supporting hyperplanes underpinning the convex Π. Compared with our asymmetric CEU where inertia also plays a role, here the modeling of inertia is through a non-CEU representation. As another example, if U satisfies the two spherical properties in Chambers and Echenique (2020), then $U(\pi)$, given by $Q(\pi) + \mathbb{E}_\pi \phi$ where Q is a *positive semi-*

[34]While we could consider a similar extension based on weakening correlation betweenness, we do not pursue this direction here.

definite quadratic, is quasiconcave.[35]

Definition 14. Π *admits an expected multi-utility representation if there exists a closed family of utilities* $u \in \mathcal{U}$ *, such that* $\pi \in \Pi \iff \inf_{u \in \mathcal{U}}(\mathbb{E}_{\pi_1} u - \mathbb{E}_{\pi_2} u) \geqslant 0$.

If we require $\phi(x, x) = 0$ as for CEU, then correlation insensitivity leads to expected multi-utility through a similar OT argument as in Theorem 2.

Proposition 10. *For* Π *that admits a correlation multi-utility representation with* $\phi(x, x) = 0$ $\forall \phi \in \Phi$*, it is correlation-insensitive if and only if it admits an expected multi-utility representation.*

As Cerreia-Vioglio et al. (2015) extend expected multi-utility to a complete and transitive *cautious expected utility*, one may consider extending the *negative certainty independence* axiom introduced by (Dillenberger, 2010)—if $p \succeq \delta_x$, then $\alpha p + (1-\alpha)q \succeq \alpha\delta_x + (1-\alpha)q$, for any other lottery q—into a correlation setting: $\bar{\pi}, \pi' \in \Pi$ implies $\alpha\bar{\pi} + (1-\alpha)\pi' \in \Pi$ for any $\bar{\pi}$ with degenerate column marginal and any π'. It follows that such partially degenerate $\bar{\pi}$'s are part of the domain of Π exhibiting correlation independence. On the flip side, one candidate for a counterpart of cautious expected utility could likewise be obtained by restricting the family Φ in a correlation multi-utility representation to those consistent with the preference on the correlation independence domain. Specifically, such a preference is correlation-complete if $U_\Phi(\pi) > 0 \iff U_\Phi(\pi^{\mathrm{T}}) < 0$.

More recently, Hara et al. (2019) has axiomatized a *coalitional expected multi-utility* with the independence axiom, but not transitivity nor completeness. One observation is that they use the three-lottery version of independence, leading to a non-convex cone $C = Cone\{p - q : p \succeq q\}$ and subsequently a family of utility sets. Our *correlation multi-utility* extends the above idea to correlation preference where convexity is restored. This is akin to the benefit from adapting Samuelson's (1952) four-lottery independence axiom to arrive at our correlation independence axiom.

Another form of correlation PS which can exhibit a form of inertia is offered in Chew, Wang, and Zhong's (2024) attention theory (AT) which models an attention-induced correlation preference for binary choice but does not belong to CEU. The

[35]Like our CEU, here we can view $U(\pi) - U(\pi')$ as "second-order" utility difference. For intuition, compared with our CEU model, the extra quadratic term Q represents DM's deviation from a linear utility preference. Since Q is positive semi-definite, it always "favors" the row marginal π_1; if we view row marginal as the status quo, then this captures DM's inertia.

AT specification takes the form of a WU representation,[36] which bears some resemblance to CBU but does not belong to that class. This leaves open the question of identifying its characteristic property and the corresponding axiomatization.

7.2 Rank Dependence

In contrast with EU, Yaari (1987) offers a *dual utility* approach by axiomatizing a rank-dependent utility model with a linear utility function defined on the unit interval. For a marginal lottery p, associate it with a random variable t and let $G_p : [0,1] \to [0,1]$ be its *de-cumulative* distribution function (DDF) with $G_p(x) = Prob\ \{t > x\}$. A DDF G_p is decreasing and right-continuous with $G_p(1) = 0$. It serves as a role of ranking: the larger outcome x is ranked "higher" by a smaller DDF value. Define the generalized inverse G_p^{-1} as $G_p^{-1}(u) = \min\{x : G_p(x) \leqslant u\}$. Such a quantile function reflects the idea of re-ranking of the outcomes by associating each rank u with the corresponding outcome $x = G_p^{-1}(u)$.

We develop here a correlation extension of dual utility. By Sklar's theorem (see, e.g., Nelsen (2006)), any bivariate joint probability distribution F on $X \times X$ can be decomposed into $F(x,y) = C(F_1(x), F_2(y))$, where F_1, F_2 are the marginals and $C : [0,1] \to [0,1]$ is a *copula* representing the correlation between the two random variables associated with F. Let $\mathcal{C}$ be the set of all copulas. For a joint density π, we can thus identify it with a triplet $(C_\pi, G_{\pi_1}, G_{\pi_2})$. In the following definition, ψ resembles the bivariate utility ϕ in CEU.

Definition 15 (Correlation Dual Utility). *The family $\{\psi^C\}_{C \in \mathcal{C}}$ is a correlation dual utility representation if each ψ^C is continuous and increasing, and for all π, $\pi \in \Pi \iff \int_{[0,1]^2} \psi^{C_\pi}(G_{\pi_1}(x), G_{\pi_2}(y))dxdy \geqslant 0$.*

Intuitively, a CDU decision maker first singles out the correlation C_π, and applies the corresponding utility function ψ^{C_π} to evaluate the two marginal quantiles/rankings $G_{\pi_1}(\cdot)$, $G_{\pi_2}(\cdot)$ at each outcome level x. If DM is correlation-insensitive, the above turns into another SSB utility form: $\pi \in \Pi \iff \int_{[0,1]^2} \psi(p,q) dG_{\pi_1}^{-1}(p) dG_{\pi_2}^{-1}(q) \geqslant 0$, where the bilinearity is now in terms of marginal ranks (quantiles). Imposing transitivity, one may arrive at a *rank-dependent weighted utility* à la Chew and Epstein (1989).

[36] Under AT, there exist utility function $v : X \to \mathbb{R}$ and bivariate weight function $\alpha : X \times X \to \mathbb{R}_+$ such that $\pi \in \Pi \iff \mathbb{E}_\pi \phi_\pi \geqslant 0$, where $\phi_\pi = \mathbb{E}_\pi[\alpha(y,x)]\alpha(x,y)v(x) - \mathbb{E}_\pi[\alpha(x,y)]\alpha(y,x)v(y)$.

When each ψ^C is separable[37], i.e., $\psi^C(r,s) = f^C(r) - f^C(s)$, the correlation dual utility preference is represented by $\pi \in \Pi \iff \int_{[0,1]} f^{C_\pi}(G_{\pi_1}(x))dx \geqslant \int_{[0,1]} f^{C_\pi}(G_{\pi_2}(y))dy$, which satisfies the following *correlation dual independence*:

Property (Correlation Dual Independence)**.** *If $\pi, \pi' \in \Pi$ and $C_\pi = C_{\pi'} = C$, then for any $\alpha \in (0,1)$, the joint density π^α corresponding to the triplet $(C, G_1^\alpha, G_2^\alpha)$ is also in Π, where $(G_i^\alpha)^{-1} = \alpha G_{\pi_i}^{-1} + (1-\alpha)G_{\pi_i'}^{-1}$ for $i = 1,2$.*

Fixing copula C, it is natural to examine *transitivity* for marginals under this specific correlation given the above separability in utility. When DM is correlation-insensitive, the axiom is simplified into Yaari's original *dual independence* axiom, which he shows to be equivalent to comonotonic independence of random variables (Schmeidler, 1989). The representation reduces correspondingly to his *dual utility* representation, i.e., there exists f such that $f^C = f$ for every C belonging to $\mathcal{C}$.

Rank dependence has also appeared in the literature on social and risk preference.[38] For example, Chew and Sagi (2012) axiomatize a social welfare function in an inequality measurement setting which has a rank-dependent form and can take into account correlations across individuals, incorporating a preference for ex-ante versus ex-post fairness. If we consider the income distribution or wealth allocation as a consequence of policy choice and economic state with uncertainty, then the different resulting outcomes must be correlated and the society or the social planner is effectively facing a correlated choice problem.[39]

7.3 Social Choice and Beyond

The Condorcet paradox, considered "*the root cause for central impossibility theorems in social choice theory such as Arrow's Theorem or the Gibbard-Satterthwaite Theorem*" (Brandt, 2017), has motivated the application of SSB utility to the study of social choice theory without transitivity. Brandl and Brandt (2020) investigate the maximal SSB preference domain in which there exists an *Arrovian* social welfare function that aggregates individual preferences into societal preference. They

[37] We formulate a specialization of ψ^C in Online Appendix D that reduces to Yaari's formula given correlation insensitivity through applying Kantorovich duality as in CEU arguments.

[38] See e.g., Sen (1973); Fleurbaey (2010); Grant et al. (2012); Saito (2013); Miao and Zhong (2018).

[39] Recently, Zhou (2024) studies a social preference model over rankings that allows for nontransitivity. She introduces an idea of blame and gratitude, derived from RT, which naturally involves correlation in the presence of randomization over rankings.

further show that the Arrovian social welfare function on that domain must be *affine utilitarian*, thus recovering Harsanyi's aggregation theorem (Harsanyi, 1955) though with an additional axiom of independence of irrelevant alternatives. Prior to that, Fishburn and Gehrlein (1987) and Turunen-Red and Weymark (1999) also attempt to restore Harsanyi's aggregation theorem in the domain of SSB preferences; however, as noted by Brandl and Brandt (2020), the results suggest that aggregating SSB utility functions is "fundamentally different" from performing this exercise under EU.

Notwithstanding these relatively recent attempts at weakening transitivity in aggregating individual preferences, correlation sensitivity has not been incorporated into the binary choice framework in this literature. Turunen-Red and Weymark (1999) *convexify* the domain of independent lottery pairs into the space of *all* joint densities between marginals, but attach a cardinal interpretation to the utility on joint densities so as to apply the result of De Meyer and Mongin (1995) who strengthen Harsanyi's result by virtue of convex structure. It will also be interesting to revisit classical aggregation results in social choice theory with correlation preferences beyond Brandl and Brandt (2020) by bringing in correlation sensitivity in the expanded domain of correlated lotteries. For example, in the symmetric CEU domain, we expect Harsanyi's aggregation theorem to hold via applying duality arguments but similar techniques may not apply to asymmetric CEU or CWU.

Beyond social choice, it would be interesting to incorporate correlation and social preference into the study of game theory and mechanism design. Observe that social preference may emerge from the ubiquitous presence of correlatedness in social networks (Bourles et al., 2017). Motivated by context-dependent choice, including status quo and framing effects, De Clippel (2014) extends the classical implementation theory to choice correspondences not necessarily consistent with maximizing a complete and transitive preference.

Our discussion of framing effects and consequentialism in subsection 6.1 points to the influence of choice perception in binary choice. After all, Savage's transformation of the original Allais problems into what he considers to be equivalent state-by-state representations appear to have taken place solely in his mind. Beyond correlation sensitivity, we may explore the role of choice perception in future research involving individual as well as social choice.

References

Abdellaoui, M., A. Baillon, L. Placido, and P. P. Wakker (2011). The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review 101*(2), 695–723.

Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica 21*(4), 503–546.

Anderson, A. and L. Smith (2024). The comparative statics of sorting. *American Economic Review 114*(3), 709–751.

Armantier, O., J. Foncel, and N. Treich (2023). Insurance and portfolio decisions: Two sides of the same coin? *Journal of Financial Economics 148*(3), 201–219.

Armantier, O. and N. Treich (2016). The rich domain of risk. *Management Science 62*(7), 1954–1969.

Arrow, K. J. (1951). *Social Choice and Individual Values*, Volume No. 12 of *Cowles Commission Monograph*. John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London.

Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics 1*(1), 67–96.

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica 55*(1), 1–18.

Ball, I. (2023). A unified theorem of the alternative. arXiv preprint arXiv:2303.07471.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research 30*(5), 961–981.

Bewley, T. F. (1986). Knightian decision theory: Part 1. Cowles Foundation Discussion Papers.

Bikhchandani, S. and U. Segal (2011). Transitive regret. *Theoretical Economics 6*(1), 95–108.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics 127*(3), 1243–1285.

Bordalo, P., N. Gennaioli, and A. Shleifer (2022). Salience. *Annual Review of Economics 14*, 521–544.

Bourles, R., Y. Bramoulle, and E. Perez-Richet (2017). Altruism in networks. *Econometrica 85*(2), 675–689.

Brandl, F. and F. Brandt (2020). Arrovian aggregation of convex preferences. *Econometrica 88*(2), 799–844.

Brandl, F., F. Brandt, and H. G. Seedig (2016). Consistent probabilistic social choice. *Econometrica 84*(5), 1839–1880.

Brandt, F. (2017). Trends in computational social choice. In *Trends in computational social choice*. AI Access.

Bruhin, A., M. Manai, and L. Santos-Pinto (2022). Risk and rationality: The relative importance of probability weighting and choice set dependence. *Journal of Risk and Uncertainty 65*(2), 139–184.

Cerreia-Vioglio, S., D. Dillenberger, and P. Ortoleva (2015). Cautious expected utility and the certainty effect. *Econometrica 83*(2), 693–728.

Cerreia-Vioglio, S., D. Dillenberger, and P. Ortoleva (2024). Caution and reference effects. Working paper.

Chambers, C. P. and F. Echenique (2020). Spherical preferences. *Journal of Economic Theory 189*, 105086.

Chambers, C. P., Y. Masatlioglu, and C. Raymond (2024). Coherent distorted beliefs. Working paper.

Chang, Y. and S. L. Liu (2024). Counterexamples to "Transitive Regret". Working paper.

Chew, S. H. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica 51*(4), 1065–1092.

Chew, S. H. (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research 19*(1), 273–298.

Chew, S. H. and L. G. Epstein (1989). A unifying approach to axiomatic non-expected utility theories. *Journal of Economic Theory 49*(2), 207–240.

Chew, S. H., L. G. Epstein, and U. Segal (1991). Mixture symmetry and quadratic utility. *Econometrica 59*(1), 139–163.

Chew, S. H., L. G. Epstein, and U. Segal (1994). The projective independence axiom. *Economic Theory 4*(2), 189–215.

Chew, S. H., B. Miao, and S. Zhong (2023). Ellsberg meets Keynes at an urn. *Quantitative Economics 14*(3), 1133–1162.

Chew, S. H. and J. S. Sagi (2006). Event exchangeability: Probabilistic sophistication without continuity or monotonicity. *Econometrica 74*(3), 771–786.

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

Chew, S. H. and J. S. Sagi (2012). An inequality measure for stochastic allocations. *Journal of Economic Theory 147*(4), 1517–1544.

Chew, S. H. and G. Tan (2005). The market for sweepstakes. *The Review of Economic Studies 72*(4), 1009–1029.

Chew, S. H., W. Wang, and S. Zhong (2024). Attention theory. Working paper.

Condorcet, M. d. (1785). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.

Coval, J. D. and T. J. Moskowitz (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance 54*(6), 2045–2073.

De Clippel, G. (2014). Behavioral implementation. *American Economic Review 104*(10), 2975–3002.

De Meyer, B. and P. Mongin (1995). A note on affine aggregation. *Economics Letters 47*(2), 177–183.

Debreu, G. (1964). Continuity properties of paretian utility. *International Economic Review 5*(3), 285–293.

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic theory 40*(2), 304–318.

Dillenberger, D. (2010). Preferences for one-shot resolution of uncertainty and Allais-type behavior. *Econometrica 78*(6), 1973–2004.

Dubra, J., F. Maccheroni, and E. A. Ok (2004). Expected utility theory without the completeness axiom. *Journal of Economic Theory 115*(1), 118–133.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics 75*(4), 643–669.

Epstein, L. G. and Y. Halevy (2019). Ambiguous correlation. *The Review of Economic Studies 86*(2), 668–693.

Fishburn, P. C. (1975). Separation theorems and expected utilities. *Journal of Economic Theory 11*(1), 16–34.

Fishburn, P. C. (1982). Nontransitive measurable utility. *Journal of Mathematical Psychology 26*(1), 31–67.

Fishburn, P. C. (1983). Transitive measurable utility. *Journal of Economic Theory 31*(2), 293–317.

Fishburn, P. C. (1984). Probabilistic social choice based on simple voting comparisons. *The Review of Economic Studies 51*(4), 683–692.

Fishburn, P. C. (1988). *Nonlinear Preference and Utility Theory.* The Johns Hopkins University Press.

Fishburn, P. C. (1989). Non-transitive measurable utility for decision under uncertainty. *Journal of Mathematical Economics 18*(2), 187–207.

Fishburn, P. C. and W. V. Gehrlein (1987). Aggregation theory for ssb utility functionals. *Journal of Economic Theory 42*(2), 352–369.

Fleurbaey, M. (2010). Assessing risky social situations. *Journal of Political Economy 118*(4), 649–680.

French, K. R. and J. M. Poterba (1991). Investor diversification and international equity markets. *The American Economic Review 81*(2), 222–226.

Frydman, C. and M. Mormann (2018). The role of salience in choice under risk: An experimental investigation. Available at SSRN 2778822.

Grant, S. (1995). Subjective probability without monotonicity: or how Machina's mom may also be probabilistically sophisticated. *Econometrica 63*(1), 159–189.

Grant, S., A. Kajii, B. Polak, and Z. Safra (2012). Equally-distributed equivalent utility, ex post egalitarianism and utilitarianism. *Journal of Economic Theory 147*(4), 1545–1571.

Hara, K., E. A. Ok, and G. Riella (2019). Coalitional expected multi-utility theory. *Econometrica 87*(3), 933–980.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy 63*(4), 309–321.

Hart, S. and D. Schmeidler (1989). Existence of correlated equilibria. *Mathematics of Operations Research 14*(1), 18–25.

Huberman, G. (2001). Familiarity breeds investment. *The Review of Financial Studies 14*(3), 659–680.

Humphrey, S. J. and N.-Y. Kruse (2023). Who accepts Savage's axiom now? *Theory and Decision 96*, 1–17.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–292.

Kantorovich, L. V. (1942). On the translocation of masses. *Dokl. Akad. Nauk. USSR (NS) 37*, 199–201.

Kőszegi, B. and M. Rabin (2007). Reference-dependent risk attitudes. *American Economic Review 97*(4), 1047–1073.

Kreweras, G. (1965). Aggregation of preference orderings. In *Mathematics and Social Sciences I: Proceedings of the seminars of Menthon-Saint-Bernard, France (1–27 July 1960) and of Gösing, Austria (3–27 July 1962)*, pp. 73–79.

Laffond, G., J.-F. Laslier, and M. Le Breton (1993). The bipartisan set of a tournament game. *Games and Economic Behavior 5*(1), 182–201.

Lanzani, G. (2022). Correlation made simple: Applications to salience and regret theory. *The Quarterly Journal of Economics 137*(2), 959–987.

Loewenfeld, M. and J. Zheng (2023). Uncovering correlation sensitivity in decision making under risk. Working paper.

Loewenfeld, M. and J. Zheng (2024). Salience or event-splitting? An experimental investigation of correlation sensitivity in risk-taking. *Journal of the Economic Science Association*, 1–21.

Loomes, G. and R. Sugden (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal 92*(368), 805–824.

Machina, M. J. and D. Schmeidler (1992). A more robust definition of subjective probability. *Econometrica 60*(4), 745–780.

Mas-Colell, A., M. Whinston, and J. Green (1995). *Microeconomic Theory*. Oxford University Press.

Masatlioglu, Y. and E. A. Ok (2005). Rational choice with status quo bias. *Journal of economic theory 121*(1), 1–29.

Miao, B. and S. Zhong (2018). Probabilistic social preference: how Machina's mom randomizes her choice. *Economic Theory 65*, 1–24.

Nakamura, Y. (1990). Bilinear utility and a threshold structure for nontransitive preferences. *Mathematical Social Sciences 19*(1), 1–21.

Nelsen, R. B. (2006). *An introduction to copulas*. Springer.

Ostermair, C. (2022). An experimental investigation of the Allais paradox with subjective probabilities and correlated outcomes. *Journal of Economic Psychology 93*, 102553.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization 3*(4), 323–343.

Quiggin, J. (1994). Regret theory with general choice sets. *Journal of Risk and Uncertainty 8*, 153–165.

Rivest, R. L. and E. Shen (2010). An optimal single-winner preferential voting system based on game theory. In *Proc. of 3rd International Workshop on Computational Social Choice*, pp. 399–410. Citeseer.

Sagi, J. S. (2006). Anchored preference relations. *Journal of Economic Theory 130*(1), 283–295.

Saito, K. (2013). Social preferences under risk: Equality of opportunity versus equality of outcome. *American Economic Review 103*(7), 3084–3101.

Samuelson, P. A. (1952). Probability, utility, and the independence axiom. *Econometrica 20*(4), 670–678.

Savage, L. J. (1954). *The foundations of statistics*. Wiley.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica 57*(3), 571–587.

Segal, U. (1992). Additively separable representations on non-convex sets. *Journal of Economic Theory 56*(1), 89–99.
Sen, A. (1973). *On economic inequality.* Clarendon Press.
Sernesi, E. (2019). *Linear algebra: a geometric approach.* Routledge.
Starmer, C. (1992). Testing new theories of choice under uncertainty using the common consequence effect. *The Review of Economic Studies 59*(4), 813–830.
Turunen-Red, A. H. and J. A. Weymark (1999). Linear aggregation of SSB utility functionals. *Theory and Decision 46*, 281–294.
Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty 5*, 297–323.
von Neumann, J. and O. Morgenstern (1944). *Theory of games and economic behavior.* Princeton University Press.
Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica 55*(1), 95–115.
Zhou, Z. (2024). Ranking blame. *Economic Theory 78*, 403–441.

Appendices

For notational ease, we write $\phi \cdot \pi$ for $\mathbb{E}_\pi \phi$ when convenient.

A Proof of Theorem 2 in Section 4

Proof. (iii) $\Rightarrow$ (i) and (ii) is easy to see.

(i) $\Rightarrow$ (iii): Let $\mathcal{D} = \{(p,q) : p \succeq^* q\}$. For simplicity, we fix a finite subset $X' \subset X$ with $|X'| = n$, and restrict attention to $\Pi|_{X' \times X'}$ and $\Delta(X' \times X')$. It is sufficient to prove the result on this restricted set as an extension to the whole X is guaranteed by uniqueness of ϕ.

Clearly the feasible set of (u,v) is convex. By standard cyclic monotonicity argument in Kantorovich duality.[40] we can further restrict (u,v) in the dual constraint to a bounded set $UV \subset \mathbb{R}^{2|X'|}$. Now, $\mathcal{D}$ and UV are both convex and compact, and $f((p,q),(u,v)) = u \cdot p - v \cdot q$ is linear in both of them. By the standard minimax theorem in the Euclidean space, we have the following

$$\min_{(p,q)\in\mathcal{D}} \max_{(u,v)\in UV} u \cdot p - v \cdot q = \max_{(u,v)\in UV} \min_{(p,q)\in\mathcal{D}} u \cdot p - v \cdot q.$$

[40]For instance, by Kantorovich duality theorem, the dual constraint holds as an equality almost surely for any optimal dual solution. Since the support is finite, we can always normalize the optimal dual solutions to a bounded set based on the given ϕ.

The LHS is non-negative by definition, and hence so is the RHS. This means $\exists(\hat{u}, \hat{v})$ such that $\hat{u}(x) - \hat{v}(y) \leqslant \phi(x, y)$ and $\hat{u} \cdot p - \hat{v} \cdot q \geqslant 0$ holds $\forall(p, q) \in \mathcal{D}$.

By setting $y = x$ in the dual constraints, we have $\hat{u}(x) \leqslant \hat{v}(x)$ $\forall x \in X'$. But at the same time, $\delta_{(x,x)} \in \Pi$ and therefore, $\hat{u}(x) - \hat{v}(x) \geqslant 0$, so that it must be $\hat{v}(x) = \hat{u}(x)$ $\forall x \in X'$. Let $u^*(x) = \hat{u}(x)$.

Now, if p, q are such that $u^* \cdot p \geqslant u^* \cdot q$, then $\max_{(u,v) \in UV} u \cdot p - v \cdot q \geqslant 0$. By Kantorovich duality, we know that for all $\pi \in \Gamma(p, q)$, $\phi \cdot \pi \geqslant 0$, so $p \succeq^* q$. Therefore, $p \succeq q \iff u^* \cdot p \geqslant u^* \cdot q$. It then suffices to set $\phi^{EU}(x, y) = u^*(x) - u^*(y)$ to represent Π.

(ii) $\Rightarrow$ (iii): Suppose ϕ is not a EU representation. Then, ϕ is not modular, meaning that there exist $x_1 < x_2, y_1 < y_2$ such that the cross-difference $s(x_1, x_2, y_1, y_2) = \phi(x_1, y_1) + \phi(x_2, y_2) - \phi(x_1, y_2) - \phi(x_2, y_1)$ is not zero. (Otherwise, for any fixed pair (x_i, x_j), $\phi(x_i, y) - \phi(x_j, y)$ is constant over y, so we can write $\phi(x, y)$ as $v(y) - u(x)$, and then $\phi(x, x) = 0$ implies $u = v$.)

Pick a $p \in \Delta X$ with x_1, x_2, y_1, y_2 in its support, and any $\pi \in \Gamma(p, p)$. If $\sum \phi(x, y)\pi(x, y) > 0$, then its transpose $\pi^{\mathrm{T}} \in \Gamma(p, p)$ while $\sum \phi(x, y)\pi^{\mathrm{T}}(x, y) < 0$ due to skew-symmetry. This means that (p, p) violates SML correlation insensitivity. If it happens that $\sum \phi(x, y)\pi(x, y) = 0$, then we can perturb π on the rectangle of the four points (x_i, y_j), $i, j = 1, 2$, so that the perturbed π' is still in $\Gamma(p, p)$ while $\sum \phi(x, y)\pi'(x, y) > 0$. For example, suppose $s(x_1, x_2, y_1, y_2) = \phi(x_1, y_1) + \phi(x_2, y_2) - \phi(x_1, y_2) - \phi(x_2, y_1) > 0$. Let $\pi' \in \Gamma(p, p)$ be the correlation identical to π on all points but the four above, and $\pi'(x_i, y_i) = \pi(x_i, y_i) + \epsilon$, $\pi'(x_i, y_{-i}) = \pi(x_i, y_{-i}) - \epsilon$, $i = 1, 2$. Then $\phi \cdot \pi' = \phi \cdot \pi + \epsilon s(x_1, x_2, y_1, y_2) = \epsilon s(x_1, x_2, y_1, y_2) > 0$. □

B Proofs for Section 5

B.1 Proof of Theorem 3 (CBU) in Subsection 5.1

Proof. For each $p \in \Delta X$, define $\Pi|_p = \{\pi \in \Pi : \pi_1 = p\}$, and $\check{\Pi}|_p, \hat{\Pi}|_p, \tilde{\Pi}|_p$ correspondingly. We assume that both $\check{\Pi}|_p$ and $\hat{\Pi}|_p$ are non-empty as the rest of the cases are straightforward. Using a similar separation argument as in Theorem 1, we can find a $\phi_p \not\equiv 0$ so that for $\pi \in \Delta_p$, $\pi \in \Pi \Rightarrow \phi_p \cdot \pi \geqslant 0$, and $\pi \in \check{\Pi} \Rightarrow \phi_p \cdot \pi \leqslant 0$. Equivalently, $\phi_p \cdot \pi > 0 \Rightarrow \pi \in \Pi$, and $\phi_p \cdot \pi < 0 \Rightarrow \pi \in \check{\Pi}$.

Now, as in Corollary 1, $\pi \in \tilde{\Pi} \Rightarrow \phi_p \cdot \pi = 0$ as otherwise, we can find some $\pi' \in \check{\Pi}|_p$

and a convex combination π'' of π and π', which should be in $\check{\Pi}$ by betweenness but $\phi_p \cdot \pi'' > 0$. So $\phi_p \cdot \pi > 0 \Rightarrow \pi \in \hat{\Pi}$.

Since $\phi_p \not\equiv 0$, we can assume without losing generality that $\exists \pi^*$ such that $\phi_p \cdot \pi^* > 0$. Then $\pi \in \check{\Pi} \Rightarrow \phi_p \cdot \pi < 0$. (Otherwise, by continuity, there is a combination π^{**} of π and π^* in $\tilde{\Pi}$, but this contradicts $\phi_p \cdot \pi^{**} > 0$.) This means $\pi \in \check{\Pi} \iff \phi_p \cdot \pi < 0$ and hence, $\pi \in \Pi \iff \phi_p \cdot \pi \geqslant 0$.

Finally, $\phi_p \cdot \pi = 0 \Rightarrow \pi \in \tilde{\Pi}$ as otherwise, $\pi \in \hat{\Pi}$, so continuity implies the existence of a combination with some $\pi^c \in \check{\Pi}$ that is in $\tilde{\Pi}$, but this clearly contradicts $\phi_p \cdot (\alpha\pi + (1-\alpha)\pi^c) < 0$. This gives $\phi_p \cdot \pi = 0 \iff \pi \in \tilde{\Pi}$ and completes the proof. □

B.2 Proofs for Subsection 5.2

Proof of Proposition 2 (CBU and CLU). By applying a duality argument similar to the proof of Theorem 2, we have for every p,

$$0 \leqslant \min_{q:\ p \succeq q} \max_{(u,v) \in UV} \sum_x u(x)p(x) - \sum_y v(y)q(y)$$
$$= \max_{(u,v) \in UV} \min_{q:\ p \succeq q} \sum_x u(x)p(x) - \sum_y v(y)q(y),$$

and hence the existence of dual $\tilde{u}_p, \tilde{v}_p \in \mathbb{R}^X$ such that $\tilde{u}_p \cdot p - \tilde{v}_p \cdot q \geqslant 0 \iff p \succeq q$. Let $v_p(\cdot) = -\tilde{u}_p \cdot p + \tilde{v}_p(\cdot)$, and assuming completeness on $\succeq$, then $p \succ$ (resp. $\sim$) if and only if $v_p \cdot q < 0$ (resp. $=$). □

Proof of Proposition 3 (CLU and BU). Under transitivity, for any $p \sim p'$ and $q \in \Delta X$, $v_p \cdot q < 0$ (resp. $=, >$) $\iff$ $v_{p'} \cdot q < 0$ (resp. $=, >$). Consequently, for any indifference set $H_\alpha \subset \Delta X$ and $\alpha \in [0,1]$ such that $H_\alpha \bigcap [\bar{r}, \underline{r}] := r_\alpha = \alpha\bar{r} + (1-\alpha)\underline{r}$, let $V(p) := \alpha$ and $u(x,\alpha) := v_{r_\alpha}(x) + \alpha$ for all $p \in H_\alpha$. By continuity and transitivity, such u and V are both well defined for $X \times [0,1]$ and ΔX respectively. Then, u and V satisfy the betweenness utility equation in Dekel (1986): $\sum_x u(x, V(p))p(x) = V(p)$.

Now, for $p \sim r_\alpha$, $\sum_x u(x,\beta)p(x) < \beta$ (resp. $=, <$) if $\beta > \alpha$ (resp. $=, <$). This is because, by construction, $\sum_x u(x,\beta)p(x) = \sum_x v_{r_\beta}(x)p(x) + \sum_x \alpha p(x) = v_{r_\beta} \cdot p + \alpha$ where $H_\beta \bigcap [\bar{r}, \underline{r}] = r_\beta$. This shows the uniqueness of solution $V(p)$ for a given p. One can verify the uniqueness of u up to a positive affine transformation, and monotonicity of $V(\lambda q + (1-\lambda)q')$ in λ. Thus, we effectively construct the implicit utility V from the conditional linear representation v_p. The opposite direction is immediate. $\square$

Proof of Proposition 4 (Generalized Skew Symmetry). The statement is trivially true for $(p,q) \in \mathcal{D}^{Ind}$ as $\phi_p \cdot \pi = \phi_q \cdot \pi^{\mathrm{T}} = 0$. We proceed in two steps with the first one being a lemma used frequently in the following proofs.

Lemma 1. *For a fixed pair $(p,q) \in \mathcal{D}^S$, there exists $b > 0$ and c, d such that $\phi_p(x,y) + b\phi_q(y,x) = c(x) + d(y)$ and $c \cdot p + d \cdot q = 0$.*

Proof. For a fixed pair $(p,q) \in \mathcal{D}^S$, observe that completeness $\iff$ the following system: $\phi_p \cdot \pi > 0$, $\phi_q \cdot \pi^{\mathrm{T}} \geqslant 0$, $\pi(x,y) \geqslant 0$, $\sum_y \pi(x,y) = p(x)$, $\sum_x \pi(x,y) = q(y)$ being infeasible. By a theorem of the alternative proved by Ball (2023), this is equivalent to the feasibility of its alternative:

$$\begin{aligned} -a\phi_p(x,y) - b\phi_q(y,x) + c(x) + d(y) &\geqslant 0, \\ c \cdot p + d \cdot q \leqslant 0,\ a, b &\geqslant 0, \\ c \cdot p + d \cdot q < 0,\ \text{or } a &> 0. \end{aligned}$$

For $(p,q) \in \mathcal{D}^S$, since the set of π with full support is dense, we can find a full-support $\pi \in \Gamma(p,q)$ such that $\phi_p \cdot \pi = 0 = \phi_q \cdot \pi^{\mathrm{T}}$. Multiply by $\pi(x,y)$ the first line of the alternative and sum across (x,y), $c \cdot p + d \cdot q \geqslant 0$. Given the second line in the alternative, this means $c \cdot p + d \cdot q = 0$. Since π has full support and the sum is zero, it follows that $-a\phi_p(x,y) - b\phi_q(y,x) + c(x) + d(y) = 0$ for all (x,y). The former then implies $a > 0$ through the third line in the alternative, so we can normalize $a = 1$ by dividing both sides of the latter equation by a. The latter now becomes

$$\phi_p(x,y) + b\phi_q(y,x) = c(x) + d(y), \tag{†}$$

which implies $\phi_p \cdot \pi + b\phi_q \cdot \pi^{\mathrm{T}} = 0$ for any $\pi \in \Gamma(p,q)$. As $(p,q) \in \mathcal{D}^S$, we know $b > 0$. $\square$

We now further normalize b to 1. If $(p,p) \in \mathcal{D}^S$, notice that (†) gives both $\phi_p(x,y) + b\phi_p(y,x) = c(x) + d(y)$ and $\phi_p(y,x) + b\phi_p(x,y) = c(y) + d(x)$. Adding them up we obtain $\phi_p(x,y) + \phi_p(y,x) = c(x) + d(x) + c(y) + d(y)$ (after dividing $c(.)$ and $d(.)$ by $b+1$). Together with $c \cdot p + d \cdot p = 0$, we obtain $\phi_p \cdot \pi + \phi_p \cdot \pi^{\mathrm{T}} = 0$ for all $\pi \in \Gamma(p,p)$.

Now we fix p and further normalize ϕ_q for all $q \neq p$ so that $\phi_p(x,y) + \phi_q(y,x) = c_{pq}(x) + d_{pq}(y)$ with $c_{pq} \cdot p + d_{pq} \cdot q = 0$, where c_{pq} and d_{pq} represent c and d derived from

($\dagger$) with the corresponding pair (p,q). Consequently, $b=1$ and $\phi_p \cdot \pi + \phi_q \cdot \pi^{\mathrm{T}} = 0$. For $q, r \neq p$, we show below that $b_{pq}b_{qr}b_{rp} = 1$ and hence b_{qr} must be equal to 1 after the other two are normalized to 1. To see this, we substitute the terms ϕ_q and ϕ_r in equation ($\dagger$) for (q,r) with ϕ_p:

$$\begin{aligned}
c_{qr}(x) + d_{qr}(y) &= \phi_q(x,y) + b_{qr}\phi_r(y,x) \\
&= \frac{1}{b_{pq}}[c_{pq}(y) + d_{pq}(x) - \phi_p(y,x)] + b_{qr}[c_{rp}(y) + d_{rp}(x) - b_{rp}\phi_p(x,y)] \\
&= -[\frac{1}{b_{pq}}\phi_p(y,x) + b_{qr}b_{rp}\phi_p(x,y)] + \frac{1}{b_{pq}}[c_{pq}(y) + d_{pq}(x)] + b_{qr}[c_{rp}(y) + d_{rp}(x)] \\
&= (\frac{1}{b_{pq}} - b_{qr}b_{rp})\phi_p(y,x)] - b_{qr}b_{rp}[\phi_p(x,y) + \phi_p(y,x)] \\
&\quad + \frac{1}{b_{pq}}[c_{pq}(y) + d_{pq}(x)] + b_{qr}[c_{rp}(y) + d_{rp}(x)] \\
&= (\frac{1}{b_{pq}} - b_{qr}b_{rp})\phi_p(y,x) - b_{qr}b_{rp}[c_{pp}(x) + d_{pp}(x) + c_{pp}(y) + d_{pp}(y)] \\
&\quad + \frac{1}{b_{pq}}[c_{pq}(y) + d_{pq}(x)] + b_{qr}[c_{rp}(y) + d_{rp}(x)]
\end{aligned}$$

Now suppose to the contrary that $\frac{1}{b_{pq}} - b_{qr}b_{rp} \neq 0$. Then, $\phi_p(y,x)$ can be expressed as terms that depend on only x or only y. This contradicts to the premise that $(p,q) \in \mathcal{D}^S$. Consequently, $\frac{1}{b_{pq}} - b_{qr}b_{rp} = 0$. □

For our analysis below, let $t^k \in \mathbb{R}^{N\times N}$, $k \in \{1,2,\dots,N\}$ be the vector with the $(x_k,\cdot)$ entries equal to 1 with the other entries being 0; t^k describes the row marginal constraints $\pi_1 = p$ with $t^k \cdot \pi = p^k$. Similarly define $\bar{t}^k \in \mathbb{R}^{N\times N}$ for the column marginal constraints. The proof of Proposition 5 shares a similar spirit with that of Lemma 2 in the next subsection.

Proof of Proposition 5. By SML insensitivity and correlation completeness, $(p,p) \in \mathcal{D}^{Ind}$ for $p \in \Delta X$. Then, the set $\Gamma(p,p)$, which is a spanning set of the affine subspace described by the linear system $[t^k \cdot \pi = p^k, \bar{t}^k \cdot \pi = p^k]$, is contained in the affine subspace described by $\phi_p \cdot \pi = 0$. Then, there exists $\{\mu_p^k\}_{k=1}^N, \{\lambda_p^k\}_{k=1}^N \in \mathbb{R}$ such that $\phi_p = \sum \lambda_p^k t^k + \mu_p^k \bar{t}^k$.[41] The first term is a constant scalar for $\pi \in \Delta_p$, so there exist a scalar $\bar{\lambda}_p \in \mathbb{R}$ and vector $\mu_p \in \mathbb{R}^N$ such that $\phi_p = \bar{\lambda}_p \cdot \mathbf{1} + \sum_1^N \mu_p^k \bar{t}^k$ represents the preference in Δ_p, i.e., (p,q') is locally insensitive for any $q' \in \Delta X$, and $p \succeq q' \iff \bar{\lambda}_p + \mu_p \cdot q' \geqslant 0$. □

[41]See e.g., Chapter 8 of Sernesi (2019).

B.3 Proofs for Subsection 5.3

Given Assumption R, the SSB representation on each connected component P^I of $\mathcal{P}^I$ follows from Fishburn's (1982) original proof. We proceed in 5 steps in the proof of Proposition 6 below.

Proof of Proposition 6 (CWU). We first restrict attention to a connected component $P^S \subset \mathcal{P}^S$. We will fix a starting marginal r and let $\phi = \phi_r$ and show that for every p, the preference on Δ_p is represented by $\phi_p = \phi + \bar{\lambda}_p \cdot \mathbf{1} + \sum_1^N \mu_p^k \bar{t}^k$ where scalars $\bar{\lambda}_p \in \mathbb{R}$ and vectors $\mu_p \in \mathbb{R}^N$ (Step 1). We then prove that the coefficients $\bar{\lambda}_p$ and μ_p are linear in p (Step 2) locally. In Step 3 we transform it to a unique skew-symmetric representation. Step 4 extends the linearity to the whole connected $\mathcal{P}^S$, with the rest of proof completed in Step 5.

We first prove the following lemma that leads to the proof of Step 1.

Lemma 2. *For $(q, q') \in \mathcal{D}^S$, there exist scalars $\bar{\lambda}_{q'} \in \mathbb{R}$ and vectors $\mu_{q'} \in \mathbb{R}^N$ such that $\phi_{q'} = \phi_q + \bar{\lambda}_{q'} \cdot \mathbf{1} + \sum_1^N \mu_{q'}^k \bar{t}^k$.*

Proof. For $(q, q') \in \mathcal{D}^S$, fix $\bar{\pi} \in \Gamma(q, q')$ and $\alpha \in [0, 1]$. Since any symmetric $\pi \in \Gamma(q, q)$ is in $\tilde{\Pi}$ by correlation completeness, correlation betweenness implies that for $q_\alpha := \alpha q + (1-\alpha)q'$, we have $(q, q_\alpha) \in \mathcal{D}^S$.

Pick any $\tilde{\pi} \in \Gamma(q, q) \bigcap \tilde{\Pi}$. By correlation completeness, $\tilde{\pi}^{\mathrm{T}} \in \tilde{\Pi}$ as well. Let $\pi_\alpha = \alpha \tilde{\pi}^{\mathrm{T}} + (1-\alpha)\bar{\pi} \in \Gamma(q, q_\alpha)$. Since $(q, q_\alpha) \in \mathcal{D}^S$, by Lemma 1, there exists $b > 0$ such that $\phi_q \cdot \pi_\alpha + b\phi_\alpha \cdot \pi_\alpha^{\mathrm{T}} = 0$, where ϕ_α is a shorthand for ϕ_{q_α}. Substituting π_α with $\pi_\alpha = \alpha\tilde{\pi}^{\mathrm{T}} + (1-\alpha)\bar{\pi}$ and noticing $\phi_q \cdot \tilde{\pi}^{\mathrm{T}} = 0$, the equation is rewritten as $\phi_\alpha \cdot \tilde{\pi} = -\frac{1-\alpha}{\alpha}(\frac{1}{b}\phi_q \cdot \bar{\pi} + \phi_\alpha \cdot \bar{\pi}^{\mathrm{T}})$. We denote the RHS constant by $C_{\bar{\pi}}$.

Notice $\tilde{\pi}$ is arbitrarily picked from $\Gamma(q, q) \bigcap \tilde{\Pi}$, which is a spanning set of the affine subspace described by the linear system $[\phi_q \cdot \pi = 0, t^k \cdot \pi = q^k, \bar{t}^k \cdot \pi = q^k]$, where the latter two are the marginal constraints $\pi_1 = \pi_2 = q$. Since $\phi_\alpha \cdot \tilde{\pi} = C_{\bar{\pi}}$, the affine subspace above is also represented by $[\phi_\alpha \cdot \pi = C_{\bar{\pi}}, t^k \cdot \pi = q^k, \bar{t}^k \cdot \pi = q^k]$. Hence, there exists $\tau_\alpha, \{\mu_\alpha^k\}_{k=1}^N, \{\lambda_\alpha^k\}_{k=1}^N \in \mathbb{R}$ such that $\phi_\alpha = \tau_\alpha \phi_q + \sum(\lambda_\alpha^k t^k + \mu_\alpha^k \bar{t}^k)$. We now further simplify the expression.

First, as ϕ_α only evaluates $\pi \in \Delta_{q_\alpha}$, for any $\pi \in \Gamma(q_\alpha, \tilde{q})$, $\phi_\alpha \cdot \pi = \tau_\alpha \phi_q \cdot \pi + \lambda_\alpha \cdot q + \mu_\alpha \cdot \tilde{q}$. So we can replace the family $\{\lambda_\alpha^k\}_{k=1}^N$ with a fixed scalar $\bar{\lambda}_\alpha = \lambda_\alpha \cdot q$. Then, we argue $\tau_\alpha > 0$: it cannot be 0 as $(q, q_\alpha) \in \mathcal{D}^S$, while the argument below shows it must be nonnegative.

We fix the second marginal as q and consider, for each α, the three sets $\hat{\Pi}\bigcap\Gamma(q_\alpha, q)$, $\tilde{\Pi}\bigcap\Gamma(q_\alpha, q)$, and $\check{\Pi}\bigcap\Gamma(q_\alpha, q)$. By local correlation sensitivity, they are non-empty. Notice that within the parallel sets $\Gamma(q_\alpha, q)$, the indifference sets $\{\tilde{\Pi}\bigcap\Gamma(q_\alpha, q) : \alpha \in [0,1]\}$, are also parallel to each other, since $\phi_\alpha \cdot \pi = \tau_\alpha \phi_q \cdot \pi + \bar{\lambda}_\alpha + \mu_\alpha \cdot q$. Hence, the sign of τ_α decides whether each "side" belongs to the strict preference set $\hat{\Pi}$ or the complement $\check{\Pi}$. Now suppose two sets $A^+ = \{\alpha : \tau_\alpha > 0\}$ and $A^- = \{\alpha : \tau_\alpha < 0\}$ are both nonempty. As A^+ and A^- is a partition of the compact interval $[0,1]$, their closures must have nonempty intersection. Let α^* be a common limit point of A^+ and A^-. Then, by continuity, the set $\hat{\Pi}\bigcap\Gamma(q_{\alpha^*}, q)$ must be on the same side as those $\hat{\Pi}\bigcap\Gamma(q_\alpha, q)$ for $\alpha \in A^+$. At the same time, it also falls on the other side which contains those $\hat{\Pi}\bigcap\Gamma(q_\alpha, q)$ for $\alpha \in A^-$, a contradiction. Therefore, A^- must be empty as $1 \in A^+$.

Put together, we can normalize $\phi_\alpha = \phi_q + \bar{\lambda}_\alpha \cdot \mathbf{1} + \sum_1^N \mu_\alpha^k \bar{t}^k$. Specifically, this holds for $\alpha = 0$, i.e., for $q_\alpha = q'$. $\square$

Step 1. For every p, the preference on Δ_p is represented by $\phi_p = \phi + \bar{\lambda}_p \cdot \mathbf{1} + \sum_1^N \mu_p^k \bar{t}^k$ where scalars $\bar{\lambda}_p \in \mathbb{R}$ and vectors $\mu_p \in \mathbb{R}^N$. For each $\tilde{p}$ in P^S, continuity implies that there is a neighborhood $B(\tilde{p}) \subset \Delta X$ such that $(\tilde{p}, p') \in \mathcal{D}^S$ for all $p' \in B(\tilde{p})$. Then, for any $r, p \in P^S$, by the open covering theorem (applied to the compact path[42] from r to q), there exists a finite sequence $(q^0 = r, q^1, q^2, \ldots, q^K, q^{K+1} = p)$ such that $(q^k, q^{k+1}) \in \mathcal{D}^S$ for all $k = 1, 2, \ldots, K$. Since $(q^k, q^{k+1}) \in \mathcal{D}^S$ for all $k = 0, 1, \ldots, K$, we can let $(q, q') = (q^k, q^{k+1})$ in Lemma 2 and obtain recursively $\phi_p = \phi + \bar{\lambda}_p \cdot \mathbf{1} + \sum_1^N \mu_p^k \bar{t}^k$.

Step 2. $\psi(p, q) =: (\bar{\lambda}_p + \mu_p \cdot q)$ is bilinear for q in some neighborhood of p.

We fix any full-support p, q sufficiently close such that there exists a open ball $\mathcal{B}$, with (p, s), $(q, s) \in \mathcal{D}^S$ for any $s \in \mathcal{B}$. We pick arbitrary $\pi_p \in \Gamma(p, s)\bigcap\tilde{\Pi}$, and similarly $\pi_q \in \Gamma(q, s)\bigcap\tilde{\Pi}$. Denote $\pi_\beta = \beta\pi_p + (1-\beta)\pi_q$, which belongs to $\tilde{\Pi}$ by correlation betweenness and correlation completeness.

By our selection, we have $0 = \phi_p \cdot \pi_p = \phi_q \cdot \pi_q = \phi_\beta \cdot \pi_\beta$, which expands to the

[42] P^S is an open and connected subset of an Euclidean space, and hence is path-connected. Formally, a path is a continuous mapping $f : [0,1] \to P^S$ with $f(0) = p$ and $f(1) = q$. Such argument also appears in Segal (1992) when he deals with transitive preference in a non-convex domain.

following equations:

$$\begin{cases} 0 = \phi \cdot \pi_p + \bar{\lambda}_p + \mu_p \cdot s, \\ 0 = \phi \cdot \pi_q + \bar{\lambda}_q + \mu_q \cdot s, \\ 0 = \phi \cdot \pi_\beta + \bar{\lambda}_\beta + \mu_\beta \cdot s, \end{cases} .$$

We substitute π_β with π_p and π_q using the first two equations and obtain

$$0 = [\bar{\lambda}_\beta - \beta\bar{\lambda}_p - (1-\beta)\bar{\lambda}_q] + [\mu_\beta - \beta\mu_p - (1-\beta)\mu_q] \cdot s.$$

Since s is chosen arbitrarily from a full-dimension subset, it must be that both terms in the square brackets equal to 0. This gives the linearity of ϕ_p in p. The equivalent representation in the Theorem can be obtained from grouping $\bar{\lambda}_p$ and μ_p as $\phi_p \cdot \pi = \phi \cdot \pi + \bar{\lambda}_p + \mu_p \cdot q$ for any $\pi \in \Gamma(p,q)$. Hence, $\psi(p,q) =: (\bar{\lambda}_p + \mu_p \cdot q)$ is bilinear.

Step 3. Uniqueness of a skew-symmetric representation in the neighborhood.

We now again apply Lemma 1 to a pair $(p,q) \in \mathcal{D}^S$. There exists $b' \in \mathbb{R}$ such that $(\phi + b'\phi^{\mathrm{T}}) \cdot \pi + \psi(p,q) + b\psi(q,p) = 0$, where $\phi^{\mathrm{T}}(x,y) = \phi(y,x)$. This means that there exists c, d such that $\phi(x,y) + b'\phi(y,x) = c(x) + d(y)$. Then b' must be common for all p, q; otherwise, taking a difference of the two equations for two different pairs gives $\phi(x,y) = c'(x) + d'(y)$ for some c', d', which contradicts $(p,q) \in \mathcal{D}^S$. Given this uniformity, it must be $b' = 1$ as we have shown $b_{pq}b_{qr}b_{rp} = 1$ for any three p, q, r, so that $\phi(x,y) + \phi(y,x) = c(x) + d(y)$, and $(\phi + \phi^{\mathrm{T}}) \cdot \pi + \psi(p,q) + \psi(q,p) = 0$. The former implies $c(x) = d(x) + C$. We can define $\phi'(x,y) = \phi(x,y) - \frac{1}{2}[c(x) + c(y) + C]$, and $\psi'(p,q) = \frac{1}{2}(c \cdot p + c \cdot q + C) + \psi(p,q)$, so that $\phi \cdot \pi + \psi(p,q) = \phi' \cdot \pi + \psi'(p,q)$, while $\phi'(x,y) + \phi'(y,x) = 0 = \psi'(p,q) + \psi'(q,p)$. Uniqueness also follows from the argument.

Step 4. Bilinearity in the whole $\mathcal{P}^S$ and a common skew-symmetric representation.

Consider any $\bar{p}, \underline{p} \in \mathcal{P}^S$. If the line segment $[\bar{p}, \underline{p}] \subset P^S$ for some connected component P^S, then through a similar open covering argument with Step 2 applied to this line segment, we can easily extend the local bilinearity to $[\bar{p}, \underline{p}]$.

Now suppose $[\bar{p}, \underline{p}] \not\subset P^S$. Consider the respectively unique CWU representations (ϕ_1, ψ_1) and (ϕ_2, ψ_2) in their neighborhoods $\mathcal{N}_{\bar{p}}$ and $\mathcal{N}_{\underline{p}}$. As the interval is not contained in P^S, we can select $\alpha \in (0,1)$ so that $p_\alpha = \alpha\bar{p} + (1-\alpha)\underline{p}$ is in some P^I.

Without loss, we can assume p_α is in its interior as otherwise, we can perturb $\bar{p}$ and $\underline{p}$ by openness of $\mathcal{P}^S$, show bilinearity after perturbation, and return to bilinearity for $\bar{p}, \underline{p}$ with continuity.

By definition of $\mathcal{P}^I$, $p_\alpha \sim p_\alpha$. SSB representation on P^I ensures that there exists a line segment $[p_\alpha, r_\alpha] \subset P^I$ such that $p_\alpha \sim q'$ for any $q' \in [p_\alpha, r_\alpha]$. We can pick a $\bar{q} \in \mathcal{N}_{\bar{p}}$ close enough to $\bar{p}$, and q' close enough to p_α such that there exists a $\underline{q} \in \mathcal{N}_{\underline{p}}$ with $q' = \alpha\bar{q} + (1-\alpha)\underline{q}$.[43]

By the above, we can pick any $\bar{\pi} \in \tilde{\Pi} \bigcap \Gamma(\bar{p}, \bar{q})$ and $\underline{\pi} \in \tilde{\Pi} \bigcap \Gamma(\underline{p}, \underline{q})$, and their $(\alpha, 1-\alpha)$ combination $\pi_\alpha \in \Gamma(p_\alpha, q_\alpha) \subset \tilde{\Pi}$. By correlation projective independence, $\forall \beta \in (0,1)$, π_β is also in $\tilde{\Pi}$. Specifically, this holds for a non-empty segment of β such that $p_\beta \in \mathcal{N}_{\underline{p}}$. By CWU representation on $\mathcal{N}_{\underline{p}}$, we have

$$
\begin{aligned}
0 &= \phi_2 \cdot \pi_\beta + \psi_2(p_\beta, q_\beta) \\
&= \beta\phi_2 \cdot \bar{\pi} + (1-\beta)\phi_2 \cdot \underline{\pi} + \beta^2\psi_2(\bar{p}, \bar{q}) + (1-\beta)^2\psi_2(\underline{p}, \underline{q}) + \beta(1-\beta)[\psi_2(\bar{p}, \underline{q}) + \psi_2(\underline{p}, \bar{q})] \\
&= [\beta\phi_2 \cdot \bar{\pi} + \beta^2\psi_2(\bar{p}, \bar{q})] + \beta(1-\beta)\phi_2 \cdot \underline{\pi} + \beta(1-\beta)[\psi_2(\bar{p}, \underline{q}) + \psi_2(\underline{p}, \bar{q})] \\
&= \beta[\phi_2 \cdot \bar{\pi} + \psi_2(\bar{p}, \bar{q})] + \beta(1-\beta)[-\psi_2(\bar{p}, \bar{q}) + \phi_2 \cdot \underline{\pi} + \psi_2(\bar{p}, \underline{q}) + \psi_2(\underline{p}, \bar{q})]
\end{aligned}
$$

It follows from existences of β in a non-empty segment that both square brackets must be equal to 0, so that $\phi_2 \cdot \bar{\pi} + \psi_2(\bar{p}, \bar{q}) = 0$. Notice that $\bar{\pi}$ is freely chosen from the indifference set in a full-dimension neighborhood of $\bar{p}$. This means (ϕ_2, ψ_2) is also a CWU representation on $\mathcal{N}_{\bar{p}}$. By uniqueness of the CWU representation, it coincides with (ϕ_1, ψ_1) (up to multiplication). Thus, there is a uniform CWU representation across all connected components of $\mathcal{P}^S$.

Step 5. Uniqueness of the SSB representation on $\mathcal{P}^I$.

The proof for unique SSB representation on $\mathcal{P}^I$ is the with a same argument. For connected components P^S_I and P^I_2 of $\mathcal{P}^S$, with their respectively unique SSB representations V_1 and V_2, we can similarly find the counterparts of $\bar{p}, \underline{p}, \bar{q}, \underline{q}$ such that for a non-empty segment of $\beta \subset (0,1)$,

$$
\begin{aligned}
0 &= V_2(p_\beta, q_\beta) \\
&= \beta^2 V_2(\bar{p}, \bar{q}) + (1-\beta)^2 V_2(\underline{p}, \underline{q}) + \beta(1-\beta)[V_2(\bar{p}, \underline{q}) + V_2(\underline{p}, \bar{q})] \\
&= \beta^2 V_2(\bar{p}, \bar{q}) + \beta(1-\beta)[V_2(\bar{p}, \underline{q}) + V_2(\underline{p}, \bar{q})]
\end{aligned}
$$

[43] For instance, we can pick two lines $\bar{l}$ and $\underline{l}$ passing $\bar{p}$ and $\underline{p}$ each, both parallel to the line of $[p_\alpha, r_\alpha]$. Then we choose q' close enough to p_α and a line passing through it, such that its intersections with $\bar{l}$ and $\underline{l}$ at $\bar{q}$ and $\underline{q}$ fall within $\mathcal{N}_{\bar{p}}$ and $\mathcal{N}_{\underline{p}}$, respectively.

Given the arbitrary choice of β, we can similarly conclude that both square brackets must be equal to 0, so that $V_2(\bar{p}, \bar{q}) = 0$, and further V_1 and V_2 coincide (up to a multiplication). $\square$

C Proof of Proposition 8 in Section 6

The proof below is adapted from that of Proposition 1 in Bikhchandani and Segal (2011).[44] They study a special form of PS called *regret-based* preference which conforms to P2.[45] For acts f, g on a state space S, let $\pi[f, g]$ be the induced joint density, and $R[f, g]$ be the distribution of $\psi(f(s), g(s))$ where $\psi(x, y)$ is a regret or rejoicing function for receiving x while foregoing y. The preference is regret-based if $(f, g) \in \Pi \iff (f', g') \in \Pi$ whenever $R[f, g] = R[f', g']$. This requirement is stronger than PS where the condition applies only to (f, g) and (f', g') when $\pi[f, g] = \pi[f', g']$. They show that a regret-based preference is transitive if and only if it reduces to EU. Our Theorem 2 is more permissive in the sense that CEU takes distributions of *outcome pairs* as primitive, whereas Bikhchandani and Segal (2011) examine distributions of *regret* over outcome pairs.

Proof of Proposition 8. For acts f_0 and f_1, we first show that $f_0 \sim f_1$ if they induce the same outcome probability distribution. By finiteness in outcome, let $f_0(s) = x_i$ for $s \in S_i$, and $f_1(s) = y_i$ for $s \in S'_i$ where $\{S_i\}_{i=1}^n$ and $\{S'_i\}_{i=1}^n$ are partitions of S.

Case 1: $S_i = S'_i$ and $\mu(S_i) = 1/n$ for all i. Then there exists a permutation σ on $\{1, 2, \ldots, n\}$ such that $f_1(S_i) = f_0(S_{\sigma(i)})$. Let $f_k(S_i) = f_0(S_{\sigma^k(i)})$, then the induced distributions $\pi[f_k, f_{k+1}]$ are the same for all k. By PS, the preferences over f_k versus f_{k+1} are the same for all k. Since there are at most n different outcomes x_i, there exists a k' such that $f_{k'} = f_0$. Then, transitivity implies that DM must be indifferent between f_k and f_{k+1} for all $k = 0, 1, \ldots, k'$. In particular, $f_0 \sim f_1$.

Case 2: For all i, j, $\mu(S_i \bigcap S_{j'})$ is a rational number. Let $\hat{n}$ be their common denominator. Rewrite f_0 and f_1 in terms of equally likely events $T_1, \ldots, T_{\hat{n}}$. This returns us to Case 1.

Case 3: There exist i, j such that $\mu(S_i \bigcap S_{j'})$ is not rational. In conjunction with continuity, a limiting argument given Case 2 suffices. $\square$

[44] Completeness is also required for the proof of Bikhchandani and Segal's (2011) Proposition 1; see errata to original paper (Chang and Liu, 2024).

[45] P2 requires $fEg \succ f'Ag \Rightarrow fEg' \succ f'Eg'$ for acts f, g and event E.

Online Appendices

A Supplementary Materials for Section 4

In Lanzani's (2022) characterization of symmetric CEU, he makes use of the following correlation version of the Archimedean property in von Neumann and Morgenstern (1944) and Herstein and Milnor (1953).

Correlation Archimedean Property. Π exhibits the correlation Archimedean property if for all $\pi \in \hat{\Pi}$ and $\pi' \in \check{\Pi}$, there exists $\alpha, \beta \in (0,1)$ such that $\alpha\pi + (1-\alpha)\pi' \in \hat{\Pi}$ and $\beta\pi + (1-\beta)\pi' \in \check{\Pi}$.

By contrast, we apply the stronger continuity property (Axiom 0) needed for the separation argument as we begin the our proof of Theorem 1. The second part of our proof uses a weaker continuity property which relates closely to correlation Archimedean property.

We next discuss Lanzani's transitivity axiom which implies correlation insensitivity under CEU.

L-Transitivity. For all $p, q, r \in \Delta X$ and $\pi \in \Gamma(p,q)$, $\pi' \in \Gamma(q,r)$, $p \succeq^{\pi} q$ and $q \succeq^{\pi'} r \Rightarrow p \succeq r$.

Lemma 3. *For Π admitting an CEU representation, L-transitivity implies correlation insensitivity.*

Proof. This can be seen from setting $p_1 = p, p_2 = q$ in the definition of L-transitivity. For a fixed pair (p,q), suppose we have $\Gamma(p,q) \cap \Pi \neq \varnothing$. Then, by pointwise reflexivity and independence, the perfect correlation between (q,q), $\sum_x q(x)\delta_{(x,x)} \in \Pi \Rightarrow \Gamma(q,q) \cap \Pi \neq \varnothing$. So by definition, $\Gamma(p,q) \in \Pi$. This is exactly correlation insensitivity. $\square$

As a consequence, we obtain Proposition 1 of Lanzani (2022) which essentially corresponds to the equivalence between (ii) and (iii) below.

Corollary 2 (CI, L-Transitivity, and EU)**.** *For a CEU preference Π, the following are equivalent:*

(i) Π *is correlation-insensitive;*

(ii) Π *is L-transitive;*

(iii) Π *has an EU representation.*

Proof of $\check{\Pi} = \hat{\Pi}^T$ *under completeness.* By correlation completeness, for any $\pi \notin \Pi$, $\pi^{\mathrm{T}} \in \Pi$. So, $\check{\Pi} = \{\pi : \pi \notin \Pi\} \subset \{\pi : \pi^{\mathrm{T}} \in \Pi\} = \{\pi^{\mathrm{T}} : \pi \in \Pi\}$, so that $\check{\Pi} \subset \{\pi^{\mathrm{T}} : \pi \in \Pi,\ \pi^{\mathrm{T}} \notin \Pi\} = \{\pi^{\mathrm{T}} : \pi \in \hat{\Pi}\} = \hat{\Pi}^{\mathrm{T}}$. Meanwhile, $\hat{\Pi}^{\mathrm{T}}$ is clearly contained in $\check{\Pi}$ by correlation completeness, so we have $\hat{\Pi}^{\mathrm{T}} = \check{\Pi}$, and equivalently, $\check{\Pi} = \{\pi^{\mathrm{T}} : \pi \notin \Pi\}$. $\square$

B Supplementary Materials for Section 5

We begin by showing that SSB implies the *ratio consistency property* stated in subsection 2.4:

Proof. Fix lotteries p, q, r as given in the definition of ratio consistency. Notice that $\beta p + (1-\beta)r \sim \gamma q + (1-\gamma)r$ implies

$$\begin{aligned}
0 &= \psi(\beta p + (1-\beta)r, \gamma q + (1-\gamma)r) \\
&= \beta\gamma\psi(p,q) + \beta(1-\gamma)\psi(p,r) + (1-\beta)\gamma\psi(r,q) + (1-\beta)(1-\gamma)\psi(r,r) \\
&= \beta(1-\gamma)\psi(p,r) + (1-\beta)\gamma\psi(r,q),
\end{aligned}$$

where the second line uses bilinearity of ψ, and reduction to the last line is due to the indifference in (p,q) and (r,r). Hence, the ratio $\frac{\beta/(1-\beta)}{\gamma/(1-\gamma)} = \frac{\psi(r,q)}{\psi(r,p)}$ is constant across (β, γ) pairs for fixed p, q, r. $\square$

Fishburn (1982) introduces the following weakened version of projective independence, called *symmetry*: ($p \succ q \succ r, p \succ r$ and, $q \sim \frac{1}{2}p + \frac{1}{2}r$) implies ($\lambda p + (1-\lambda)r \sim \frac{1}{2}p + \frac{1}{2}q \iff \lambda r + (1-\lambda)r \sim \frac{1}{2}r + \frac{1}{2}q$) for $\lambda \in (0,1)$. Under betweenness, symmetry can be deduced from projective independence through a simple geometry argument, as the latter implies that the straight indifferent curves are projective, resulting in a bilinear utility representation. In Section 5.2, we obtain the conditional linearity result of *Lemma 3* in Fishburn (1982). Fishburn then calibrate v_p to arrive at a skew-symmetric $\psi(p,q)$ defined as $\psi(p,q) = v_p(q)$.

Notice that Fishburn's proof actually implies the following key observation: given pairwise non-indifferent p, q, r, we can find α, β, γ (not necessarily positive) such that for $t = p, q, r$, the *linearly extended* v_t to the affine set[46] satisfies $v_t(\alpha p + \beta q + \gamma r) = 0$. When there is a preference cycle $p \succ q \succ r \succ p$, the combi-

[46] The set $\bar{\Delta}X = \{\bar{p} : \sum_x \bar{p}(x) = 1\}$ where each coordinate is not necessarily positive.

nation is within the simplex of $\{p, q, r\}$. Otherwise, for the transitive preference $p \succ q \succ r$, and $p \succ r$, the combination falls outside the simplex but within the extended affine set. Geometrically, this means the three (*extended*) indifference curves p, q, r are projective. Now, if we start from a reference r^* and define $\psi(r^*, p) = -\psi(p, r^*) = v_{r*}(p)$, letting $r = r^*$ gives the desired calibration $v_p(q) = -v_q(p)$.

C Supplementary Materials for Section 6

C.1 Fishburn's Nontransitive SEU

As in the main text, let S, X, F be the sets of states, outcomes, and acts with $F \subset X^S$. Let $x \in X$ also denote the constant act with the same outcome x. Fishburn studies an asymmetric preference relation $\succ$ on F (and its induced weak preference $\succeq$ and indifference $\sim$). For $E \subset S$, the conditional preference $f \succ_E g$ means $fEh \succ gEh$ for any h. Denote by $\mathcal{N}$ the null events as usual.

Fishburn requires $\succ$ to satisfy the following axioms corresponding to those of S54 except for the last axiom S1* which can be viewed as a conditional form of STP.

P1*. $\succ$ is asymmetric and, for all $x, y \in X$, $\succ$ is a weak order on $F_{xy} = \{f \in F : f(s) \in \{x, y\}\}$.

P2. $fEg \succ f'Ag \Rightarrow fEg' \succ f'Eg'$.

P3. For $E \notin \mathcal{N}$, $xEf \succ xEf \iff x \succ y$.

P4. $(x \succ y, z \succ w) \Rightarrow (xAy \succ xBy \iff zAw \succ zBw)$.

P5. $x' \succ y'$ for some $x', y' \in X$.

P6*. $(f \succ g; x, y \in X) \Rightarrow$ there exists a finite partition of S such that for every event E in the partition, $fEx \succ gEy$, $fEx \succ g$, and $f \succ gEy$.

S1*. $(A \bigcap B) = \varnothing, f \succeq_A g, f \succeq_B g) \Rightarrow f \succeq_{A \bigcup B} g$; if in addition, $f \succ_A g$, then $f \succ_{A \bigcup B} g$.

Fishburn shows that, if the preference $\succ$ satisfies all seven axioms above, then there is a probability measure μ on S and a skew-symmetric $\varphi : X \times X \Rightarrow \mathbb{R}$ such that, μ satisfies parts (i) to (iii) of Savage's original theorem, along with the nontransitive SEU representation (iv):

(i) $A \in \mathcal{N} \iff \mu(A) = 0$;

(ii) for $A \subset S, \lambda \in (0,1)$, there exists $A' \subset A$ such that $\mu(A') = \lambda \cdot \mu(A)$;

(iii) for any $x \succ y$, $xAy \succ xBy \iff \mu(A) > \mu(B)$;

(iv) $f \succ g \iff \int_S \varphi(f(s), g(s))d\mu > 0$.

In his proof, Fishburn first restricts attention to the set of acts with binary outcomes $\{F_{xy} : x, y \in X\}$ and borrows Savage's theorem to obtain a candidate probability measure μ. He then shows that μ together with a calibrated φ represents the preference. Notice that nontransitive SEU satisfies our PS as $\varphi(f(\cdot), g(\cdot))$ effectively translates binary acts the into joint densities over outcome pairs. Hence, while nontransitive SEU is set to describe nontransitive preferences over Savagean acts under uncertainty, it corresponds to Lanzani's correlation-sensitive representation, or equivalently, our symmetric CEU preference, in the risk domain.

C.2 Comparing Fishburn (1989) and Lanzani (2022)

Assuming that DM is PS, we can induce, through μ, a joint density $\pi \in \Delta(X \times X)$ for each pair of acts f, g. Then, the original preference $\succ$ under uncertainty induces a corresponding correlation preference $\Pi^{\succ}$ under risk. The proposition below formally connects the strong independence of the induced $\Pi^{\succ}$ and S1*.

Proposition 11. *Assuming PS and Fishburn's P1* and P6*, S1* is equivalent to strong independence of the induced correlation preference.*

Proof. P1* and P6*, as in their original versions, imply completeness and continuity of preference. It is then easy to check that strong independence of the induced $\Pi^{\succ}$ implies S1* on $\succ$.

To see that S1* implies strong independence, suppose we have two joint densities π, π', induced from act pairs (f, g) and (f', g'). Then, $\pi, \pi' \in \Pi$ is equivalent to $f \succeq g, f' \succeq g'$. We now show that $\pi_{1/2} = \frac{1}{2}\pi + \frac{1}{2}\pi' \in \Pi$. Notice that by property (ii) of μ, we can find $A \subset S$ with $\mu(A) = \frac{1}{2}$ and acts f'', g'', f''', g''', such that (1) the joint densities induced by (f'', g'') *conditional on* A and A^c equal π, π' respectively, and similarly (2) the joint densities induced by (f''', g''') *conditional on* A and A^c equal π', π respectively. As a result, the (full) joint densities induced by $(f'', g''), (f''', g''')$ are $\pi'' = \pi''' = \pi_{1/2}$.

Suppose $g'' \succ_A f''$. Then, by PS, $g''' \succ_{A^c} f'''$. Define $\tilde{f}$ as $\tilde{f} =_A f''$ and $\tilde{f} =_{A^c} f'''$, and $\tilde{g}$ similarly. Now S1* implies $\tilde{g} \succ \tilde{f}$, which violates PS as we have assumed that $f \succeq g$ and that $(f,g), (\tilde{f}, \tilde{g})$ induce the same joint density π. Hence, it must be that $f'' \succeq_A g''$. Similarly, $f'' \succeq_{A^c} g''$, so $f'' \succeq g''$ again by S1*. This means $\pi_{1/2} = \pi'' \in \Pi$ as desired.

We can now go on to show that $\pi_\alpha = \alpha\pi + (1-\alpha)\pi' \in \Pi$ for all rational $\alpha \in (0,1)$ with $\alpha = \frac{m}{2^n}$. By continuity given P6*, we can further extend to all rational and real α. This completes the proof. □

D Supplementary Materials for Section 7

Here, we focus on *correlation Dual Utility.* Consider the following representation:

$$\pi \in \Pi \iff U(\pi) = \int_{[0,1]^2} \psi(G_{\pi_1}(x), G_{\pi_2}(y)) dC^\pi(1-x, 1-y) \geqslant 0.$$

Using a change of variable $p = G_{\pi_1}(x)$, $q = G_{\pi_2}(y)$, we can write the representation as $U(\pi) = \int_{[0,1]^2} \phi(p,q) dC^\pi(1 - G_{\pi_1}^{-1}(p), 1 - G_{\pi_2}^{-1}(q))$. This resembles a CEU representation with quantiles playing the role of marginals.

Observe that copulas have uniform marginals and suppose Π is correlation-insensitive. By a Kantorovich duality argument, we can arrive at a dual representation: $\int_{[0,1]} u^{G_{\pi_1}, G_{\pi_2}}(p) d[1 - G_{\pi_1}^{-1}(p)] - \int_{[0,1]} v^{G_{\pi_1}, G_{\pi_2}}(q) d[1 - G_{\pi_2}^{-1}(q)]$ for each fixed pair of marginals. A minimax argument then delivers constancy of both u and v in the marginals, thereby reducing to Yaari's original *dual utility* representation.

E Extremal Correlation, Matching, and Optimal Transport

An important implication of Theorem 2 is that any non-EU preference within the CEU framework is sensitive to the correlation between the two lotteries being compared. An interesting direction in the non-EU literature is to assess directly the axioms underpinning the models considered. We show in the next subsection how to check DM's robustness to correlation sensitivity for a given lottery pair (p,q) through the well-known *network simplex* algorithm (see, e.g., Chapter 7 of Bertsimas and Tsitsiklis (1997)), which efficiently finds one, if not the unique solution to the OT problem. This is a special case of the simplex algorithm first discovered by

Dantzig (1951).[47,48]

The network simplex algorithm iterates on extreme points of $\Gamma(p, q)$ by updating according to cycles created and removed in the support graph, which is closely related to *cyclic monotonicity*, one of the equivalent optimality conditions in the Kantorovich duality (see, e.g., Villani (2009)). Cyclic monotonicity and its variants are widely observed in optimality conditions in the economics literature. A well-known application is for optimization in quasi-linear settings (Rochet, 1987), particularly in mechanism design and in auction (Vohra, 2011). Lin and Liu (2024) study the problem of credibly persuading a receiver who can observe and verify the final signal distribution when the sender is not able to commit to the information structure. In an OT setup, they show that any implementable policy must be cyclic monotone, and furthermore *comonotonic* if the sender's utility involves a certain supermodularity. Anderson and Smith (2024) study the comparative statics of a matching problem using a formulation very similar to the OT problem. They propose a *synergy* function: the cross-difference $\phi(x_i, y_i) + \phi(x_{i+1}, y_{i+1}) - \phi(x_i, y_{i+1}) - \phi(x_{i+1}, y_i)$, behind which the idea is closely related to the cycle modification in the network simplex algorithm.[49]

We have considered examples of extremal correlations when we discuss the ERD model after Theorem 2, where we borrow the idea of assortative matching from the matching literature for SML under submodularity. In fact, a discrete OT problem can be equivalently viewed as an assignment problem (Shapley and Shubik, 1971).

[47]Many other algorithms are developed in the operations research literature. The classical Hungarian method for assignment problem is developed by Kuhn (1955), also based on cross-difference and cyclic improvements. This lies at the foundation of the well-known multi-item auction algorithm (Demange et al., 1986), also closely connected to the job matching process (Crawford and Knoer, 1981).

[48]When CEU is symmetric, Theorem 2 (ii) implies that we can restrict ourselves to SML, i.e., $\pi \in \Gamma(p, p)$ for $p \in \Delta X$. Then the optimization problem, viewed as a zero-sum game, relates to von Neumann's observation that for any finite two-person zero-sum game, there is a feasible linear programming problem whose saddle points yield the equilibria of the game. We can do better by invoking the following known result on fictitious play: if both players use a no-regret learning algorithm to respond to their opponent's strategies, then the average payoff of the players converge to their minimax value, and their average strategies constitute an approximate minimax equilibrium, with the approximation converging to 0; see Chapter 7 of Cesa-Bianchi and Lugosi (2006). No-regret learning algorithm is not only a natural iterative dynamics: its rate of convergence is fast, at the order of the average regret in T rounds and recently improved to $O(lnT/T)$ for zero-sum games (Daskalakis et al., 2015).

[49]They show that this synergy function solely determines the optimal matching plan; for instance, an everywhere positive synergy, corresponding to a supermodular ϕ, leads to the famous positive assortative matching (Becker, 1973). Comparative statics given varying ϕ is examined and shown to be related to the positive quadrant order (Lehmann (1966)).

Algorithms such as the well-known auction mechanism of Crawford and Knoer (1981) and Demange et al. (1986) for matching can then fit in directly.[50] For CEU, choice reversal for (p, q) is a generic property. We can, without losing generality, focus on lotteries with rational probabilities. For rational marginals, we can "replicate" the points in the support thereby reducing them to uniform densities along with reducing the feasible set of joint densities to doubly stochastic matrices.[51] We then refer to the Birkhoff-von Neumann Theorem which identifies all permutation matrices.

E.1 Test Joint Densities for a Fixed CEU Kernel

Suppose that a decision maker exhibits correlation sensitivity with respect to a pair of lotteries (p, q). We demonstrate how to find $\pi_+, \pi_- \in \Gamma(p, q)$ such that $\pi_+ \in \Pi$ while $\pi_- \notin \Pi$, i.e., her preference over (p, q) changes when correlation varies from π^+ to π^-. If in addition $\pi_-^{\mathrm{T}} \in \Pi$ (which is true for a skew-symmetric ϕ), then she prefers p to q under π^+ and conversely under π^-.

Specifically, we can pick $\pi_- \in \arg\min_{\pi \in \Gamma(p,q)} \sum \phi(x, y)\pi(x, y)$; in terms of correlation sensitivity, this π_- is the worst-case correlation in $\Gamma(p, q)$ to ϕ. It is a "test" joint density determining whether DM's preference over (p, q) is robust to correlation variations — if, for any pair (p, q), DM prefers p to q even under the worst-case correlation $\pi_- \in \Gamma(p, q)$. Similarly, pick $\pi_+ \in \arg\max_{\pi \in \Gamma(p,q)} \sum \phi(x, y)\pi(x, y)$ — this π_+ "favors" the lottery p most among all joint densities in $\Gamma(p, q)$. It is the best-case correlation that determines whether DM would, under any correlation, prefer p to q. These best-case and worst-case test joint densities are of importance for testing whether DM's preference is sensitive to correlation.[52]

Observe that the problem is linear and boils down to finding the extreme points of $\Gamma(p, q)$ and then solving the OT problem, which can be done through the network simplex algorithm.

[50] Bertsekas and Castanon (1989) convert the transport problem to an assignment problem. A rich strand of literature extends and improves on the early methods in the general transport problem, see, e.g., Kim (2010), Chapter 1.5 for a review.

[51] A matrix is doubly stochastic if the sum of entries of each row and column is always 1.

[52] In fact, the two extremal correlations correspond to the two possible ways of defining the induced preference in Lanzani (2022). In his main text, he defines it as our correlation-insensitive preference relation $\succeq^*$, with a "for all" quantifier. Meanwhile, he proposes another in the footnote right after it, where he defines $p \succeq^* q \iff \exists \pi \in \Pi$ such that $\sum \phi(x, y)\pi(x, y) \geqslant 0$. While we have shown that the former preference is generally incomplete, the latter often involves too much indifference as it is not too demanding that both $p \succeq^* q$ and $q \succeq^* p$ hold.

Procedure (Network simplex algorithm). *Start with a tree* $\pi^0 \in \Gamma(p, q)$.

1. *Given the tree* π *from previous iteration, compute the dual* (u, v) *according to complementary slackness:* $v_{pq}(y) = \min_x \{\phi(x, y) + u_{pq}(x)\}$, $u_{pq}(x) = \max_y \{-\phi(x, y) + v_{pq}(y)\}$.
2. *Compute the reduced costs* $\hat{\phi}(x, y) = \phi(x, y) - (u(x) - v(y))$ *for all* (x, y) *not in the support. If none of them is negative, terminate as we already obtain an optimal solution; otherwise, choose a node* (x', y') *with* $\hat{\phi}(x', y') < 0$ *and add it to the support graph.*
3. *The entering* (x', y') *forms a unique cycle* C *with other nodes and edges in* G_π. *For a positive number* θ, *a new cycle* C_θ *is obtained from* C *by alternatively adding* θ *to and subtracting* θ *from* $\pi(x, y)$ *along the cycle* C, *starting from* (x', y'). *Let* $\theta^* = \arg\max_\theta \{C_\theta$ *remains a non-negative cycle*$\}$.
4. *There is a unique node in* C_{θ^*} *with zero value. Replace it with* (x', y') *and we obtain a new tree* π'. *Return to Step 1.*

For example, Bordalo et al. (2012) test the significance of correlation in Allais paradox in their appendix. After the classical (independent) common-consequence Allais test $L_1^0 = (2500, 33\%;\ 0, 67\%)$ versus $L_2^0 = (2400, 34\%;\ 0, 66\%)$, they ask the participants to decide on the same, but correlated pair of lotteries.

By the results in the next subsection, we will see that π^0 in our example is one (of the two) extreme point of $\Gamma(p, q)$. In fact, it survives the network simplex algorithm when ϕ satisfies the salience conditions in Bordalo et al. (2012). Hence, it is the correlation for which a salience minded decision maker will "most likely" not prefer L_1^0 to L_2^0. Their experimental results confirms this: more than half (26% out of 46%) of those who chooses L_1 over L_2 now reverses to L_2^0 over L_1^0.

Interestingly, for another, non-extremal correlation $\hat{\pi}^0$ using the same pair of marginal lotteries, subjects exhibit a similar preference pattern as for the classical one. This is also consistent with our theory: binary choice switches are most likely observed at extremal correlations.

π_{max}	0	24
0	66%	1%
24		
25		33%

Table 8: π^0, L_1^0 vs L_2^0, maximally correlated distribution

E.2 Extremal Correlations

We fix a pair of lotteries (p, q). As the problem is linear, the algorithm always finds an extreme point of $\Gamma(p, q)$. Denote by $\mathcal{E}(p, q)$ the set of extreme points of $\Gamma(p, q)$. Diego and Germani (1972) studies this set and shows that for $\pi \in \Gamma(p, q)$, it is an extreme point if and only if it is a set of uniqueness.[53]

Definition 16 (Set of uniqueness)**.** *For a fixed pair of marginals (p, q), a set $A \subset X \times X$ is a set of uniqueness if for any π, π' supported within A, $(\pi_1, \pi_2) = (\pi'_1, \pi'_2) = (p, q) \Rightarrow \pi = \pi'$.*

For a $\pi \in \Delta(X \times X)$, its associated *support graph* G_π is defined as following: each node $(i, j) \in V_\pi$ corresponds to a point $(x_i, x_j) \in \text{supp}\pi$, considered as a cross in the $n \times n$ grid corresponding to $X \times X$; two nodes are linked by an edge if (i) the two nodes are in the same row or column of the grid and (ii) they are adjacent, i.e., there are no other nodes lying between the two.

We say π is a forest if G_π is a forest, meaning it does not contain a cycle. If $\pi \in \Gamma(p, q)$ is a forest, then values of each $\pi(x, y)$ can be determined uniquely according to the marginals (p, q) by starting from the roots to the leafs. The Lemma below shows that such forest joint densities are vital in our analysis.

Lemma 4 (Diego and Germani (1972), Theorems 1 and 2)**.** *$\pi \in \mathcal{E}(p, q)$ if and only if $\pi \in \Gamma(p, q)$ and π is a forest.*

Though perhaps based on different reasoning and foundations, there is a tendency to adopt extreme correlations observed in research studying the robustness of correlation reference and the corresponding laboratory tests. For example, all of Lanzani's numerical examples are forests, and notably among them is the correlated Allais paradox example. As a motivation for Lanzani (2022), Bordalo et al. (2012) test whether DM's Allais behavior is dependent on the level of correlation. Most joint densities in their experimental tests are also forests.

A forest is a tree if it is connected, or equivalently, it is not a proper subset of another forest. Then, by well know results in graph theory, π is a tree if and only if $|V_\pi| = 2n-1$. A joint density supported on a forest is similar to a perfectly correlated joint density given the non-existence of cycles, while a perfectly correlated joint

[53]Sets of uniqueness also play a part in He et al. (2022) for the study of private information structures.

density in the current terms is supported on a forest with exactly n disconnected nodes. For the sake of testing DM's robustness to correlation, it is sufficient to focus on extreme points supported by trees, the reason being that a generic pair (p, q) always admits at least one such extreme point.[54]

Lemma 5 (Diego and Germani (1972), Theorems 6 and 7)**.** *For a generic pair (p, q), $\pi \in \mathcal{E}(p, q)$ if and only if (i) it is a tree and (ii) $p(X^1) + q(Y^1) < 1$ for all non-empty X^1, Y^1 satisfying $(X^1 \times Y^1) \bigcap supp\ \pi = \varnothing$.*

Diego and Germani (1972) also provides a procedure for finding all trees corresponding to the extreme points of $\mathcal{E}(p, q)$ in their Section 6.B.

References for Online Appendices

Becker, G. S. (1973). A theory of marriage: Part I. *Journal of Political Economy 81*(4), 813–846.

Bertsekas, D. P. and D. A. Castanon (1989). The auction algorithm for the transportation problem. *Annals of Operations Research 20*(1), 67–96.

Bertsimas, D. and J. N. Tsitsiklis (1997). *Introduction to linear optimization*, Volume 6. Athena scientific Belmont, MA.

Cesa-Bianchi, N. and G. Lugosi (2006). *Prediction, learning, and games.* Cambridge University Press.

Crawford, V. P. and E. M. Knoer (1981). Job matching with heterogeneous firms and workers. *Econometrica 49*(2), 437–450.

Dantzig, G. B. (1951). Application of the simplex method to a transportation problem. In *Activity Analysis and Production and Allocation.*

Daskalakis, C., A. Deckelbaum, and A. Kim (2015). Near-optimal no-regret algorithms for zero-sum games. *Games and Economic Behavior 92*(C), 327–348.

Demange, G., D. Gale, and M. Sotomayor (1986). Multi-item auctions. *Journal of Political Economy 94*(4), 863–872.

Diego, A. and A. Germani (1972). Extremal measures with prescribed marginals (finite case). *Journal of Combinatorial Theory, Series A 13*(3), 353–366.

He, K., F. Sandomirskiy, and O. Tamuz (2022). Private private information. Working Paper.

Herstein, I. N. and J. Milnor (1953). An axiomatic approach to measurable utility. *Econometrica 21*(2), 291–297.

[54] Suppose an extreme point $\pi \in \mathcal{E}(p, q)$ is not a tree, and its associated support graph has (at least) two connected components. Pick any one of them as G^1_π and let X^1, Y^1 denote the sets of those x_i and y_j such that there exist some nodes of G^1_π of the form (i, j') or (i', j). Then, π puts zero mass on nodes in $X^1 \times \bar{Y^1} \cup \bar{X^1} \times Y^1$ where the bar represents the complement set. This implies $\sum_{i \in X^1} p_i = \sum_{j \in Y^1} q_j$, and hence such (p, q) are non-generic. If we indeed run into one such (p, q), then we can simply replace it with another generic (p', q') in its neighborhood by perturbation.

Kim, E. D. (2010). Geometric combinatorics of transportation polytopes and the behavior of the simplex method. arXiv preprint arXiv:1006.2416.

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly 2*(1-2), 83–97.

Lehmann, E. L. (1966). Some concepts of dependence. *The Annals of Mathematical Statistics 37*(5), 1137–1153.

Lin, X. and C. Liu (2024). Credible persuasion. *Journal of Political Economy 132*(7), 000–000.

Rochet, J.-C. (1987). A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics 16*(2), 191–200.

Shapley, L. S. and M. Shubik (1971). The assignment game I: The core. *International Journal of Game Theory 1*, 111–130.

Villani, C. (2009). *Optimal transport: old and new*, Volume 338. Springer.

Vohra, R. V. (2011). *Mechanism Design: A Linear Programming Approach*, Volume 47. Cambridge University Press.