Correlation Preference

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

August 26, 2024

Abstract

We propose a correlation utility (CU) representation of correlation preference without requiring transitivity nor completeness. Under a *correlation independence* axiom, CU specializes to *correlation expected utility* (CEU) which is not compatible with the *extended* Allais paradox. This motivates our *correlation betweenness* and *correlation projective independence* axioms, which characterize *correlation weighted* utility (CWU). In the absence of correlation sensitivity, CEU reduces to EU while CWU reduces to *skew-symmetric bilinear* utility which reduces further to *weighted utility* under transitivity. Finally, we characterize *correlation probabilistic sophistication*, subsuming two major directions of generalization of SEU: maintaining Savage's Postulate 2 without transitivity, and vice versa.

[*]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 in a preference aggregation setting. (See, e.g., Brandl et al. (2016); Brandl and Brandt (2020).)

The Condorcet paradox is considered "the root cause for central impossibility theorems in social choice theory such as Arrow's Theorem or the Gibbard-Satterthwaite Theorem" (Brandt, 2017). At the same time, there does not appear to be works in the literature on mechanism design without requiring transitivity for mechanism participants. 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 players with the ability to perceive strategy recommendations potentially *correlated* with each other through opponents' strategies. Subsequently, Aumann (1987) recast correlation equilibrium as EU maximization in line with Savage (1954).

Motivated by the Allais (1953) paradox which inspires much of the subsequent development of transitive NEU models, Fishburn (1982) proposes the nontransitive model of *skew-symmetric bilinear* utility (SSB) which is determined by a skew-symmetric bilinear[1] kernel $\psi(\cdot,\cdot)$ on pairs of lotteries over outcomes X: the decision maker weakly prefers one lottery $p \in \Delta X$ over another q if $\psi(p,q) \geqslant 0$; when the set of outcomes is finite, $\psi(p,q)$ can be written as $\sum_{x,y\in X} p(x)q(y)\psi(x,y)$ where x, y stand for the corresponding degenerate lotteries. The SSB utility has been applied extensively to social choice theory where individual and social preferences are not necessarily transitive. In particular, the *maximal lottery* (Fishburn, 1984), a *probabilistic* social choice function (SCF), is guaranteed to exist for an SSB utility by applying the minimax theorem due to von Neumann (1928). Moreover, maximal lottery is the *unique* (probabilistic) SCF satisfying two influential axioms: population-consistency and composition-consistency (Brandl et al., 2016).[2] 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 (PC), where the SSB kernel $\psi(x,y)$ only takes one of the three values $\{0,1,-1\}$. Intuitively, p is preferred to q for PC 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.[3]

The SSB utility resembles the payoff in a zero-sum game where a row player chooses a mixed strategy p over actions $x \in X$ and the opposite column player chooses q. Then, one might contemplate Aumann's idea of correlated equilibrium in this setting by replacing $p(x)q(y)$ with a joint density $\pi(x,y)$, where $\pi \in \Delta(X\times X)$. Two papers attempt to extend SSB utility to pairs of correlated lotteries. Fishburn (1989) accomplishes this in a Savagean setting leading to a skew-symmetric additive utility (SSA), where the incomprehensive consideration on joint densities is also reflected when he describes the relation between SSB and his newly proposed SSA as "both parent and child". Recently, the problem is revisited in the risk setting when

[1]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 $\psi(r,\alpha p+\beta q) = \alpha\psi(r,p)+\beta\psi(r,q)$.

[2]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.

[3]In individual choice, the validity of transitivity assumption has also been in doubt given similar problems such as the *intransitive dice* (Gardner, 1970) and its original and more formal form as a statistical paradox (Steinhaus and Trybula, 1959); the problem has been followed by literature in applied statistics on voting paradox, (e.g., Usiskin (1964), Blyth (1972)), which is closely connected to the idea of PC.

Lanzani (2022) axiomatizes regret theory (Bell, 1982; Loomes and Sugden, 1982) and salience theory (Bordalo et al., 2012) in the form of comparing marginals of *joint densities* π taken from $\Delta(X \times X)$, as opposed to lottery pairs $(p, q) \in \Delta X \times \Delta X$.[4]

From the emerging literature on nontransitive preferences, we seek to distill and develop a theory of choice towards what Arrow suggests, in the epigraph, as "cannot be foreseen". To this end, we characterize a general utility representation for binary choice through a form of continuity without requiring transitivity nor completeness. We begin by incorporating correlation sensitivity to Samuelson's four-lottery version of the independence axiom. This yields a characterization of the *correlation expected utility* (CEU) in which utility difference $\phi(x, y)$ is defined on pairs of outcomes and then taken expectation over joint density π: the decision maker (weakly) prefers the row marginal lottery to the column lottery *under* π if $\mathbb{E}_\pi \phi$ is nonnegative. Notably, CEU encompasses the nontransitive models of regret theory, salience theory and Lanzani's correlation sensitive representation, which corresponds to CEU under skew-symmetry of the bivariate utility function. Without skew-symmetry, CEU includes additionally an extension of the *expectations-based reference dependent* (ERD) model to setting of correlation preference.

	1/3	1/3	1/3
A	l	h	m
B	h	m	l

Table 1: SML with 3 outcomes $l < m < h$.

As example of correlation preference, consider the following binary choice problem (Loewenfeld and Zheng, 2023) bearing some resemblance to the Condorcet paradox. Subjects are asked to choose between a pair of *same-marginal lotteries* (SML) A and B, each with three equally likely outcomes (Table 1). Loewenfeld and Zheng (2023) find a significant majority of subjects exhibit preference for A over B, even after a premium is added to B so that it first-order stochastically dominates A. Such pattern can be accommodated in a simple CEU model, but not in models ignoring correlation between lotteries.[5] In the absence of correlation sensi-

[4]We will revisit this issue in Section 6 where we axiomatize a correlation counterpart of probabilistic sophistication.

[5]As discussed in Loomes and Sugden (1987) and Starmer (2000), imagine a third alternative C paying (m, l, h) for the three states: if A is strictly preferred to B, then it would be reasonable for B to be strictly preferred to C, and C to A, giving rise to a preference cycle. Note that the preference for A over B is *opposite* to the prediction of regret theory; see detailed discussion on the experiment in Section 3.

tivity, CEU reduces to *transitive* expected utility. The corresponding reduction of skew-symmetric CEU to EU boils down to the weaker requirement of for correlation insensitivity over SMLs.

Despite its greater flexibility, CEU cannot account for the range of reported choice behaviour in the Allais common-consequence problem (Allais (1953); see subsection 3.2) which is traceable to an exchange between Allais and Savage during the 1952 Paris Colloquium and revisited in Savage (1954). The literature documents a wide range of proportions of EU violating behavior even when the Allais gambles are presented in a *correlated* manner.[6] Recently, Frydman and Mormann (2018) 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 choose from the two pairs of correlated lotteries $L_1^z = (\$25, 33\%;\ z, 66\%;\ 0, 1\%)$ versus $L_2^z = (\$24, 34\%;\ z, 66\%),\ z \in \{0, \$24\}$, under zero, intermediate, and maximum correlation illustrated in Table 2.[7]

$\pi_\perp$	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 2: Joint densities used in Frydman and Mormann (2018) when $z = 0$, with zero (stochastically independent), intermediate, maximal correlation.

A significant *decrease* in rates of Allais type reversal is observed as the level of correlation increases (Table 3).[8] Despite being correlation sensitive, CEU cannot account for the significantly positive rate of Allais behavior as it satisfies sure-thing principle: common-consequence effect is *uniform* with $\phi(x, x) = 0$ required for all $x \in X$.[9] The pattern is also incompatible with any correlation insensitive utility model, including SSB and all transitive NEU models (see discussions in Section 3).

To accommodate the above *extended* Allais paradox, we formulate a *correlation betweenness* axiom to axiomatize a non-CEU preference, called *correlation between-*

[6]Reported proportions range from 33-36% in Starmer (1992), 20% in Bordalo et al. (2012), 14-36% in Ostermair (2022), 19-40% in Esponda and Vespa (2023),40% in Humphrey and Kruse (2023).

[7]Scaled by 1:100 from the values in Problems 1 and 2 of Kahneman and Tversky (1979).

[8]Decreasing rates of Allias behavior from "independent" to "maximally correlated" are also documented in Bruhin et al. (2022) and Loewenfeld and Zheng (2024) for other probability and outcome parameters.

[9]If we were to allow for non-zero $\phi(x, x)$, we would have strict preference at the degenerate choice problem $\delta_{(x,x)}$ under completeness.

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 3: Allais behavior decreases with correlation.

ness utility (CBU), which, under correlation insensitivity and transitivity, reduces to betweenness conforming utility (Dekel, 1986; Chew, 1989). In conjunction with an alternative weakening of correlation independence, called *correlation projective independence*, CBU specializes to a representation, called *correlation weighted utility* (CWU), given by $\mathbb{E}_\pi \phi + \psi(p, q)$ which comprises a CEU kernel defined on π and a SSB kernel defined on the row and column marginals p and q of π.

		Correlation Sensitivity	
		Insensitive	Sensitive
Linearity of representation in (joint) probability	Linear	EU	CEU - SML preference
	Bilinear	SSB/WU - indep. Allais	CWU - extended Allais

Table 4: Correlation sensitivity, linearity, and Allais behavior.

Table 4 summarizes how different generalizations of EU can account for the experimental evidence in the literature relating to SML and Allais behavior. Under correlation insensitivity, while SSB and WU can both exhibit Allais behavior, the rates of violations do not depend on the degree of correlation. Although CEU can exhibit SML preference, it is incompatible with Allais behavior in correlated common-consequence problem. CWU is the only model which can both exhibit SML preference and account for the extended Allais paradox. The specific arguments are discussed in Section 3.

2 Preview of Main Findings

In binary choice, 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. Although we are cognizant of the potential incidence of correlation between any pair of lotteries, this is typically ignored in the presence of transitivity part of the decision maker's preference make up.

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 $\Delta(X \times X)$. For a joint density $\pi \in \Delta(X \times X)$, illustrated in Table 5 below, we adopt the convention $\pi \in \Pi$ if the decision maker weakly prefers receiving its row marginal $\pi_1 = p$ paying x_i with probability $p(x_i)$ than receiving its column marginal $\pi_2 = q$ paying y_j with probability $q(y_j)$, where $\Pi \subset \Delta(X \times X)$ comprises all joint densities π for which the row marginal π_1 is weakly preferred to the corresponding column marginal π_2.

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

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 $\succeq^\pi$ in terms of the preference set $\Pi \subset \Delta(X \times X)$: we denote by $p \succeq^\pi q$ the decision maker's (nonstrict) preference for "p over q *under correlation* π" with p and q being the row and column marginals of π. We refer to both $\{\succeq^\pi\}_{\pi \in \Pi}$ and Π as the decision maker's *correlation preference.*

2.1 Correlation Utility Representation

We offer a general representation theorem for binary choice which can accommodates correlation between lotteries. A particularly simple form of correlation utility is given by an expectation utility form: $U(\pi) = \mathbb{E}_\pi \phi = \sum_{(x,y) \in supp\ \pi} \phi(x, y)\pi(x, y)$. 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 value to Π and negative value to its complement suffices.

The following notations will come in handy. 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. The table below explains the sets with the corresponding equivalent notations. 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, two arbitrary points in $\tilde{\Pi}$ do not satisfy such an equivalence relation. Let δ_x stands for

$\pi \in \cdot$	Π	$\hat{\Pi}$	$\tilde{\Pi}$	$\check{\Pi}$
p vs. q	$p \geq^{\pi} q$	$p >^{\pi} q$	$p \sim^{\pi} q$	$p \ngeq^{\pi} q$

Table 6: Equivalent notations for correlation preference.

the degenerated probability density putting probability mass 1 at the point x alone. In the sequel, we maintain the following assumption throughout the paper.

Assumption 0. *The preference* Π *is*

(i) *non-trivial: neither the complement* $\check{\Pi}$, *nor* Π°, *the interior of* Π *(relative to the probability simplex* $\Delta(X \times X)$*) is empty;*

(ii) *pointwise reflexive: each degenerate joint lottery* $\delta_{(x,x)}$ *lies on the boundary of* Π, *i.e.,* $D = \{\delta_{(x,x)} : x \in X\} \subset \Pi \backslash \Pi^{\circ}$.

Intuitively, pointwise reflexivity requires that the decision maker *marginally* prefers the degenerate lottery x to itself, in the sense that a slight perturbation away from it *could* change the decision maker's preference. Mathematically, this is saying for any $x \in X$, $\delta_{(x,x)} \in \Pi$ and for any of its neighbourhood N, $N \bigcap \check{\Pi} \neq \varnothing$.

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 in which 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 $p \succeq q \iff \Gamma(p, q) \subset \Pi$: the decision maker 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)**.** *A correlation preference* Π *exhibits correlation insensitivity if 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, q) regardless of their correlation. Such a decision maker's preference at one correlation of (p, q) pins down the preference at any other correlation of (p, q).

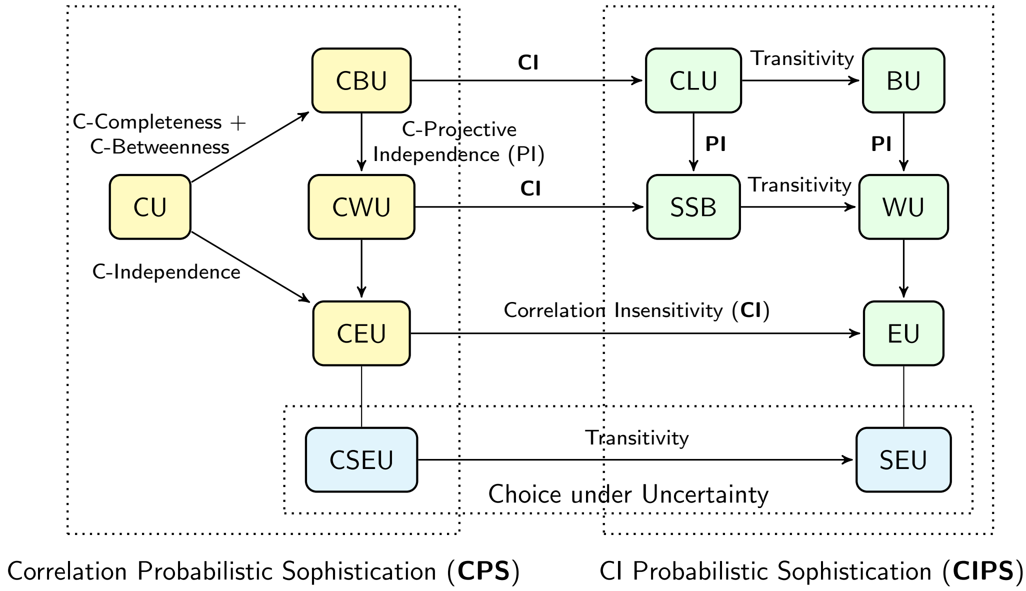
Remark 1. *As a corollary of Proposition 0, for correlation insensitive preference* Π *(and the induced* $\succeq$*), there exists a continuous* $\psi : \Delta X \times \Delta X \to \mathbb{R}$ *such that* $p \succeq q \iff \psi(p, q) \geqslant 0$*. Fishburn's SSB utility is a specialization with* ψ *being both skew-symmetric and bilinear in its arguments* (p, q)*.*

The correlation insensitive preference is further complete and transitive if and only if it admits a classical transitive utility : 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)). As we will show in Section 4, CEU will reduce to EU under correlation insensitivity.*

The following weaker notion of correlation insensitivity restricts to same-marginal lotteries (SML, Loewenfeld and Zheng (2023)) which will be further discussed in Section 3.

Axiom 1* (SML Correlation Insensitivity)**.** *A correlation preference* Π *exhibits SML correlation insensitivity if for all* $p \in \Delta(X)$*,* $\Gamma(p, p) \subset \Pi$*.*

If a decision maker 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 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.



Correlation Utility (CU)			
Correlation Expected Utility	CEU	CSEU	Correlation SEU
Correlation Betweenness Utility	CBU	CLU	Conditional Linear Utility
Correlation Weighted Utility	CWU	SSB	Skew-Symmetric Blinear Utility

Figure 1: Relation among Axiomatic Preference Models.

Figure 1 illustrates the interrelations between the transitive models on the right and their correlation counterparts on the left. The next two sub-sections will preview CEU and CWU before they are formally discussed in Sections 4 and 5 after section 3 on experimental evidence. The relations between CPS and CIPS and the other models will be developed and discussed in Section 6.

2.3 Correlation Expected Utility

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

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

The kernel $\phi(x, y)$ may be interpreted as the utility of receiving x while foregoing

y. Notice that linearity of the CEU representation implies continuity of the correlation preference Π and its representation. Specific restrictions can be imposed on the CEU kernel ϕ corresponding to respective behavioral properties. For example, by *pointwise reflexivity*, we have $\phi(x,x) = 0$ for all $x \in X$. In the following we assume the CEU representation is *non-trivial*: there exists $(x,y),(x',y')$ such that $\phi(x,y) > 0$ and $\phi(x',y') < 0$.

We say the CEU preference is *symmetric* if ϕ is skew-symmetric, i.e., $\phi(x,y) = -\phi(y,x)$, as the decision maker experiences the same level of utility change whether receiving x in foregoing y or vice versa. Symmetry of CEU preference implies $\mathbb{E}_\pi \phi > 0 \iff \pi \in \hat{\Pi}$,[10] 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* either through taking the average (Bordalo et al., 2022) or the maximum over other candidate lotteries (Quiggin, 1994).

The CEU class includes the following important utility models which we will refer to subsequently.

- Expected Utility (EU) model: $\phi^{EU}(x,y) = u(x) - u(y)$.
- Regret Theory (RT) by Bell (1982) and Loomes and Sugden (1982): ϕ^{RT} is often assumed to satisfy *regret aversion* $\phi(x,y) > \phi(x,z) + \phi(y,z)$ for all $x > z > y$.
- Salience Theory (ST) by Bordalo et al. (2012): $\phi^{ST}(x,y) = \sigma(x,y)(x-y)$ with the *salience function* σ satisfying a set of *salience properties* (their Definition 1).[11]
- The correlation sensitive representation by Lanzani (2022): This is the symmetric CEU model.
- Expectations-based-Reference-Dependent (ERD) model from Kőszegi and Rabin (2007): $\phi^{ERD}(x,y) = (\lambda \cdot \mathbf{1}_{x>y} + \mathbf{1}_{x \leqslant y})(x-y)$, where $\lambda > 1$ captures decision maker's aversion to the *anticipated* loss by switching from x to y when $x > y$. Since $\lambda > 1$, ϕ^{ERD} is not skew-symmetric.

Theorem 2 in Section 4.3 relates CEU and EU through correlation sensitivity

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

[11] Though with different behavioral/psychological fundamentals, ST representation also satisfies the regret aversion condition. An example would be $\sigma(x,y) = \frac{|x-y|}{|x|+|y|+\theta}$ as provided in the original paper. Herweg and Müller (2021) propose a generalized RT and demonstrate that it includes ST as a special case.

without involving directly transitivity: for a correlation preference 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.

Theorem 2 also implies that SML correlation insensitivity is equivalent to *full* insensitivity for symmetric CEU under correlation completeness. In conjunction with Theorem 1, Theorem 2 can be seen as providing an alternative proof of the Expected Utility Theorem.

2.4 Correlation Weighted Utility

The Allais (1953) paradox has provided the primary impetus for the development of transitive NEU models since Kahneman and Tversky (1979). 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 (Weighted Utility)**.** *The preference relation admits a weighted utility representation if there exist continuous* $u : X \to \mathbb{R}$ *and weight function* $w : \Delta X \to \Delta X$*, such that* $p \succeq^{\pi} q \iff \mathbb{E}_{w[p]}u > \mathbb{E}_{w[q]}u$ *, where* $w[p](x) = \sum \frac{\tilde{w}(x)p(x)}{\sum \tilde{w}(x')p(x')}$ *for some continuous* $\tilde{w} : X \to \mathbb{R}_{>0}$.

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

Definition 3 (SSB Utility)**.** *The preference relation admits an SSB representation if there exists a continuous and skew-symmetric bilinear* $V : \Delta X \times \Delta X \to \mathbb{R}$ *such that* $p \succeq^{\pi} q \iff \psi(p, q) > 0$.

Nakamura (1990) points out that SSB implies a characteristic property of WU called *ratio consistency*,[12] which we restate in Online Appendix B where we offer proofs of our main results relating to correlation betweenness. To develop an ac-

[12]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 $\frac{\gamma_1/1-\gamma_1}{\beta_1/1-\beta_1} = \frac{\gamma_2/1-\gamma_2}{\beta_2/1-\beta_2}$.

count of the *extended* Allais paradox, we observe that SSB has a central role in being both correlation insensitive and reduces to WU under transitivity. This points to the following correlation weighted utility proposed formally in Section 5.2.

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

In Section 5.2, we will show that for a correlation preference Π that admits a CWU representation on $\Delta(X \times X)$, it is correlation insensitive if and only if it admits an SSB representation. 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 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. This observation is however not compatible with the equivalence between correlation-insensitive CEU and transitive EU in our Theorem 2 as well as the axiomatic foundation of SSB (Fishburn, 1982) reducing to WU under transitivity (Fishburn, 1983).[13]

A CWU decision maker combines a correlation-sensitive symmetric CEU kernel on joint densities with a correlation-insensitive SSB kernel on pairs of marginals. While SSB containing WU can exhibit correlation-insensitive Allais behavior, CEU can exhibit correlation-sensitive Allais behavior except in the case of maximal correlation. We demonstrate in Section 3 that the resulting CWU combining these two kernels can exhibit correlation sensitivity in Allais behavior including the case of maximal correlation thereby accounting for the extended Allais paradox.

In terms of parametric form, as an example, we can adopt the *double exponential* specification of WU in Chew and Tan (2005) as transitive SSB kernel $\psi(x, y) = w(x)w(y)[u(x) - u(y)]$ with $w(x) = e^{\rho x}$, $u(x) = -e^{-\lambda x}$, delivering correlation insensitive Allais behavior. At the same time, the ST kernel

[13]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., a belief as correlation $\pi \in \Delta(X \times X)$) and binary utility, 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.

$\phi(x,y) = \sigma(x,y)(x-y)$ with $\sigma(x,y) = \frac{|x-y|}{|x|+|y|+\theta}$ in Bordalo et al. (2012) exhibits correlation sensitive Allais behavior except at maximal correlation. Chew (1983) has derived the Arrow-Pratt measure of local risk attitude for WU as comprising an EU term $-u''/u'$ plus a weighted term $-2w'/w$. The corresponding Arrow-Pratt measure for CWU also comprises a CEU term and a SSB term 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 $\bar{\phi}_{22}$.[14] Interestingly, now the ϕ term is dominated by ψ in the Arrow-Pratt measure: $A(y) = -\frac{\psi_{22}(y,y)}{\psi_2(y,y)} = \lambda - 2\rho$ is the same as in Chew and Tan (2005).

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

Our CEU preference requires that a decision maker *linearly* evaluate a joint density, i.e., $\pi \in \Pi$ if and only if $\mathbb{E}_\pi \phi \geqslant 0$. Recently, Loewenfeld and Zheng (2023) test a key assumption in 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 $\phi(l,h) + \phi(h,m) + \phi(m,l) < 0$. They devise a same-marginal lottery (SML) test where subjects are asked to choose between two correlated lotteries with the same marginal. The following properties are closely related to the SML test. We say the decision maker is *increasingly* (resp. *decreasingly*, *constantly*) *sensitive to payoff difference*, denoted ISPD (resp. DSPD, CSPD) if $\phi(l,h) + \phi(h,m) + \phi(m,l) <$ (resp. $>, =$) 0 for $l < m < h$.

Probability	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
DSPD	l	h	m
ISPD	h	m	l

Table 7: ISPD vs. DSPD, SML with 3 outcomes $l < m < h$.

One can easily verify that EU is CSPD, RT and ST are both ISPD while ERD is

[14]Notation ϕ_2 means the partial derivative of the second marginal of ϕ, and similarly for others.

DSPD. Earlier results, mostly for testing the RT and ST theories, report evidence with choice patterns supporting the ISPD assumption (see Loewenfeld and Zheng (2023) for a complete reference.) Loewenfeld and Zheng (2023) ask each subject to choose one from the following two SML, where the lotteries are named so that the ISPD (resp. DSPD) lottery will be chosen if the subject's preference satisfies ISPD (resp. DSPD). The experimental results confirm the existence of SML correlation sensitivity (37.4%) in their subjects' choice behavior and thereby rejects EU as well as SML correlation insensitivity of any symmetric CEU preference by Theorem 2. Moreover, their evidence also shows that choices display moderate DSPD while finding no ISPD pattern at the individual level, thereby rejecting the ISPD assumption in regret and salience theory. The results could be interpreted either as supporting DSPD (rejecting ISPD) under skew-symmetry, or rejecting skew-symmetry under ISPD. As our axiomatization covers both symmetric and asymmetric CEU preferences, it is compatible with the evidence from both directions and calls for further experiments to distinguish the two possibilities.

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 the Table 8.[15] He then describes a thought process involving the application of a con-

[15]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.

	Ticket Number Drawn		
	1	*2-11*	*12-100*
Gamble 1	5	5	5
Gamble 2	0	25	5
Gamble 3	5	5	0
Gamble 4	0	25	0

Table 8: Savage's presentation of Allais paradox, prizes in units of $100,000.

tingent reasoning principle, called the sure-thing principle (STP).[16] 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 (see Table 9 below), they report a declining rate of Allais behavior from "independent" (uncorrelated) to "intermediate" to "maximally correlated" (see Table 3 in Introduction), 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 implying no violation as well as any transitive NEU (PS) preference implying uniform rates of violation.

As previewed in Section 2.4, a CWU decision maker can be seen as evaluating joint densities linearly, but with different thresholds $\psi(p, q)$ when the marginals (p, q) change, thus account for *extended* Allais-type behavior. Let $\phi(x, y) = \sigma(x, y)(x - y)$ where $\sigma(x, y) = \frac{|x-y|}{|x|+|y|+\theta}$ as in Bordalo et al. (2012), and $\psi(x, y)$ be a

[16] "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)

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 9: Correlation between lotteries increasing with β

WU kernel. Then, $\partial U(\pi^0)/\partial\beta = -24(2\theta^2+123\theta+1200)/(\theta+24)(\theta+25)(\theta+49) < 0$, so that the chance of choosing L_1^0, and hence the rate of exhibiting Allais behavior, decreases with correlation β. Meanwhile, a non-degenerating ψ kernel would allow for a non-zero rate when there is maximal correlation $\beta = 1$.

4 Correlation Expected Utility

Our development of *correlation expected utility* is closely related to Lanzani's *correlation-sensitive representation* which parallels Fishburn's (1989) axiomatization of *skew-symmetric additive* (SSA) utility[17] representation in a Savagean setting (CSEU). Distinct from Lanzani (2022), our axiomatization does not rely on completeness nor transitivity while convexity continues to play a key role.

4.1 Correlation Independence

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 $\alpha \in (0,1)$.[18] Below is the correlation counterpart to this independence axiom.

Axiom 2 (Correlation Independence)**.** *For any* $\alpha \in (0,1)$,

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

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

[17]In a Savagean framework, let S, X, F be the sets of states, outcomes, and acts with $F \subset X^S$. Fishburn axiomatizes a preference relation $\succ$ on F represented by a probability measure μ on S and a skew-symmetric $\varphi : X \times X \Rightarrow \mathbb{R}$ such that $f \succ g \iff \int_S \varphi(f(s), g(s))d\mu > 0$

[18]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.

The above may be stated, equivalently, in a form that resembles closely the original four-lottery independence axioms: $p \succeq^{\pi} (\not\succeq^{\pi})\ q$ and $p' \succeq^{\pi'} (\not\succeq^{\pi'})\ q'$ implies $p_\alpha \succeq^{\pi_\alpha} (\not\succeq^{\pi_\alpha})\ q_\alpha$, 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.[19]

We further refine correlation independence 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 $p_\alpha \succ^{\pi_\alpha} (\not\succeq^{\pi_\alpha})\ q_\alpha$. The axiom below is a restatement of what is proposed in Lanzani (2022) by the same label.

Axiom 2* (Correlation 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) $\pi \in \check{\Pi} \bigcup \tilde{\Pi}, \pi' \in \check{\Pi} \Rightarrow \alpha\pi + (1-\alpha)\pi' \in \check{\Pi}$.

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)**.** *A continuous correlation preference* Π *satisfies correlation independence, if and only if it admits a CEU representation.*

Proof of Theorem 1. The proof of sufficiency is immediate. To prove necessity, notice 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 $c = 0$.[20]

Suppose there exists a $\pi' \in \check{\Pi}$ such that $\mathbb{E}_{\pi'}\phi = 0$ $(*)$. Pick a $\pi \in \Pi^\circ$, the nonempty interior of Π (relative to the simplex), with $\mathbb{E}_\pi \phi > 0$ $(**)$. By continuity[21],

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

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

[21] Here we make use of a weaker version of continuity which is close to Lanzani's (2022)

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, $\mathbb{E}_{\pi}\phi \geqslant 0 \iff \pi \in \Pi$. □

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 vital 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 $\tilde{\Pi}|_{ind}$, 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 $\Pi|_{ind}$ first through spanning[22], and further by $\tilde{\Pi}|_{ind}$ 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.

4.2 Correlation Completeness

We next demonstrate how *correlation completeness* stated below together with strong correlation independence leads to the symmetric CEU representation in Lanzani (2022).

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

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 $\tilde{\Pi}$, and further we show in Online Appendix A that $\check{\Pi} = \hat{\Pi}^{\mathrm{T}} := \{\pi^{\mathrm{T}} : \pi \in \hat{\Pi}\}$.

Corollary 1 (Axiomatization of Symmetric CEU)**.** *A continuous correlation preference* Π *satisfies correlation completeness and correlation 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, $\mathbb{E}_{\pi}\phi <$

Archimedean continuity axiom. See details in Online Appendix A.

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

$0 \Leftrightarrow \pi \in \check{\Pi}$. 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 makes use of 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, CSEU (Fishburn, 1989). We will elaborate on their close connection in Section 6.

4.3 Correlation Insensitivity and Expected Utility

To facilitate our characterization of classical EU in terms of correlation insensitivity, we introduce some notions from *optimal transport* theory (OT). Recall that the correlation insensitive preference relation $\geq$ is defined as $p \geq q \iff \Gamma(p,q) \subset \Pi$, which is equivalent to $\min_{\pi \in \Gamma(p,q)} \sum_{x,y} \phi(x,y)\pi(x,y) \geqslant 0$. The Kantorovich (1942) duality delivers the following equation[23] which will bridge CEU and EU:

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

Hence, $p \geq q \iff \max_{u(x)-v(y) \leqslant \phi(x,y)} \mathbb{E}_p u - \mathbb{E}_q v \geqslant 0$.

Theorem 2 (Correlation Insensitivity and Expected Utility)**.** *For a correlation preference* 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.*

Points (i) and (ii) in Theorem 2 strengthen Lanzani's Proposition 1 in two ways:

[23] The optimal (u,v) on the RHS satisfies $v_{pq}(y) = \min_x \{\phi(x,y) + u_{pq}(x)\}$, $u_{pq}(x) = \max_y \{-\phi(x,y) + v_{pq}(y)\}$, and $v_{pq}(y) - u_{pq}(x) = \phi(x,y) \ \forall (x,y) \in \operatorname{supp} \pi^*$, where π^* is any optimal solution to the LHS.

part (i) relaxes symmetry, while part (ii) limits the examination of correlation insensitivity to SML. Both correlation insensitivity and SML correlation insensitivity appear weaker than the classical notion of transitivity.[24] Indeed, correlation insensitivity only concerns a pair of (p, q) 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 being developed in the next section. One would also see how transitivity *alone* appears a more natural property for a rank-dependence correlation preference in Section 7.2.

Theorems 1 and 2 together imply that it is impossible to have non-EU preferences that satisfy correlation independence and are fully dictated by comparing marginals. In other words, non-EU preferences such as SSB or WU, becomes incomplete if we require, following the idea of Fishburn (1988) and Machina and Schmeidler (1992), that a lottery be valued only according to its distribution.

It is apparent that (iii) implies (i) and (ii). Here we first sketch how (i) implies (iii) by invoking the Kantorovich duality. For Π admitting CEU representation ϕ, consider the following minimax problem on a finite support $X' \times X'$:

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

where $\mathcal{P} = \{(p, q) : p \succeq q\}$. By the equivalent definition, correlation insensitivity and pointwise reflexivity implies the problem has value 0. By correlation independence and correlation insensitivity, $\mathcal{P}$ 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 $(p, q) \in \mathcal{P}$. 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

[24] 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 $q \succeq^{\pi'} r \Rightarrow p \succeq r$. Under *pointwise reflexivity* and correlation independence, it implies CI and hence the reduction from CEU to EU; we defer the details to Online Appendix A.

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. A similar argument also appears in Machina and Schmeidler (1992) due to a private correspondence with Frank Gul.

As observed in subsection 2.3, Theorem 2 together with Theorem 1 can be seen as providing an alternative proof of the Expected Utility Theorem.[26] Three versions of linear duality play the main role in the establishment. *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 extending Nash Equilibrium to Correlated Equilibrium (Aumann, 1974). The idea resembles the proof by Hart and Schmeidler (1989) on the existence of correlated equilibrium and explicitly reveals 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 with expected utility form. More recently, Hara et al. (2019) axiomatize 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 C and subsequently a family of utility sets. We discuss in Appendix 7.1 a *correlation multi-utility* that extends the above ideas to correlation preference where convexity is restored. This is akin to the benefit from adapting the Samuelson (1952) four-lottery independence axiom to arrive at our correlation independence. Moreover, following Cerreia-Vioglio et al. (2015) who obtain a complete multi-utility model via the *negative certainty independence* axiom introduced

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

[26]Under CI, we can induce a correlation preference $\Pi^{\succeq}$ from an uncorrelated preference $\succeq$. The EU axioms (e.g., Mas-Colell et al. (1995)) *ordering* and *independence* would imply our assumptions and correlation independence/correlation strong independence, while *continuity* implies correlation continuity. Therefore, Theorem 1/Corollary 1 leads to a CEU representation and then EU is given by Theorem 2.

by (Dillenberger, 2010), we discuss in Section 7.1 a correlation extension of their *cautious expected utility* and negative certainty independence.

4.4 Asymmetric CEU Preference

One may wonder whether SML correlation insensitivity in Theorem 2 suffices to guarantee the reduction to EU without requiring symmetry. The example below based on the ERD model shows that this is not the case.

Consider the ERD model with representation function proposed by Lanzani (2022): $\phi^{ERD}(x,y) = (\lambda \cdot \mathbf{1}_{x>y} + \mathbf{1}_{x\leqslant y})(x-y),\ \lambda > 1$. It satisfies $\phi^{ERD}(x,x) = 0$, but is not skew-symmetric when $\lambda > 1$.

Recall that the function $\phi^{ERD}(x,y)$ is submodular. Then, when $q = p$, to minimize the OT primal, we should choose the *perfectly positive* correlation $\pi^{perfect} = \sum p(x)\delta_{(x,x)}$. This gives the primal objective value 0, i.e. $\min_{\pi\in\Gamma(p,p)} \sum \phi(x,y)\pi(x,y) = 0$, so the decision maker 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, which can be seen from our proof of Theorem 2.

Definition 5 (Generalized ERD)**.** *The correlation 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. *For* Π *admitting an asymmetric CEU representation with* $\phi(x,x) = 0$*, the following are equivalent:*

(i) Π *exhibits SML correlation insensitivity;*

(ii) $\succeq$ *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 $\pi' \in \hat{\Pi}$, the decision maker's preference could remain unchanged. This difference between the two independence axioms and their induced representations is akin to that which is reflected in Fishburn's SSB utility versus Nakamura's (1990) non-SSB utility; there, Nakamura also relaxes the constraint on combining pairs of lotteries across strict preference and indifferent sets, thereby resulting 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, and reference point effect.[27] One way to incorporate reference point involving correlation is discussed in Cerreia-Vioglio et al. (2024). They propose a *cautious utility* model where the decision maker 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 (Section 5.1) and projective independence (Section 5.3). The resulting axiomatizations of correlation betweenness utility and correlation weighted utility are discussed in the sequel.

5.1 Correlation Betweenness

The statement of our correlation betweenness axiom below bears some resemblance to the betweenness property in Chew (1989) requiring probability mixtures of a

[27]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.

pair of indifferent lotteries to also be indifferent each of them. The resulting planar indifference surfaces each divides ΔX into convex better-than and worse-than sets.

Axiom 4 (Correlation 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)**.** *The correlation preference* Π *admits a correlation betweenness utility (CBU) representation if for each* $p \in \Delta X$*, there exists a continuous* ϕ_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)**.** *A continuous correlation preference* Π *satisfies correlation betweenness if and only if it admits a CBU representation.*

The proof is similar to the one for Theorem 1 (details 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 $P \subset \Delta X$.

We now demonstrate that a correlation insensitive CBU reduces, under transitivity, to the class of *betweenness utility* due to Dekel (1986) and Chew (1989). To see this, suppose now the decision maker is correlation insensitive. By applying a duality argument similar to that in the proof of Theorem 2, we obtain the following result:

Proposition 2 (Conditional Linear Utility). *For* Π *admitting a CBU representation, it is correlation insensitive and complete if and only if it admits a conditional linear utility representation.*

While correlation insensitivity implies transitivity for CEU by Theorem 2, this is not the case for a CBU preference. In Appendix B, we show that under completeness and transitivity, if we assume there is a *maximal* lottery $\bar{r}$ and a *minimal* one $\underline{r}$ as in Dekel (1986) such that $\bar{r} \succeq p \succeq \underline{r}$ for all $p \in \Delta(X)$, then in a similar spirit as his original calibration, one can construct BU through implicit utility function V and the auxiliary function u (from v_p) defined as $\sum_x u(x, V(p))p(x) = V(p)$.

Proposition 3. *The correlation insensitive preference* $\succeq$ *admits a betweenness utility representation if and only if it is complete, transitive and admits a conditional linear utility representation.*

From the two propositions, we arrive at the following.

Theorem 4 (Correlation Insensitivity and Betweenness Utility). *For* Π *admitting a CBU representation, it is correlation insensitive, complete and transitive if and only if it admits a betweenness utility representation.*

5.2 Local Correlation (In-)Sensitivity

Assume that the outcome space X is finite with $|X| = N$. 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 7 (Local Strict Correlation Insensitivity). *The decision maker 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 the decision maker 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 $(p, q) \in Q$, $\mathbb{E}_\pi \phi_p + \mathbb{E}_{\pi^\top} \phi_q = 0$ 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.

Proposition 4 (Generalized Skew-Symmetry)**.** *For a correlation complete CBU preference, it 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, which is implied by Corollary 2 in Appendix B. This echoes a similar equivalence result obtained for symmetric CEU preference in Theorem 2.

Proposition 5 (SML Correlation Insensitivity)**.** *For a correlation complete CBU preference, 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 neighbourhood 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).

A key step in the overall axiomatization is to observe that correlation completeness guarantees a CWU representation in a connected component of $\mathcal{P}^S$ where the decision maker is nowhere SML insensitive.

Proposition 6 (Correlation Completeness and CWU Representation)**.** *Within a connected component* $P^S \subset \mathcal{P}^S$, *the CBU preference* Π *is correlation complete if and only if it admits a CWU representation.*

The CWU representation, unique up to multiplication by a positive constant, implies that indifference sets in each $\Gamma(p, q)$ are parallel if viewed as affine subspaces. Proposition 6 parallels Fishburn's (1982) Lemma 4, where he also derives ratio consistency and hence projective independence *directly* from betweenness and completeness for *nontransitive* triplets of marginals.[28] Notice that our previous result assumes existence of correlation sensitivity, in the absence of which we cannot

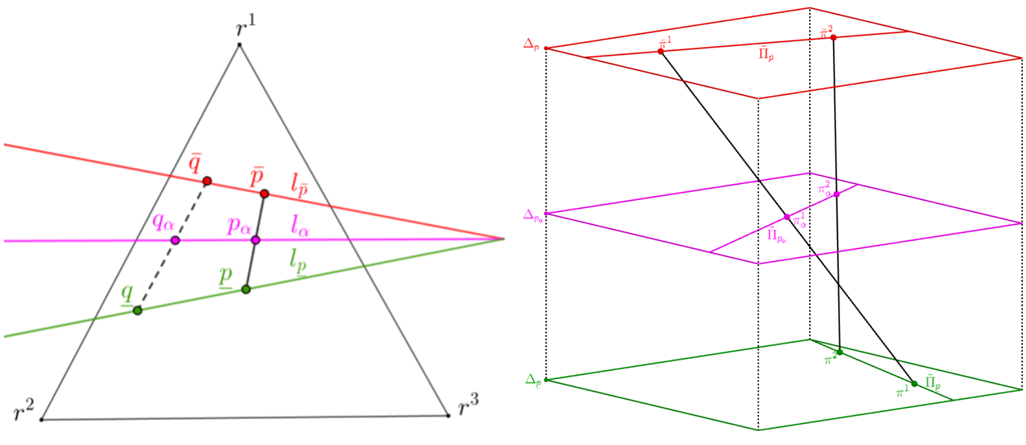
[28]This is a direct application of Ceva's Theorem in affine geometry.

invoke generalized skew-symmetry to begin with. This also corresponds to the necessity of projective independence for SSB utility (Fishburn's remaining proof) and WU when preference is *transitive*. This observation helps elucidate the tension between transitivity and correlation sensitivity. In this regard, correlation sensitivity appears to be driving nontransitive preference.

5.3 Correlation Projective Independence

Now we are ready to axiomatize the intended bilinear representation on the entire domain. First, notice that if the decision maker admits a CWU representation on a subdomain, then Π satisfies a correlation counterpart of the (uncorrelated) *projective independence* axiom (Chew et al., 1994), which requires that 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 2 below). Together with betweenness, it implies Fishburn's SSB utility and further WU under transitivity as previewed in Section 2.4.

Axiom 5 (Correlation 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* $\beta \in (0,1)$*.*



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 2: Projective independence and its correlation counterpart.

We next show that correlation projective independence (illustrated also in Figure 2) 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.

Assumption 1. *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 7. *Under Assumption 1, if the CBU preference Π satisfies correlation completeness and correlation projective independence, then it admits a CWU representation on $\mathcal{P}^S$ and an SSB utility representation 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).

Theorem 5. *Under Assumption 1, the correlation preference Π satisfies correlation betweenness, correlation completeness and correlation projective independence if and only if it admits a CWU representation (ϕ, ψ) on $\mathcal{P}^S$ and an SSB utility representation V on $\mathcal{P}^I$, that satisfy 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 the decision maker is correlation insensitive, a CWU representation reduces to SSB utility, and further to WU under transitivity given that (uncorrelated) projective independence is satisfied. This can also be viewed as the limit of a se-

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

quence 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 8 (SSB Utility and Weighted Utility)**.** *For a correlation preference* Π *that admits a CWU representation on* $\Delta(X \times X)$*, it satisfies correlation insensitivity if and only if it admits an SSB representation. It is further transitive if and only if it admits a WU representation.*

When developing his CSEU, Fishburn (1989) states that "*SSB theory is both a parent and a child of*" CSEU as, in particular, the skew-symmetric representation of SSB utility "*emerges as the specialization*" of the CSEU 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 earlier 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.

6 Correlation Preference in a Savagean Setting

As illustrated in Figure 3 below, there are two directions of research following Savage (1954 – S54). The direction on the left, initiated in Fishburn (1989 – F89), drops transitivity while maintaining Postulate 2 (P2). This leads to a characterization of a *skew-symmetric additive* representation which we label as *correlation subjective expected utility* (CSEU). On the right is the better known *(transitive) probabilistic sophistication* (TPS) direction, originating with Machina-Schmeidler (1992 – MS92), which maintains *transitivity* without imposing P2. Subsuming both directions of generalizing S54 is the present paper (CCZ24).

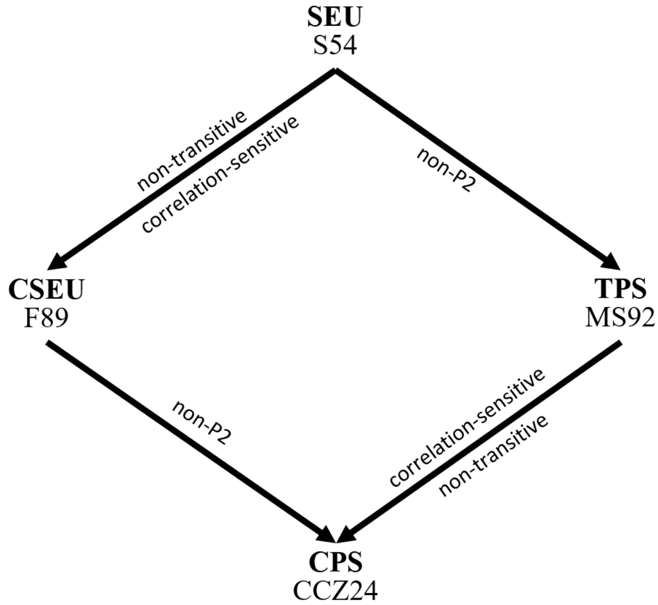


Figure 3: Parallelisms involving P2 and correlation sensitivity

Fishburn (1988. p.27) posits a decision making principle based on the following example of the two choice situations below:

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

Fishburn then states the following decision making principle which is in essence equivalent to TPS (MS92, p.747) which is shared by SEU in S54.

Distribution Reduction (DR). Lotteries which reduce to the same probability distributions are valued the same.

Under DR, choice behavior in the two situations – interdependent SI and independent SII – ought to be the same. In other words, how probabilities in different alternatives arise does not affect choice behavior. Suggesting that this implication does not accord well with intuition, Fishburn proceeds to axiomatize CSEU which is suggestive of an extension of DR principle in terms of requiring the preference between pairs of lotteries to remain the same when they arise from the same joint density. Observe that in the absence of correlation sensitivity, CPS reduces to correlation-insensitive PS as exemplified by SSB, which further reduces to WU under transitivity (Fishburn, 1983), exemplifying TPS.

6.1 Correlation Probabilistic Sophistication

We are ready to extend TPS to CPS by adapting Chew and Sagi's (2006) characterization of TPS which does not rely on continuity nor monotonicity of 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 the decision maker 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 $f \geq g$ for $(f, g) \in \Pi$ when there is no confusion in notation.

Definition 8 (Correlation Probabilistic Sophistication)**.** *The decision maker exhibits correlation probabilistic sophistication if there exists a probability measure μ on Σ such that for any two pairs of acts (f, g) and (f', g') inducing the same joint density under μ, $f \geq g \iff f' \geq g'$.*

The decision maker is said to exhibit CPS 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 9 (Correlation Event Exchangeability)**.** *Disjoint events E and E' are correlation exchangeable, denoted $E \approx E'$, if $xEx'E'f \geq yEy'E'g \iff x'ExE'f \geq y'EyE'g$ for all $x, y, x', y' \in X$ and acts f, g.*

Like its non-correlated counterpart, correlation exchangeability can be viewed as a notion of equal likelihood: if $E \approx E'$, the decision maker does not change her preference over f and g if the outcomes on E and E' are swapped. This yields the definition of *correlation 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 10 (Correlation Comparative Likelihood)**.** *For any events* E, E', $E \succsim^C E'$ *whenever there exists a sub-event* $e \subset E \backslash E'$ *such that* $e \approx E' \backslash E$.

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

Axiom 6. *Axioms C, A, and N of Chew and Sagi (2006):*

- *Completeness (C): given any disjoint pair of events* E *and* E', *either* $E \succsim^C E'$ *or* $E' \succsim^C E$.
- *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* $e \approx E \bigcup A$.

Adapting the Chew-Sagi proof slightly, we arrive at the following.

Theorem 6 (Characterization of CPS)**.** *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* $\succsim^C$, *(ii) any two events have the same measure if and only if they are correlation exchangeable, and (iii) the decision maker exhibits CPS with respect to* μ.

We omit the proof which is straightforward and somewhat repetitious. Compared with the original proof, notice that equivalence relation in Definition 8 delivers the role of transitive indifference played by Chew and Sagi (2006) in the derivation of an agreeing probability measure.

Bikhchandani and Segal (2011) study a special form of CPS called *regret-based* preference which conforms to P2[30]. For acts f, g on a probability 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 the regret or rejoicing about x over 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 CPS 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. As it turns out, the proof of Proposition 1 in Bikhchandani and Segal (2011) can be modified (see Online Appendix C) to show, as stated below, that CPS and TPS are equivalent under transitivity for any correlation-complete

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

preference in the Savagean setting.[31]

Proposition 9 (CPS and TPS)**.** *If a correlation-complete preference satisfying CPS is transitive, then it satisfies TPS.*

The above proposition shows that imposing transitivity would obliterate any influence of correlation in binary choice. Notice that correlation insensitivity serves an intermediate role between CPS and TPS as reflected in the relation between SSB and WU illustrated in Figure 1. Under CPS, 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*, a precursor to Lanzani's (2022) correlation strong independence, which he views as being complementary to P2. (The relevant definitions and statements are formally stated in Online Appendix C.) Notwithstanding the key role of completeness in F89, Fishburn's arguments do not extend to general, non-symmetric CEU. As our CPS axiomatization does not rely on P2, we can perform a similar exercise to arrive successively at CBU and CWU in the Savagean setting by applying correlation betweenness followed by adding correlation projective independence.[32]

6.2 Small Worlds

Under the Savagean perspective in which all uncertainties arise from a single grand world, the decision maker always faces correlated lotteries. Yet, this does not matter given the maintained assumption of transitivity of the preference ordering. Our development of CPS reveals how requiring a single grand world may constitute an undue limitation in modeling choice under risk and uncertainty. This leads naturally to the need to explore the implications of having multiple sources of uncertainty following the axiomatization of small worlds PS in Chew and Sagi (2008). Building on the approach taken in the preceding subsection, we can adapt the definition of correlation event exchangeability to the small worlds setting. This yields a definition of small worlds correlation comparative likelihood on which to apply the axioms of Completeness, Archimedean, and Non-satiation. This process can then deliver a context dependent CPS in which the decision maker possesses varying degrees of

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

[32]We can mechanically translate the axioms in the risk setting back to those in the Savagean setting, just as the translation for CEU and CSEU in the online appendix.

correlation sensitivity, ranging from being transitive and fully correlation insensitive to being SML correlation sensitive depending on the small world associated with the specific choice situation.

In the experimental literature, several papers[33] provide empirical support for the incidence of source preference including, in particular, familiarity bias which relates to the decision maker's identification with the source of uncertainty when evaluating a choice situation. As pointed out in Simon (1947), a sense of identity often arises from loyalty and identification with groups and organizations. In this sense, people generally possess multiple identities through memberships in different groups, e.g, family, nationality, jobs, and clubs relating to schools and hobbies. Like small worlds PS and in contrast with CPS, small worlds CPS points to the possibility of the decision maker exhibiting distinct correlation preference depending on how she may identify with the small world underpinning the choice context.

Small worlds CPS may give rise to a fresh perspective to account for 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). Such considerations may also apply to the differentiation between investment and insurance, discussed in Armantier et al. (2023), exemplifying a more general reference-dependent differentiation across gain- versus loss-oriented choice situations. Finally, it seems that labor and matching markets could provide a rich source of small worlds CPS given the inherent heterogeneity in how diverse individuals may relate to jobs and other occupational arrangements taking up large portions of their daily lives.

6.3 Correlation Ambiguity

Recently, Epstein and Halevy (2019) bring in correlation sensitivity to the study of ambiguity attitude in the setting of two-urn Ellsberg problem. By incorporating choice comparisons from simultaneous draws from both urns, they find an incremental aversion towards correlation ambiguity. Given their focus on inconsistency between correlation ambiguity and PS under transitivity, there is value of a more general investigation of correlation ambiguity building on the CPS approach taken in the present paper. As an example, consider the following thought experiment in their Section 2.

[33]See Chew et al. (2008), Abdellaoui et al. (2011), Armantier and Treich (2016), Chew et al. (2023), among others.

There are two urns numbered 1 and 2. Each urn contains two balls, each of which is either red or black; there is no more additional information about the compositions. One ball is to be drawn from each urn simultaneously, and thus 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 colours of the two balls, each giving a same prize x if being true: for instance, bet *Same* corresponds to the event of balls being in the same color, $\{R_1R_2, B_1B_2\}$, and similarly *Diff* for balls being different colors; bet R_1 denotes the event of urn 1 ball being red, $\{R_1R_2, R_1B_2\}$, and similarly B_1 denotes the opposite.

Consider the following choice pattern: $R_1 \succ \textit{Same}$, and $B_1 \succ \textit{Diff}$. As argued by Epstein and Halevy (2019), there exists no measure P on $\mathcal{S}$ such that $P(R_1) > P(\textit{Same})$, $P(B_1) > P(\textit{Diff})$, while maintaining $P(R_1) + P(B_1) = 1 = P(\textit{Same}) + P(\textit{Diff})$. They then resolve this paradox by assuming that the decision maker is averse to the ambiguity in the possible correlation between the ball compositions of two urns: bets R_1 and B_1 involves *one* urn, while *Same* and *Diff* both involves *two* urns simultaneously.

Such paradox also exists for correlation preference when there is a single grand world. Suppose the decision maker attaches the following joint probability to the bet pairs.[34] Then, preferences $R_1 \succ \textit{Same}$ and $B_1 \succ \textit{Diff}$ is equivalent to both joint densities π and $\pi^{\mathrm{T}} \in \hat{\Pi}$, which contradicts the definition of strict preference itself. On the other hand, the paradox can be resolved if we extend the framework to allow for one *correlation-insensitive* small world for one-urn bets R_1 and B_1, and another *correlation-sensitive* small world for two-urn bets *Same* and *Diff*.

	Same	*Diff*
R_1	$\pi(x, x)$	$\pi(x, 0)$
B_1	$\pi(0, x)$	$\pi(0, 0)$

Table 10: Joint probability for comparing bets R_1 and *Same*.

[34]This can be translated into the joint density of ball composition of two urns. For instance, $\pi(x, x) = Prob(R_1R_2), \pi(x, 0) = Prob(R_1B_2)$, etc.

7 Discussions

7.1 Inertia

Recall that Section 4.4 links asymmetric CEU preference to the idea of *inertia*. Here, we consider another approach to model inertia through a non-CEU correlation utility. As an example, consider the following multi-linear representation based on weakening of correlation independence.[35]

Definition 11. *The correlation preference* Π *admits a correlated multi-utility representation if there exists a closed family* $\{\phi : \phi \in \Phi\}$ *, such that*

$$\pi \in \Pi \iff \inf_{\phi \in \Phi} \sum_{(x,y) \in supp\ \pi} \phi(x,y)\pi(x,y) \geqslant 0.$$

The family Φ correspond to the *indecisiveness* of the decision maker (Dubra et al., 2004) in the sense that she compares each correlated pair with multiple ϕ, and the determining ϕ may be different at different π. For such a preference, let $U_\Phi(\pi) := \inf_{\phi \in \Phi} \sum_{(x,y) \in supp\ \pi} \phi(x,y)\pi(x,y)$. Preferences with a correlation multi-utility representation are characterized by the following convexity axiom.

Correlation 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 correlation convex if U is quasi-concave. By an immediate supporting hyperplane argument, a correlated multi-utility representation is characterized by correlation convexity.

Proposition 10. *A continuous correlation preference* Π *satisfies correlation convexity if and only if it admits a correlated multi-utility representation.*

One family 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.[36]

The correlation multi-utility model naturally links to the original expected multi-

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

[36] As an example, if U satisfies the two spherical properties in Chambers and Echenique (2020), then U is quadratic (in π) with a *positive semi-definite* quadratic term Q and linear part ϕ: $U(\pi) = Q(\pi) + \phi \cdot \pi$. 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 the decision maker'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 the inertia of the decision maker.

utility representation by Dubra et al. (2004).

Definition 12. *The correlation preference* Π *admits an expected multi-utility representation if there exists a closed family* $\{u_\phi : \phi \in \Phi\}$ *, such that* $\pi \in \Pi \iff \inf_{\phi\in\Phi} u_\phi \cdot \pi_1 - u_\phi \cdot \pi_2 \geqslant 0$*.*

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

Proposition 11. *For* Π *that admits a correlated 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 utility to a complete and transitive *cautious expected utility*, one may consider extending their *negative certainty independence* axiom—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 $U_\Phi(\pi) > 0 \iff U_\Phi(\pi^{\mathrm{T}}) < 0$.

Besides Bikhchandani and Segal (2011) and correlation dual utility above, another form of CPS is offered in Chew, Wang, and Zhong's (2024) attention theory (AT) model of attention induced correlation preference which encompasses RT as well as ST but does not belong to CEU. The AT specification takes the form of a weighted utility representation bearing some resemblance to CBU, but does not belong to that class.[37] This leaves open the question of identifying its characteristic property and the corresponding axiomatization.

[37]For an AT preference, there exists a outcome utility $v : X \to \mathbb{R}$ and a bivariate attention function function $\alpha : X \times X \to \mathbb{R}_+$ such that $\pi \in \Pi \iff \mathbb{E}_\pi\phi_\pi \geqslant 0$, where $\phi_\pi = \mathbb{E}_\pi[\alpha(y,x)]\alpha(x,y)v(x) - \mathbb{E}_\pi[\alpha(x,y)]\alpha(y,x)v(y)$.

7.2 Rank Dependence

We propose a correlation-sensitive extension of Yaari's (1987) *dual utility* representation, who axiomatizes a rank-dependent utility linear in monetary outcome. Following his notation, we now consider the monetary outcome space $X = [0, 1]$ and define the following.

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 $x = G_p^{-1}(u)$.

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 $(C_\pi, G_{\pi_1}, G_{\pi_2})$. The representation ψ below resembles the bivariate utility ϕ in CEU.

Definition 13 (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 the decision maker 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).

When each ψ^C is separable[38], 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*:

[38]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.

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 a copula C, it is natural to examine *transitivity* for marginals under this specific correlation given the above separability in utility. On the other hand, when the decision maker is correlation insensitive, the axiom is simplified into Yaari's original *dual independence* axiom, which he shows being 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 $\mathcal{C}$.

7.3 Social Preference

There is a rich literature concerned with social and risk preference.[39] For example, bringing together the theories of individual choice with risk and allocation inequality measurement, Chew and Sagi (2012) take into account the correlations across individuals and axiomatize a social preference for both ex-ante and ex-post fairness. Their social choice function takes a rank-dependent form, addressing the inequality concerns. Now, if we consider the income distribution/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.[40]

Given such understanding, we revisit the correlation rank-dependent preference f^C discussed in the previous subsection. The correlation utility function f^C takes as argument a ranking over wealth/income after detaching it from the correlation C. It is now natural to examine the preference for different correlations when the marginal rankings are fixed. For example, if f^C is linear in C, then it further satisfies the following.

Property (Dual Strong Independence). *If $\pi, \pi' \in \Pi$ and $\pi_i = \pi_i'$ for $i = 1, 2$, then for any $\alpha \in (0,1)$, the joint density π^α corresponding to the triplet $(C^\alpha, G_{\pi_1}, G_{\pi_2})$ is*

[39] See e.g., Fleurbaey (2010); Fleurbaey and Zuber (2017); Gajdos and Maurin (2004); Grant et al. (2012); Saito (2013); Miao and Zhong (2018).

[40] Recently, Zhou (2024) studies a social preference model over rankings that allows for non-transitivity. She introduces an idea of blame and gratitude revised from RT, which will naturally involve correlation when randomization over rankings is allowed.

also in Π*, where* $C^{\alpha} = \alpha C_{\pi} + (1-\alpha)C_{\pi'}$

It would be interesting to incorporate correlation and social preference into the study of game theory and mechanism design. For example, Rubinstein and Salant (2016) report evidence of individuals positively correlating her belief over others with her own action in a two-player game, a term they coined as *self-similarity*. On the other hand, designing mechanisms for agents with social preference in the sense of Chew and Sagi (2012) could bring about fresh challenge on social related issues.

7.4 Social Choice

As mentioned in the Introduction, SSB utility has been applied to the study of social choice theory for intransitive preferences. Recently, in seeking to resolve Arrow's impossibility, Brandl and Brandt (2020) investigate the maximal SSB preference domain in which there exists an *Arrovian* social welfare function (SWF) that aggregates individual preferences into social preference. They further show that the Arrovian SWF 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 (IIA). 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 the exercise for EU functions. None of these attempts incorporate correlation preference into the binary choice framework, even transitivity is dispensed with at the very beginning. Noticeably, Turunen-Red and Weymark (1999) *convexify* the domain of *independent* lottery pairs into the space of *all* joint densities between marginals, but still 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. For example, in the symmetric CEU domain, we expect Harsanyi's aggregation theorem to hold via applying the duality argument of Turunen-Red and Weymark (1999) but similar techniques would not be applicable to asymmetric CEU or CWU. Likewise, we may also study Arrovian aggregation of nontransitive preferences *à la* Brandl and Brandt (2020) in the CEU/CWU domain.

Such investigation would be a rich avenue for future research.

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.

Blyth, C. R. (1972). Some probability paradoxes in choice from among random alternatives. *Journal of the American Statistical Association 67*(338), 366–373.

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.

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*, 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., 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 (1994). The projective independence axiom. *Economic Theory 4*(2), 189–215.

Chew, S. H., K. K. Li, R. Chark, and S. Zhong (2008). Source preference and ambiguity aversion: Models and evidence from behavioral and neuroimaging experiments. In *Neuroeconomics*, pp. 179–201. Emerald Group Publishing Limited.

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.

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 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.

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

Esponda, I. and E. Vespa (2023). Contingent thinking and the Sure-Thing principle: Revisiting classic anomalies in the laboratory. accepted at The Review of Economic Studies.

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, 10). 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.

Fleurbaey, M. and S. Zuber (2017). Fair management of social risk. *Journal of Economic Theory 169*, 666–706.

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.

Gajdos, T. and E. Maurin (2004). Unequal uncertainties and uncertain inequalities: An axiomatic approach. *Journal of Economic Theory 116*(1), 93–118.

Gardner, M. (1970). The paradox of the nontransitive dice and the elusive principle of indifference. *Scientific America 223*, 110–114.

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.

Herweg, F. and D. Müller (2021). A comparison of regret theory and salience theory for decisions under risk. *Journal of Economic Theory 193*, 105226.

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), 363–391.

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.

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.

Loomes, G. and R. Sugden (1987). Some implications of a more general form of regret theory. *Journal of Economic Theory 41*(2), 270–287.

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.

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

Rubinstein, A. and Y. Salant (2016). "Isn't everyone like me?": On the presence of self-similarity in strategic interactions. *Judgment and Decision Making 11*(2), 168–173.

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.

Sernesi, E. (2019). *Linear algebra: a geometric approach*. Routledge.

Simon, H. A. (1947). *Administrative behavior: A study of decision-making processes in administrative organization*. Macmillan.

Starmer, C. (1992). Testing new theories of choice under uncertainty using the common consequence effect. *The Review of Economic Studies 59*(4), 813–830.
Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature 38*(2), 332–382.
Steinhaus, H. and S. Trybula (1959). On a paradox in applied probabilities. *Bulletin de l'Academie polonaise des sciences 7*, 67–69.
Turunen-Red, A. H. and J. A. Weymark (1999). Linear aggregation of SSB utility functionals. *Theory and Decision 46*, 281–294.
Usiskin, Z. (1964). Max-min probabilities in the voting paradox. *The Annals of Mathematical Statistics 35*, 857–862.
von Neumann, J. (1928). Zur theorie der gesellschaftsspiele. *Mathematische annalen 100*(1), 295–320.
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. accepted at the Economic Theory.

Appendices

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

A Proofs for Section 4

Proof of Theorem 2. (iii) ⇒ (i) and (ii) is easy to see.

(i) ⇒ (iii): Let $\mathcal{P} = \{(p, q) : p \geq q\}$. 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[41], we can further restrict the (u, v) in the dual constraint to a bounded set $UV \subset \mathbb{R}^{2|X'|}$. Now, $\mathcal{P}$ 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

[41] 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.

minimax theorem in the Euclidean space, we have the following

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

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 $\forall(p, q) \in \mathcal{P}$.

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): We can also use a similar OT proof here. However, the following alternative proof demonstrates the power of symmetry.

Suppose that 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 $u = v$.)

Now, we 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

Proof of Theorem 3 (CBU). 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 $\pi' \in \check{\Pi}$ 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. □

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

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

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 (CBU, 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 easily verify the uniqueness of u up to a positive affine transformation, continuity of u in the second argument, 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. $\square$

For our analysis of CWU, let $t^k \in \mathbb{R}^{N\times N}$, $k \in \{1, 2, \ldots, N\}$ 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.

Proof of Proposition 4 (GSS). The statement is trivially true for $(p, q) \in \mathcal{D}^{ind}$ as $\phi_p \cdot \pi = \phi_q \cdot \pi^{\mathrm{T}} = 0$. For a fixed pair $(p, q) \in \mathcal{D}^S$, we proceed in two steps.

Step 1. A theorem of the alternative.

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 $(p, q) \in \mathcal{D}^S$, we can find a full-support $\pi \in \Gamma(p, q)$ such that $\phi_p \cdot \pi = 0 = \phi_q \cdot \pi^{\mathrm{T}}$.[42] 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 $c \cdot p + d \cdot q = 0$, and $-a\phi_p(x, y) - b\phi_q(y, x) + c(x) + d(y) = 0$ for all (x, y). The former then implies $a > 0$ through the third line in the alternative, so we can normalize $a = 1$. The latter now becomes

$$\phi_p(x, y) + b\phi_q(y, x) = c(x) + d(y), \qquad (\dagger)$$

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 $b > 0$.

Step 2. Normalizing b to 1.

If $(p, p) \in \mathcal{D}^S$, notice that ($\dagger$) 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

[42] The set of full-support π is dense.

$c \cdot p + d \cdot p = 0$, 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 $c_{pq} \cdot p + d_{pq} \cdot q = 0$. 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 † 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, $\frac{1}{b_{pq}} - b_{qr}b_{rp} = 0$. □

Proof of Proposition 6 (CWU). For each $\tilde{p}$ in this P^S, continuity implies that there is a neighbourhood $B(\tilde{p}) \subset \Delta X$ 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[43] 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 $k = 1, 2, \ldots, K$. 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). Finally we show the uniqueness of a CWU representation (Step 3).

Step 1. Representation for a fixed p.

[43] 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.

Lemma 1. *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 $\tilde{\Pi}$, correlation betweenness implies that $q_\alpha := \alpha q + (1-\alpha)q'$, $(q, q_\alpha) \in \mathcal{D}^S$. Pick any $\tilde{\pi} \in \Gamma(q, q) \bigcap \tilde{\Pi}$. 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 Step 1 in proof of GSS, there exists $b > 0$ such that $\phi_q \cdot \pi_\alpha + b\phi_\alpha \cdot \pi_\alpha^{\mathrm{T}} = 0$, where ϕ_α is short for ϕ_{q_α}. By our selection, LHS is equal to $[\phi_q \cdot \bar{\pi} + b(1-\alpha)\phi_\alpha \pi^{\mathrm{T}}] + b\alpha\phi_\alpha \cdot \tilde{\pi}$. We denote the latter 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 $\phi_\alpha = \tau_\alpha \phi_q + \sum \lambda_\alpha^k t^k + \mu_\alpha^k \bar{t}^k$.[44] We now further simplify the expression.

First, as ϕ_α only evaluates $\pi \in \Delta_\alpha$, 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, notice that $\tau_\alpha > 0$: it cannot be 0 as $(q, q_\alpha) \in \mathcal{D}^S$, while also nonnegative by continuity and correlation completeness (betweenness on the second marginal). 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 $q_\alpha = q'$. □

Now 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 the Lemma above and obtain recursively the desired formula $\phi_p = \phi + \bar{\lambda}_p \cdot \mathbf{1} + \sum_1^N \mu_p^k \bar{t}^k$.

Step 2. Linearity in 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

[44] See e.g., Chapter 8 of Sernesi (2019). Corollary 2 below, which implies Proposition 5, is shown through a similar argument.

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 skew-symmetric representation.

We now again apply GSS to a fixed pair $(p,q) \in \mathcal{D}^S$. There exists a $b \in \mathbb{R}$ 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 this b must be common for all q as correlation sensitivity requires that ϕ cannot be written into some $c'(x) + d'(y)$. Then the GSS consistency implies that $b = 1$, 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. This completes the proof. □

The Corollary below shares a similar spirit with the Lemma in Step 1 above and implies Proposition 5.

Corollary 2. *For a correlation complete CBU preference, if $(p,q) \in \mathcal{D}^{ind}$, then 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 $p \succeq q' \iff \bar{\lambda}_p + \mu_p \cdot q' \geqslant 0$.*

Proof. As in the previous proof, pick an arbitrary $\tilde{\pi} \in \Gamma(p,p) \bigcap \tilde{\Pi}$, which is a spanning set of the affine subspace described by the linear system $[t^k \cdot \pi = p^k, \bar{t}^k \cdot \pi = p^k]$, also satisfies $\phi_p \cdot \pi = 0$ by SML insensitivity. Then there exists $\{\mu_p^k\}_{k=1}^N, \{\lambda_p^k\}_{k=1}^N \in \mathbb{R}$

such that $\phi_p = \sum \lambda_p^k t^k + \mu_p^k \bar{t}^k$. The first term is then normalized to a scalar with a same argument. □

Proof of Proposition 7 (General CWU). Given Assumption 1, the SSB representation on each connected component P^I of $\mathcal{P}^I$ follows from Fishburn's (1982) original proof and the similar open covering argument in the above proof of Proposition 6. We are now left to show the validity of uniform CWU and SSB representation across different components.

Consider two connected components P_1^S and P_2^S of $\mathcal{P}^S$, with their respectively unique CWU representations (ϕ_1, ψ_1) and (ϕ_2, ψ_2). Pick any $\bar{p} \in P_1^S$ and $\underline{p} \in P_2^S$. By continuity and SML sensitivity of $\bar{p}$, we find a open neighbourhood $\mathcal{N}$ of $\bar{p}$ such that $(\bar{p}, \bar{q}) \in \mathcal{D}^S$ for every $\bar{q} \in \mathcal{N}$. By disconnectedness between two components, the line segment $[\bar{p}, \underline{p}]$ is not contained in the union of two. We can select $\alpha \in (0, 1)$ so that p_α 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 two components.

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 $q' \in [p_\alpha, r_\alpha]$. Fix a $\bar{q} \in \mathcal{N}$ close enough to $\bar{p}$, We can select an q' close enough to p_α such that there exists a $\underline{q} \in P_2^S$ such that (1) $(\underline{p}, \underline{q}) \in \mathcal{D}^S$ and (2) $q' = \alpha\bar{q} + (1-\alpha)\underline{q}$.

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 $p_\beta \in P_2^S$. By CWU representation on P_2^S, 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 neighbourhood of $\bar{p}$. This means (ϕ_2, ψ_2) is also a CWU representation on P_1^S. 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 $\mathcal{P}^S$. The proof for unique SSB representation on $\mathcal{P}^I$ is the with a same argument. $\square$

The proof of Theorem 5 follows from the observation that for both correlation betweenness and correlation projective independence, cross-domain completeness is guaranteed by their respective validity in each subdomain.

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 $\beta\pi + (1-\beta)\pi' \in \check{\Pi}$.

Lanzani (2022) makes use of correlation Archimedean throughout his axiomatization while we apply the stronger continuity axiom 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 2. *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), $\sum_x q(x) \cdot \delta_{(x,x)} \in \Pi \Rightarrow \Gamma(q,q) \cap \Pi \neq \varnothing$. So by definition, $\Gamma(p,q) \in \Pi$. This is exactly correlation insensitivity. □

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

Corollary 3 (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, $\check{\Pi} = \{\pi^{\mathrm{T}} : \pi \notin \Pi\}$. □

B Supplementary Materials for Section 5

We discuss the geometry behind Fishburn's SSB utility.

SSB Implies Ratio Consistency. 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. □

Fishburn (1982) introduces the following weakened version of projective independence, which he terms *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 $\lambda \in (0, 1)$. Assuming betweenness, symmetry can be deduced from projective independence through a simple geometry argument, as the latter implies that the straight indifferent curves are projective. Both then results in a bilinear utility representation. In Section 5.2, we have 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: for 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[45] satisfies $v_t(\alpha p + \beta q + \gamma r) = 0$. When

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

there is a preference cycle $p \succ q \succ r \succ p$, the combination is within the simplex of $\{p, q, r\}$. Otherwise, for the transitive preference $p \succ q \succ r$, $p \succ r$, the combination falls outside of the simplex but in the extended affine set. Geometrically, this means the three (*extended*) indifference curves crossing 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 CSEU

As in the main text, let S, X, F be the sets of states, outcomes, and acts with $F \subset X^S$. Fishburn studies a preference relation $\succ$ on F (and its induced weak preference $\succeq$ and indifference $\sim$) satisfying the following set of axioms.

For $x \in X$, let x also represent the constant act with the same outcome x. For $E \subset S$, the conditional preference $f \succ_E g$ means $fEh \succ gEh$ for any h. Denote $\mathcal{N}$ as the null events as usual.

Axiom 7. *Fishburn's (1989) axioms:*

- *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 the following (i) to (iii) of Savage's original theorem, along with the CSEU 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 to the set of binary acts $\{F_{xy} : x, y \in X\}$, and borrow Savage's theorem to obtain a candidate probability measure μ. He then shows that this μ together with a calibrated φ represents the preference. Notice that CSEU satisfies our CPS as $\varphi(f(\cdot), g(\cdot))$ effectively translates binary acts the into joint densities over outcome pairs. Hence, while CSEU is set to describe nontransitive preferences over Savagean acts under uncertainty, it corresponds exactly to Lanzani's correlation sensitive representation, or equivalently, our symmetric CEU preference, in the risk domain.

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

Now, assuming the decision maker is CPS, 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 12. *Assuming CPS and Fishburn's P1* and P6*, S1* is equivalent to strong independence of the induced correlation preference.*

Proof of Proposition 12. P1* and P6*, as their original versions, implies completeness and continuity of preference. It is then easy to check that strong independence of the induced $\Pi^{\succ}$ implies Fishburn's S1* on $\succ$.

To see 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 are equal to π, π' respectively, and similarly (2) the joint densities induced by (f''', g''') *conditional on* A and A^c are equal to 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 CPS, $g''' \succ_{A^c} f'''$. 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 CPS as we already assume $f \succeq g$ and $(f, g), (\tilde{f}, \tilde{g})$ induce the same joint density π. Hence, it must be $f'' \succeq_A g''$. 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 for strong independence. □

C.3 Relating CPS to TPS

The following proof is adapted from that of Proposition 1 in Bikhchandani and Segal (2011).

Proof of Proposition 9. For acts f_0 and f_1 inducing the same outcome probability distribution, we will show that $f_0 \sim f_1$. 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 the states 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 CPS, 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 the decision maker 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 a common denominator of all these fractions. Then we can rewrite f_0 and f_1 as in Case 1 with events $T_1, \ldots, T_{\hat{n}}$ of equal probabilities.

Case 3: There exist i, j such that $\mu(S_i \bigcap S_{j'})$ is irrational. A limiting argument given Case 2 would suffice together with continuity. □

D Supplementary Materials for Section 7

This section focuses on the discussion of 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 $U(\pi) = \int_{[0,1]^2} \phi(p,q) dC^{\pi}(1 - G_{\pi_1}^{-1}(p), 1 - G_{\pi_2}^{-1}(q))$. It resembles a CEU representation with the role of marginals now played by the quantiles.

Suppose Π is correlation-insensitive. Noticing that copulas have uniform marginals, by a similar 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 would give both u and v are constant in the marginals, so that it reduces 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 compared. Among the considerable effort made in the literature to test departure from EU, an interesting direction is to assess/test the axioms where the EU theory is built on. We show in the next subsection how to check a decision maker's robustness to correlation for a fixed pair (p, q) through the famous *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. It is a special case of the simplex algorithm, and was first discovered by Dantzig (1951).[46,47]

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

[47]When the CEU is symmetric, then 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 (LP) problem whose saddle points yield equilibria of the game. We

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 various economic literature. A well-known application is for optimization in the quasi-linear settings (Rochet, 1987), and in particular mechanism design and auctions (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 optimal transport setup, they show that any implementable policy must be cyclic monotone, and further *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 optimal transport 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.[48]

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 optimal transport problem can be equivalently viewed as an assignment problem (Shapley and Shubik, 1971). Then, algorithms from the matching literature can directly fit in, including the famous auction mechanisms by Crawford and Knoer (1981) and Demange et al. (1986).[49] Note that for a correlation sensitive CEU representation, binary choice reversal is a generic property for (p, q). Hence, we can, without loss,

can actually do better by invoking the following well known result on fictitious play: if both players use a no-regret learning algorithm to adapt their strategies to their opponent's strategies, then the average payoff of the players converge to their minmax value, and their average strategies constitute an approximate min-max 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; more precisely, it is at the order of the average regret in T rounds and recently improved to $O(lnT/T)$ for zero-sum games by Daskalakis et al. (2015).

[48]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 are examined and shown to be related to the positive quadrant order (PQD, Lehmann (1966)).

[49]Bertsekas and Castanon (1989) convert the transport problem to an assignment problem. A rich strand of literature extend and improve on the early methods in the general transport problem, see, for instance, Kim (2010), Chapter 1.5 for a short survey.

focus on lotteries with rational probabilities. For rational marginals, we can "replicate" the points in the support, and then reduce the rational lotteries to uniform densities and the feasible set of joint densities to doubly stochastic matrices.[50] We then refer to the famous Birkhoff-von Neumann theorem, which identifies all permutation matrices[51] as the extreme points of the set of doubly stochastic matrices (see, e.g., Budish et al. (2013)).

E.1 Test Correlations for a Fixed CEU Kernel

Suppose that a decision maker exhibits correlation sensitivity with respect to a pair of (p, q). We now 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 further $\pi_-^\intercal \in \Pi$ (which is true for a skew-symmetric ϕ), then she prefers p to q under π^+ but 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" correlation that determines whether the decision maker's preference over (p, q) is robust to correlation variations — if, for any pair (p, q), the decision maker 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 correlations in $\Gamma(p, q)$. It is the best-case correlation that determines whether the decision maker would, under any correlations, prefer p to q. These best-case and worst-case test correlations are of importance for testing whether the decision maker's preference is sensitive to correlation.[52] Observe that the problem is linear. Now, the problem boils down to finding the extreme points of $\Gamma(p, q)$ and then solving the OT problem, which can be done through the famous network simplex algorithm.

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

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

[51]A permutation matrix has, in each row and column, one entry equal to 1 and other entries equal to 0.

[52]In fact, the two extreme 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 $p \succeq^{*\Pi} q \iff \exists \pi \in \Pi$ 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 $p \succeq^{*\Pi} q$ and $q \succeq^{*\Pi} p$ hold.

1. *Given the tree π from previous iteration, compute the dual (u, v) according to complementary slackness equation (23).*
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 (uncorrelated) 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 the 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. The experiment results confirms this: more than half of (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 $\hat{\pi}^0$ with 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 11: π^0, L_1^0 vs L_2^0, maximally correlated distribution

E.2 Extremal Correlations

We fix a pair of (p, q). As the problem is linear, the algorithm always finds one extreme point of $\Gamma(p, q)$. Denote $\mathcal{E}(p, q)$ 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 14 (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 $X \times X$; an edge is linked between two nodes if (i) the two nodes are in the same row or column of the grid and (ii) they are adjacent in the concerning row or column (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 correlations are vital in our analysis.

Lemma 3 (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 of adopting extreme correlations observed in research studying a decision maker's preference robustness to correlation 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 decision makers' exhibited inconsistency in Allais paradox are partly related to different correlations. Most correlations in their the 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

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

$|V_\pi| = 2n - 1$. A correlation supported on a forest is similar to a perfect correlation given the non-existence of cycles, while a perfect correlation in the current terms is supported on a forest with exactly n disconnected nodes. For sake of testing a decision maker's robustness to correlation, it is sufficient to focus on extreme points supported on trees, the reason being that a generic pair (p, q) always admits at least one such extreme point.[54]

Lemma 4 (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 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.

Budish, E., Y.-K. Che, F. Kojima, and P. Milgrom (2013). Designing random allocation mechanisms: Theory and applications. *American Economic Review 103*(2), 585–623.

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

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

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.

[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 neighbourhood by perturbation.

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.

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.