A Theory of Choice Bracketing under Risk*

Mu Zhang †

November 20, 2021

Job Market Paper

[Click here for the latest version]

Abstract

In this paper, we study two heuristics, narrow bracketing and correlation neglect, that decision makers adopt to simplify the evaluation of risk from multiple sources and axiomatize them as behavioral deviations from the expected utility benchmark by relaxing the independence axiom. Our axiomatization reveals decision rules that (i) allow for narrow bracketing or correlation neglect or both and (ii) allow for different levels of narrow bracketing. We interpret the different sources as different streams of income to explain experimental evidence on violations of first order stochastic dominance. We propose a rationality ranking of decision makers based on their tendency to violate first order stochastic dominance and rank our models accordingly. Then, we interpret one source as background risk and show that narrow bracketing can explain risk aversion over small gambles. Finally, we interpret the different sources as consumption in different periods to explore the implications of narrow bracketing and correlation neglect on the optimal saving behavior. We also show that an Epstein-Zin type utility function can emerge because of narrow bracketing. This new utility function achieves separation between time and risk preferences without inducing an implausibly high early resolution premium as observed by Epstein, Farhi, and Strzalecki (2014).

*I am deeply indebted to my advisor, Faruk Gul, for his continued encouragement and support throughout this project, and to my committee members, Wolfgang Pesendorfer, Pietro Ortoleva and Xiaosheng Mu for their invaluable advice and guidance. I also thank Roland Bénabou, Modibo Camara, Sylvain Chassang, Xiaoyu Cheng, Joyee Deb, Francesco Fabbri, Shaowei Ke, Shengwu Li, Annie Liang, Alessandro Lizzeri, Dan McGee, Lasse Mononen, Evgenii Safonov, Ludvig Sinander, Rui Tang, Dmitry Taubinsky, João Thereze, Can Urgan, Leeat Yariv, Chen Zhao and seminar participants at EC'21, D-TEA 2021, RUD 2021, NASMES 2021, ESEM 2021, CSMES 2021, YES 2021 and Princeton Theory Lunch for helpful comments and discussions.

†Department of Economics, Princeton University, muz@princeton.edu

1 Introduction

Decision makers often have to deal with multiple interacting choice problems simultaneously. For example, an investor might need to manage accounts in different financial markets; a worker with both labor income and financial assets has to consider the risk of getting laid off and the risk of a financial crisis, as well as the interaction between them; a consumer might face income risk and interest rate risk in multiple periods; a subject with social preferences might care about the risk in both her own payoff and her opponent's payoff in an experiment. However, aggregating multi-source risk is challenging in practice and decision makers often adopt heuristics to simplify the evaluation process. In this paper, we focus on two such heuristics, narrow bracketing and correlation neglect.

Narrow bracketing, formalized by Thaler (1985) and Read, Loewenstein, and Rabin (1999), describes a situation where a decision maker (DM) faced with multiple choice problems attempts to solve each in isolation, without taking account of the interactions between the problems. When an agent confronts two choice problems over risky prospects simultaneously, the physical outcome of one choice will affect her ranking of the alternatives in the other problem. A narrow bracketing agent ignores this effect and may end up choosing a stochastically dominated prospect over one that dominates it. While all narrow bracketing agents will choose stochastically dominated prospects in some instances, some forms of narrow bracketing render decision makers more susceptible to violations of stochastic dominance than others. In particular, with the narrow bracketing model that is often used to interpret experimental evidence, it is possible for a decision maker to choose X dollars for sure over $Y > X$ dollars for sure when these amounts are suitably divided across the two income sources. Building such extreme departures from rationality into agents' behavior might lead to a theory that explains certain anomalies in data at the expense of creating others that are not likely to be present.

Each of the new models of narrow bracketing that we present below allows for a specification that precludes violations of stochastic dominance when there is no uncertainty. Moreover, we provide a ranking of these models based on how likely each one is to generate violations of stochastic dominance; that is, we provide a rationality ranking for narrow bracketing models. We hope that this ranking will be useful for calibrating the level of narrow bracketing and irrationality to specific applications.

The second heuristic, correlation neglect, describes the tendency of agents to ignore the statistical interdependence among different decision problems. When an agent faces two-dimensional risk, the outcome on one dimension will alter the payoff distribution associated with the other dimension. A correlation neglecting agent may underestimate or overestimate the hedging effect that one dimension provides over the other. Such an agent may appear overly risk averse or unwilling to take advantage of diversification opportunities. Our analysis reveals that the most extreme form of narrow bracketing, which we call full narrow bracketing, implies correlation neglect. However, in general, it is possible to have a weaker form of narrow bracketing, which we call asymmetric narrow bracketing, with or without correlation neglect. This flexibility is useful because there is plenty of evidence that not all individuals ignore correlation and, in fact, many are averse to it. When evaluating income from two separate sources, this aversion to correlation is an immediate consequence of risk aversion.

We characterize the two above heuristics as behavioral deviations from the expected utility benchmark by relaxing the independence axiom of that model. Our axiomatization reveals decision rules that (i) allow for narrow bracketing or correlation neglect or both and (ii) allow for different degrees of deviations from the benchmark without narrow bracketing. For instance, we allow the DM to violate stochastic dominance in complicated problems but not in simple problems that only involve deterministic options. We show that our asymmetric narrow bracketing model without correlation neglect facilitates a dynamic model that distinguishes between time preference and risk preference without imposing a preference for timing of resolution of uncertainty.

We consider preferences over lotteries with two-dimensional outcome profiles. An outcome profile can be interpreted as the consequences of two decision problems, such as choices over two different classes of risky assets or consumption choices involving two periods. We will call each dimension a *source* of risk and the marginal distribution in some source a *marginal lottery*. Our main formal results are two representation theorems. The first builds in correlation neglect and encompasses three cases, (pure) correlation neglect (CN), asymmetric narrow bracketing with correlation neglect (ANB-CN) and full narrow bracketing (FNB). Pure correlation neglect is described by the following type of utility function:

$$V^{CN}(P) = \sum_{x,y} w(x,y) P_1(x) P_2(y)$$

for all $P \in \mathcal{P}$ where P_i is the marginal lottery of P in source i. Hence, the decision maker identifies every P with its marginal distributions and ignores the correlation.

The second case covered by Theorem 1 is asymmetric narrow bracketing with correlation neglect. Let v be a utility index over one-dimensional gambles and let $CE_v(P_2)$ be the certainty equivalent of the marginal distribution P_2 of P; that is, $CE_v(P_2)$ is the unique x that solves $v(x) = \sum_z v(z) P_2(z)$. Then, the asymmetric narrow bracketing with correlation neglect utility function is:

$$V^{ANB-CN}(P) = \sum_{x} w(x, CE_v(P_2)) P_1(x)$$

for all $P \in \mathcal{P}$. In this case, the decision maker not only ignores the statistical dependence between the two sources of risk but also takes limited account of the physical interaction of the outcomes from two sources. For example, if the two sources are different sources of income, then the utility function above means that the DM is only considering the income effect that arises from the second source “on average”.

The last case covered by Theorem 1 is full narrow bracketing:

$$V^{FNB}(P) = w(CE_u(P_1), CE_v(P_2))$$

for all $P \in \mathcal{P}$. In this case, the DM evaluates both sources of risk in isolation, and therefore, correlation neglect is implied.

Theorem 1 characterizes the above three types of utility functions with correlation neglect. They follow from two standard assumptions, weak order and monotonicity, weaker versions of standard continuity, independence axioms and the axiom of correlation neutrality.

For our second theorem, we relax the correlation neutrality axiom to a correlation consistency axiom which permits correlation neglect but does not require it. The characterization of Theorem 2 permits the three representations of Theorem 1 described above and two other classes of utility functions: the standard, “fully rational”, expected utility function

$$V^{EU}(P) = \sum_{x,y} w(x,y)P(x,y)$$

for all $P \in \mathcal{P}$, and the asymmetric narrow bracketing utility function without correlation neglect defined as follows: for any x such that $P(x,y) > 0$ for some y, let $P_{2|x}$ denote the conditional distribution of the outcome in source 2 given x in source 1. Then, the asymmetric narrow bracketing utility function is

$$V^{ANB}(P) = \sum_{x} w(x, CE_v(P_{2|x}))P_1(x)$$

for all $P \in \mathcal{P}$. Compared to the asymmetric narrow bracketing with correlation neglect utility function, here the decision maker takes into account the statistical dependence between the two sources of risk. Suppose that the two sources represent consumption levels in two periods. Given each realization of today’s consumption, the decision maker evaluates the conditional risk in tomorrow’s consumption narrowly using its certainty equivalent. Then she averages over the utility of today’s consumption and tomorrow’s certainty equivalent. Asymmetric narrow bracketing suggests a heuristic that mimics backward induction in a manner that does not take into account the effect of today’s consumption on tomorrow’s preferences. This asymmetry reflects the nature of intertemporal problems that future risk is typically more complicated and less accessible than current risk.

Unlike a standard representation theorem that derives a general functional form to include all models, each of our theorems characterizes several seemingly distinct functional forms. In each utility function, the DM either adopts some heuristic, or does not use it at all. This feature follows from our axioms, and is not an ex ante assumption. For instance, intermediate cases such as the weighted average of two utility functions V^{EU} and V^{FNB} are ruled out by our axioms. We argue that such exclusion is consistent with our interpretation of narrow bracketing and correlation neglect as simplifying heuristics, since evaluating a lottery using the weighted average of V^{EU} and V^{FNB} is more complex and demanding than using either V^{EU} or V^{FNB}. Ellis and Freeman (2020) also show that allowing for such intermediate levels of narrow bracketing does not significantly help to explain the data in their experiments.

In Section 6, we consider applications of our models. First, we interpret the different sources as different streams of income to explain experimental evidence on violations of first order stochastic dominance. We focus on models that satisfy stochastic dominance when there is no uncertainty. We also propose a rationality ranking of these models based on the readiness with which they yield violations of stochastic dominance. This reveals a novel connection between violations of stochastic dominance and deviations from the independence axiom of the expected utility model. Then, we interpret one source as background risk and show that narrow bracketing can explain risk aversion over small gambles.

Finally, we interpret the different sources as consumption in different periods to explore the implications of narrow bracketing and correlation neglect on intertemporal risk attitudes. We show that the full narrow bracketing model is essentially identical to the time preference model studied by Selden (1978) and Selden and Stux (1978). We also show that an Epstein and Zin (1989) type utility function can emerge because of narrow bracketing. This new utility function achieves separation between time and risk preferences without inducing an unreasonably high timing premium as observed by Epstein, Farhi, and Strzalecki (2014). Then, we study how narrow bracketing and correlation neglect affect the optimal saving behavior of a consumer in a simple two-period model. To isolate the effect of narrow bracketing, we consider the standard "certain $\times$ uncertain" setup in the literature of precautionary savings, where there is only one saving decision (in the first period) and income is uncertain in a single period (in the second period). We show that when the consumer's relative risk aversion is larger than the reciprocal of her elasticity of intertemporal substitution, narrow bracketing leads to more risk aversion over future consumption and hence more current saving compared to the expected utility benchmark. To study the effect of correlation neglect, we modify the previous setup by assuming that the consumer knows her future income in the first period and bases her saving decision on this information. Compared to a consumer with correlation neglect, a correlation averse consumer adjusts the current saving levels to compensate for the difference in future income. She saves more when the future income is low and less when the future income is high.

The above applications reveal a novel and possibly surprising connection between the Epstein-Zin type utility function, the most widely used model that separates time and risk preferences in the literature, and violations of first order stochastic dominance, a phenomenon typically considered as an irrational bias. Both the separation of time and risk preferences and violations of stochastic dominance can result from narrow bracketing and be captured by relaxing the independence axiom.

Related Literature

Tversky and Kahneman (1981), Rabin and Weizsäcker (2009) and Ellis and Freeman (2020) provide experimental evidence for narrow bracketing. The experiments in the first two papers isolate narrow bracketing by considering lotteries with independent marginals. Hence, there is no correlation for the DM to ignore. In Ellis and Freeman (2020), each of two assets pays off in only one of two equally likely states, which induces perfect negative correlation between payoffs of the two assets. The authors show that 73% of their subjects are best described by full narrow bracketing, 14% by broad bracketing (expected utility), and only 5% by intermediate cases. However, they do not consider models that separate correlation neglect and narrow bracketing. There is also plenty of experimental evidence for correlation neglect in belief formation (Enke and Zimmermann, 2019), portfolio allocation (Eyster and Weizsäcker, 2016, Kallir and Sonsino, 2009) and school choice (Rees-Jones, Shorrer, and Tergiman, 2020). Those experiments do not address narrow bracketing.

Our paper is related to the growing literature on rationales for narrow bracketing. For instance, Lian (2020) develops an informational theory of narrow bracketing, where the DM cares less about her other decisions when making each decision because of imperfect knowledge of other decisions. Camara (2021) studies the expected utility model with high-dimensional decisions and shows that computational tractability requires the utility index to satisfy a slightly weaker version of additive separability. Complementary to those papers which provide rationales for narrow bracketing, our paper characterizes the observable implications of narrow bracketing and axiomatizes it as a deviation from expected utility. Moreover, our model can be applied to simple economic settings such as the experiment in Rabin and Weizsäcker (2009), without informational frictions or high-dimensional decisions.

In Vorjohann (2021), each DM is characterized by a broad preference and a narrow preference, both of which can be observed and admit an expected utility representation. Vorjohann (2021) axiomatizes the expected utility index of the narrow preference, which is additive separable across brackets and can be derived from the utility index of the broad preference. Our work differs from Vorjohann (2021)'s in two aspects. First, the DM in our model has only one preference, which determines whether she is an expected utility maximizer, or adopts some simplifying heuristics. Second, we allow for more flexible forms of narrow bracketing and distinguish it from correlation neglect, while the notion of narrow bracketing in Vorjohann (2021) subsumes correlation neglect.

A strand of recent literature studies the axiomatization of correlation-sensitive models. Ellis and Piccione (2017) models a decision maker who chooses a portfolio of assets among which she may misperceive the joint returns. She has a (possibly misspecified) subjective belief over the correlation of returns and maximizes expected utility. Ellis (2021) further extends the framework and considers a decision maker who deems possible a set of possible correlation structures and evaluates each alternative by the worst expected utility in that set. In our paper, we only consider the extreme form of correlation misperception, i.e., correlation neglect, since our goal is to separate it from narrow bracketing. We leave the analysis of general correlation misperception for future study. Lanzani (2021) studies a decision maker who is sensitive to the correlation of two risky prospects she needs to compare. By relaxing the transitivity axiom of the expected utility theory, he provides an axiomatization for the salience theory model in Bordalo, Gennaioli, and Shleifer (2012). Our treatment of correlation is complementary to Lanzani (2021). One can interpret that the decision maker in Lanzani (2021) is comparing the *two marginal lotteries* of the same lottery over two-source outcome profiles, while the decision maker in our model is choosing between *two lotteries*.

Our application to portfolio choices is related to the recent literature on risk aversion over small gambles. Many remedies have been proposed since the critique of Rabin (2000) on the expected utility theory. However, the results in Barberis, Huang, and Thaler (2006) and Mu, Pomatto, Strack, and Tamuz (2021) suggest that narrow bracketing is necessary to explain commonly observed irregularity in choice under risk. These results motivate our work.

Our paper is also related to the recent work on Epstein-Zin preferences. Epstein, Farhi, and Strzalecki (2014) show that calibrating the Epstein-Zin preferences to macroeconomic and finance data yields an implausibly high early resolution premium. We propose an Epstein-Zin type utility function that emerges from narrow bracketing and addresses this critique.

2 Primitives

Let $Z = X_1 \times X_2$, where X_i is the set of outcomes in source $i \in \{1, 2\}$. We assume that X_i is a nontrivial compact (closed and bounded) interval on the real line. We call $(x_1, x_2) \in Z$ an *outcome profile* and x_i is the outcome in source i for $i \in \{1, 2\}$.

A *(joint) lottery*, P, is a probability distribution over Z with a finite support. We endow the set of all lotteries, $\mathcal{P} = \mathcal{L}^0(Z)$, with the topology of weak convergence and the standard mixture operation. For $P \in \mathcal{P}$, the *marginal lottery* of P in source 1 is denoted $P_1 \in \mathcal{L}^0(X_1)$.

That is, $P_1(x) = \sum_{y \in X_2} P(x,y)$ for all $x \in X_1$. The marginal lottery P_2 in source 2 is defined similarly. Let $\hat{\mathcal{P}} = \mathcal{L}^0(X_1) \times \mathcal{L}^0(X_2) \subset \mathcal{P}$ denote the set of *product lotteries*. For $P \in \mathcal{P}$, the pair $(P_1, P_2) \in \mathcal{P}$ is a product lottery with the same marginals as P. When there is no risk of confusion, we use (x_1, x_2) to denote the degenerate lottery that yields the outcome profile $(x_1, x_2) \in Z$ for sure and interpret Z as a subset of $\mathcal{P}$.

The primitive of our analysis is a binary relation $\succsim$ on $\mathcal{P}$.

3 Representations

In this section, we introduce different decision rules adopted by a decision maker faced with two-source risk. We start with the expected utility model as the benchmark. A binary relation $\succsim$ is said to admit an *expected utility* (EU) representation if it can be represented by V^{EU} with

$$V^{EU}(P) = \sum_{x,y} w(x,y)P(x,y)$$

for all $P \in \mathcal{P}$, where w is a continuous and strictly monotonic utility index on Z. In Section 6, we will study specifications of w when outcomes represent income from two sources or consumption in two periods.

As discussed in the introduction, people in practice usually deviate from the benchmark systematically by adopting some simplifying heuristics. In the following we will focus on two such heuristics: choice bracketing and correlation neglect. We say that a binary relation $\succsim$ admits a *full narrow bracketing* (FNB) representation if it is represented by V^{FNB}:

$$V^{FNB}(P) = w(CE_u(P_1), CE_v(P_2))$$

for all $P \in \mathcal{P}$, where w is a continuous and strictly monotonic utility index on Z and u, v are continuous and strictly monotonic utility indices on X_1, X_2 respectively. $CE_u(P_1)$ is the certainty equivalent of the marginal distribution P_1 of P in source 1 under utility index u; that is, $CE_u(P_1)$ is the unique x that solves $u(x) = \sum_z u(z)P_1(z)$. $CE_v(P_2)$ is defined similarly. Hence, a full narrow bracketing decision maker evaluates each marginal lottery, P_i, independently and aggregates the two evaluations to determine the overall utility of lottery P. It follows that she also ignores the correlation between the marginals.

The full narrow bracketing model might be too restrictive as it necessitates correlation neglect. Recent experimental literature has shown the prevalence of correlation non-neutrality in different economic scenarios. For instance, Andersen, Harrison, Lau, and Rutström (2018) find strong evidence for correlation aversion over lotteries with intertemporal income profiles, that is, subjects are averse to positive correlation between income in different periods. Lanier, Miao, Quah, and Zhong (2020) confirm this finding using non-parametric analysis and a different experimental design. For atemporal settings, Ebert and van de Kuilen (2016) and Attema, L'haridon, and van de Kuilen (2019) find correlation non-neutrality for two-source risk in monetary outcomes and outcomes in terms of waiting time or longevity. Thus, it is worthwhile to develop a model that separates choice bracketing from correlation neglect.

This is our model called asymmetry narrow bracketing, where the decision maker only narrowly bracket risk in one source, instead of both sources. Without loss of generality, we denote by source 2 the source of narrow bracketing. This labeling makes sense in practice when the risk in source 2 is more complicated and/or less accessible than the risk in source 1. For instance, we interpret the outcome in source 1 as today's consumption and the outcome in source 2 as tomorrow's consumption, because evaluating tomorrow's risk requires the decision maker to assess the hypothetical circumstances she is not facing at present. Future risk is also more complex since the distribution over consumption paths is usually presented in the form of a probability tree starting with the current consumption. Similarly, in social preferences, we interpret the outcomes in source 1 and source 2 as the payoffs of the decision maker and her opponent respectively, because it is more complicated for the decision maker to put herself in the opponent's shoes. In the case with background risk, it is less clear whether or not the decision maker is more inclined to narrowly bracket the background risk. One might argue that outcomes generated by the prospect the decision maker is facing now are more accessible than background wealth and hence source 2 should represent the background risk. However, since the background risk is fixed and the decision maker is choosing over different risky prospects at hand, she might evaluate a risky prospect conditional on the realization of background wealth. In this case, source 1 should represent the background risk.

For each $P \in \mathcal{P}$ and $x \in X_1$ such that $P(x, y) > 0$ for some $y \in X_2$, let $P_{2|x}$ denote the conditional distribution of the outcome in source 2 given x in source 1. A binary relation $\succsim$ admits an *asymmetric narrow bracketing* (ANB) representation if it is represented by V^{ANB}:

$$V^{ANB}(P) = \sum_x w(x, CE_v(P_{2|x}))P_1(x)$$

for all $P \in \mathcal{P}$, where w, v are continuous and strictly monotonic utility indices on Z, X_2 respectively. To assesses the overall utility of lottery P, a decision maker with asymmetric narrow bracketing first replaces each conditional lottery in source 2 with its certainty equivalent and then evaluates the new lottery using expected utility. The decision maker narrowly brackets the risk in source 2 but not in source 1. Moreover, she takes into account the correlation between outcomes in two sources. Since the notion of correlation aversion typically involves lotteries with degenerate conditional lotteries in source 2 (e.g., Bommier (2007)), V^{ANB} and V^{EU} share the same attitude towards correlation if they have the same utility index w over deterministic outcome profiles.

When each outcome profile is a consumption path, the asymmetric narrow bracketing model combines narrow bracketing with backward induction and resembles the functional form of recursive preferences such as Epstein and Zin (1989). In Section 6.2, we show that an Epstein-Zin type utility function can arise due to narrow bracketing instead of a preference for timing of resolution of uncertainty.

Now we consider a decision maker with correlation neglect and no narrow bracketing. A binary relation $\succsim$ admits an *(pure) correlation neglect* (CN) representation if it is represented by V^{CN}:

$$V^{CN}(P) = \sum_{x,y} w(x,y)P_1(x)P_2(y)$$

for all $P \in \mathcal{P}$, where w is a continuous and strictly monotonic utility index on Z. A decision maker with correlation neglect rationally aggregates independent risk from two sources, but fails to appreciate the correlation.

Our final utility function combines asymmetric narrow bracketing with correlation neglect (ANB-CN), and is described below:

$$V^{ANB-CN}(P) = \sum_{x} w(x, CE_v(P_2))P_1(x)$$

for all $P \in \mathcal{P}$, where w, v are continuous and strictly monotonic utility indices on Z, X_2 respectively. Compared to V^{ANB}, here the decision maker ignores the statistical dependence between the two sources of risk: she first determines the certainty equivalent of the marginal lottery of P in source 2 and then assesses the overall utility of P as if it yields that certainty equivalent in source 2 for sure.

Figure 1 depicts the heuristics involved in the five models introduced above. Given the expected utility benchmark, V^{CN} exhibits correlation neglect and no narrow bracketing while V^{ANB} exhibits narrow bracketing in source 2 and no correlation neglect. V^{ANB-CN} combines the previous two heuristics, and V^{FNB} further involves narrow bracketing in source 1.

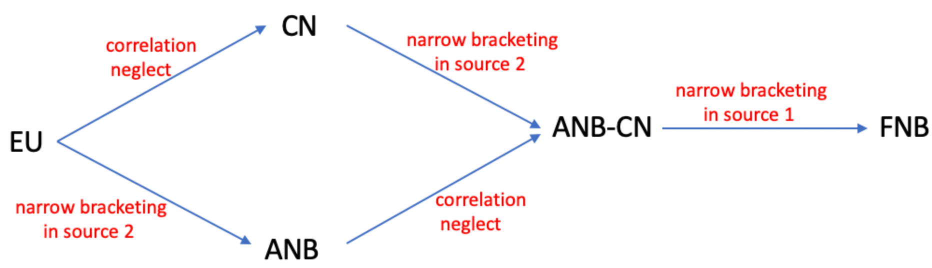

Figure 1: Relationship between models in terms of heuristics.

4 Axioms

We will present two representation theorems. Our Theorem 1 identifies axioms that ensure $\succsim$ can be represented by one of these three utility functions with correlation neglect: V^{CN}, V^{ANB-CN} or V^{FNB}. Moreover, any preference that can be represented by one of these utility functions must satisfy the axioms in Theorem 1. In Theorem 2, we weaken correlation neglect to incorporate the other two utility functions: the expected utility benchmark V^{EU} and asymmetric narrow bracketing V^{ANB}.

4.1 Narrow Bracketing with Correlation Neglect

In this section, we present our axioms for representations that exhibit correlation neglect. The first of the four common axioms shared by our two main theorems is rationality.

Weak Order: $\succsim$ is complete and transitive.

The next axiom is a version of monotonicity for our two-dimensional setting. With single-source risk, there is an agreed notion first order stochastic dominance. However, it is possible to extend this notion to a setting with multiple sources in several ways. We will choose a weak notion of dominance, which only involves comparing each lottery P with degenerate lotteries $(x, y) \in Z$. For any marginal lottery $p \in \mathcal{L}^0(X_1) \cup \mathcal{L}^0(X_2)$, let $\text{supp}(p) := \{x \in X_1 \cup X_2 : p(x) > 0\}$ be the support of p. For any $P, (x, y) \in \mathcal{P}$, we say P *dominates* (x, y) if $P \neq (x, y)$ and $x' \geq x, y' \geq y$ for all $x' \in \text{supp}(P_1), y' \in \text{supp}(P_2)$. Symmetrically, we say (x, y) *dominates* P if $P \neq (x, y)$ and $x' \leq x, y' \leq y$ for all $x' \in \text{supp}(P_1), y' \in \text{supp}(P_2)$. Axiom 2 states that the preference $\succsim$ is monotonic with respect to this notion of dominance.

Monotonicity: For $P \in \mathcal{P}$ and $(x, y) \in Z$, $P \succ (x, y)$ if P dominates (x, y) and $(x, y) \succ P$ if (x, y) dominates P.

Our continuity axiom combines three weaker notions of the standard continuity axiom: Continuity 1 is mixture space continuity (i.e., continuity in probabilities); Continuity 2 is continuity in outcomes while Continuity 3 combines the two but is restricted to product lotteries.

Continuity 1: For $P, R, Q \in \mathcal{P}$, the sets $\{\alpha \in [0, 1] : \alpha P + (1 - \alpha)Q \succ R\}$ and $\{\alpha \in [0, 1] : R \succ \alpha P + (1 - \alpha)Q\}$ are open subsets of $[0, 1]$.

Continuity 2, below, asserts that if increasing (decreasing) every outcome in the support of P by $(\varepsilon^1, \varepsilon^2)$ renders P better (worse) than Q, for any positive $\varepsilon^1, \varepsilon^2$, then P must also be better than Q. Since the set Z is bounded, we need to deal with the possibility that increasing some of the prizes in the support of P may not be feasible. Formally, for any $P \in \mathcal{P}$, choose $\eta > 0$ small enough so that $P(x,y) \cdot P(x',y') > 0$, $|x - x'| \leq \eta$ and $|y - y'| \leq \eta$ implies $(x,y) = (x',y')$. Then, for $\varepsilon = (\varepsilon^1, \varepsilon^2)$ such that $\varepsilon^1, \varepsilon^2 \in (-\eta, \eta)$, define $\phi : Z \to Z$ as follows: $\phi_\varepsilon(x,y) = (x + \varepsilon^1, y + \varepsilon^2)$ if $(x + \varepsilon^1, y + \varepsilon^2) \in Z$; otherwise, $\phi_\varepsilon(x,y) = (x',y')$ such that (x',y') is the element of Z closest to (x,y) with respect to the distance $d((x,y),(x',y')) = |x - x'| + |y - y'|$. Since we have chosen $\varepsilon^1, \varepsilon^2$ sufficiently small, the restriction of ϕ to the support of P is one-to-one. Then, define $P_\varepsilon \in \mathcal{P}$ as follows

$$P_\varepsilon(\phi_\varepsilon(x,y)) = P(x,y)$$

if $P(x,y) > 0$ and set $P_\varepsilon(x',y') = 0$ if $(x',y') \neq \phi_\varepsilon(x,y)$ for any (x,y) in the support of P.

Continuity 2: For all $P, Q \in \mathcal{P}$ and ε_n converging to $(0,0)$,

$$P_{\varepsilon_n} \succsim Q \text{ for all } n \text{ implies } P \succsim Q \text{ and } Q \succsim P_{\varepsilon_n}, \text{ for all } n \text{ implies } Q \succsim P.$$

For product lotteries, we will assume strong continuity:

Continuity 3: For all $Q \in \mathcal{P}$, the sets $\{P \in \hat{\mathcal{P}} : P \succ Q\}$ and $\{P \in \hat{\mathcal{P}} : Q \succ P\}$ are open subsets of $\hat{\mathcal{P}}$.

We will refer to the conjunction of the three notions above as **Continuity**. Note that Continuity is weaker than the strong continuity axiom of expected utility theory with monetary prizes, which is Continuity 3 by replacing $\hat{\mathcal{P}}$ with $\mathcal{P}$. Note that Continuity 1 and Continuity 2 would be implied by this strong continuity. However, Continuity does not imply strong continuity.

Below, we consider an alternative to the von Neumann-Morgenstern (vNM) independence axiom. Recall that expected utility theory preferences satisfy the following stronger version of the independence axiom: $P \succ Q$, $R \sim S$ implies $\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$ for $\alpha \in (0,1)$. In our alternative imposed on product lotteries $\hat{\mathcal{P}}$, we restrict the condition above to three situations in which P-R, Q-S or R-S are sufficiently similar. In order to guarantee that the mixture of two product lotteries remain a product lottery, we will require that P and R are identical in some source and so are Q and S.

We define the *conditional preference in source 1* given marginal lottery $q \in \mathcal{L}^0(X_2)$ in source 2, $\succsim_{1|q}$, as the restriction of $\succsim$ on $\mathcal{L}^0(X_1) \times \{q\}$, that is, $p \succsim_{1|q} p'$ if and only if $(p,q) \succsim (p',q)$ for each $p, p' \in \mathcal{L}^0(X_1)$. Symmetrically, for each $p \in \mathcal{L}^0(X_1)$, $\succsim_{2|p}$ denotes the conditional preference in source 2 given p in source 1. For $i \in \{1,2\}$, we denote $-i \neq i$ and $-i \in \{1,2\}$. Then we say P and Q are *conditionally indifferent in source* i if there exists $z \in X_{-i}$ such that $P_i \sim_{i|z} Q_i$. Conditional indifference captures the notion of sufficient similarity of marginal lotteries in one source.

The first notion of sufficient similarity requires R to be the same as S and P, Q and

R to have the same marginal lottery in one source. The second notion requires P to be conditionally indifferent to R in one source and identical to R in the other; similarly, Q and S need to be conditionally indifferent in one source and identical in the other. The third notion requires that R is the same as S, identical to P in one source and identical to Q in the other, and that P and Q are conditionally indifferent in source 2.

Independence*: For $P, Q, R, S \in \hat{\mathcal{P}}$ and $\alpha \in (0,1)$,

$$P \succ Q \Longrightarrow \alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S \quad \text{if}$$

(i) $R = S$ and for some $i \in \{1,2\}$, $P_i = Q_i = R_i$ or

(ii) $R \sim S$ and for some $i, j \in \{1,2\}$, $x \in X_i$ and $y \in X_j$, $P_i = R_i$, $Q_j = S_j$, $P_{-i} \sim_{-i|x} R_{-i}$ and $Q_{-j} \sim_{-j|y} S_{-j}$ or

(iii) $R = S = (P_1, Q_2)$ or (Q_1, P_2) and for some $x \in X_1$, $P_2 \sim_{2|x} Q_2$.

Now we briefly discuss the three parts in Independence*. Part (i) guarantees that given a fixed marginal lottery in one source, the conditional preference in the other must admit an expected utility representation. Hence, we maintain the independence axiom within each source and only relax inter-source independence. Part (ii) becomes a trivial implication of part (i) if P is indifferent to R or Q is indifferent to S. Suppose that these two conditions do not hold, then the requirement that P and R are conditionally indifferent in source $-i$ implies that how the DM ranks P_{-i} and R_{-i} depends on the marginal risk in the other source since $P_{-i} \not\sim_{-i|P_i} R_{-i}$ and $P_{-i} \sim_{-i|x} R_{-i}$ for some $x \in X_2$. This implies that the DM does not *narrowly bracket* P_{-i} and R_{-i} in source $-i$. Similarly, the DM does not narrowly bracket Q_{-j} and S_{-j} in source $-j$. Thus, part (ii) essentially states that the independence axiom holds when the DM does not narrowly bracket marginal lotteries in P-R and Q-S. Part (iii) strengthens part (ii) when R is the same as S, identical to P in one source and identical to Q in the other. In this case, the independence axiom holds if the DM does not narrowly bracket the P_2 and Q_2. Part (iii) is important to rule out a symmetric counterpart of V^{ANB-CN} in Theorem 1 where the roles of two sources are reversed.

Our final axiom asserts that the DM identifies every lottery with the product of its marginals; that is, she ignores correlation.

Correlation Neutrality: For all $P \in \mathcal{P}$, $P \sim (P_1, P_2)$.

Now we are ready to state our first representation theorem under correlation neglect.

Theorem 1. *A binary relation on $\mathcal{P}$ satisfies Weak Order, Monotonicity, Continuity, Independence* and Correlation Neutrality if and only if it can be represented by one of the following types of utility functions: V^{CN}, V^{ANB-CN} or V^{FNB}.*

4.2 Narrow Bracketing with Correlation Sensitivity

In this section, we discard the correlation neglect axiom to incorporate representations that are sensitive to the correlation structure of risk in different sources, that is, the expected utility benchmark V^{EU} and asymmetric narrow bracketing V^{ANB}.

The following axiom relaxes Correlation Neutrality.

Correlation Consistency: For $P, Q, R, S \in \mathcal{P}$ and $\alpha \in (0,1)$,

$$\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S \quad \text{if}$$

$P \succ Q$, $R \sim S$, for $i = 1, 2$, $P_i = Q_i$ and $supp(P_1) \cap supp(R_1) = supp(Q_1) \cap supp(S_1) = \emptyset$.

To understand Correlation Consistency, note that since P and Q have the same marginal lotteries, $P \succ Q$ means that the DM prefers the correlation structure of P to that of Q. Then, the axiom states that this preference is preserved when R is indifferent to S, and the mixture of P and R, Q and S does not affect the correlation structures of P and Q. Here, we interpret the correlation structure of P as the profile of conditional lotteries $\{P_{2|x}\}_{x \in \text{supp}(P_1)}$ in source 2, i.e., the source that the DM is more inclined to narrowly bracket. [1]

Our Theorem 2 states that, by relaxing Correlation Neutrality in Theorem 1 to Correlation Consistency, we allow for two more utility functions V^{EU} and V^{ANB}.

Theorem 2. *A binary relation on $\mathcal{P}$ satisfies Weak Order, Monotonicity, Continuity, Independence* and Correlation Consistency if and only if it can be represented by one of the following types of utility functions: V^{EU}, V^{ANB}, V^{CN}, V^{ANB-CN} or V^{FNB}.*

Theorem 1 and Theorem 2 characterize behavioral decision rules differing in the involved behavioral heuristics: correlation neglect, narrow bracketing in source 1 and narrow bracketing in source 2. In each representation, a heuristic is either adopted or completely ruled out. This feature results from our axioms, instead of an ex ante assumption, and is in contrast to a standard representation theorem in the literature, which axiomatizes a general functional form that incorporates all models.

4.3 Identification

The identification of utility functions characterized in Theorem 1 and Theorem 2 is clear: except for the utility index over deterministic outcome profiles w in V^{FNB}, all other utility indices are unique up to a positive affine transformation.

[1] If we strengthen Correlation Consistency by allowing the mixture to alter the correlation structure in P and Q, then we need to discard V^{ANB} in Theorem 2.

Proposition 1. *i). If $\succsim$ is represented by V^{EU} or V^{CN} with indices w and w', then there exist $a > 0$ and $b \in \mathbb{R}$ such that $w'(x, y) = aw(x, y) + b$ for all $(x, y) \in Z$.*

ii). If $\succsim$ is represented by V^{ANB} or V^{ANB-CN} with indices (w, v) and (w', v'), then there exist $a, a' > 0$ and $b, b' \in \mathbb{R}$ such that $w'(x, y) = aw(x, y) + b$ for all $(x, y) \in Z$ and $v'(y) = a'v(y) + b'$ for all $y \in X_2$.

iii). If $\succsim$ is represented by V^{FNB} with indices (w, u, v) and (w', u', v'), then there exist $a, a' > 0$ and $b, b' \in \mathbb{R}$ and $\phi : w(Z) \to \mathbb{R}$ continuous and strictly increasing such that $w'(x, y) = \phi(w(x, y))$ for all $(x, y) \in Z$, $u'(x) = au(x)+b$ for all $x \in X_1$ and $v'(y) = a'v(y)+b'$ for all $y \in X_2$.

5 Proof Sketch

In this section, we briefly sketch the proofs of Theorem 1 and Theorem 2. We start with Theorem 1, the representation theorem for models that exhibit correlation neglect. It will serve as an important intermediate result for the proof of Theorem 2.

By Correlation Neutrality, it suffices to focus on $\succsim$ restricted to product lotteries $\hat{\mathcal{P}}$. Recall that $\succsim_{i|q}$ is the conditional preference over marginal lotteries in source i given q in the other source. By Continuity and part (i) of Independence*, $\succsim_{i|q}$ admits an expected utility representation.

If $\succsim_{i|q}$ is independent of $q \in \mathcal{L}^0(X_{-i})$ for each $i = 1, 2$, then the decision maker adopts narrow bracketing in both sources and evaluates the marginal distributions in isolation. Suppose $\succsim_i := \succsim_{i|q}$ is represented by the expected utility index v_i. Continuity and Monotonicity guarantee that v_i is continuous and strictly increasing and the certainty equivalent function CE_{v_i} is well defined. Hence, the decision maker is indifferent between $P \in \mathcal{P}$ and $(CE_{v_1}(P_1), CE_{v_2}(P_2))$, and her preference is represented by some V^{FNB}.

If $\succsim_{2|q}$ is independent of $q \in \mathcal{L}^0(X_1)$, and $\succsim_{1|p} \neq \succsim_{1|p'}$ for some $p, p' \in \mathcal{L}^0(X_2)$, then the decision maker adopts narrow bracketing in source 2 but not in source 1. We can focus on the restriction of $\succsim$ to $\mathcal{L}^0(X_1) \times X_2$ and our goal is to show the the following independence axiom:

$$P \succsim Q \iff \alpha P + (1 - \alpha)R \succsim \alpha Q + (1 - \alpha)S, \ \forall \ \alpha \in (0, 1),$$

if $R \sim S$ and $P_2 = R_2 = y_1, Q_2 = S_2 = y_2$ with $y_1, y_2 \in X_2$. This property guarantees that the decision maker admits a utility function which is linear in the marginal distributions of the outcome in source 1. We can show that it is exactly the V^{ANB-CN} representation.

Part (ii) of Independence* implies a local version of the above property; that is, it holds under an additional condition that P and R, Q and S are conditionally indifferent in source 1.

This requires the existence of $y_1', y_2' \in X_2$ with $P_1 \sim_{1|y_1'} R_1$ and $Q_1 \sim_{1|y_2'} S_1$ when $P \not\sim R$ and $Q \not\sim S$. Then the technical part of the proof is to bridge the gap between the local property and the desired global one. Intuitively, if the decision maker never exhibits narrow bracketing in source 1 locally, then she does not exhibit narrow bracketing in source 1 globally.

If $\succsim_{2|q} \neq \succsim_{2|q'}$ for some $q, q' \in \mathcal{L}^0(X_1)$, then the decision maker does not adopt narrow bracketing in source 2. Combining the above arguments with part (iii) of Independence*, which states that the decision maker is more inclined to narrowly bracket risk in source 2, we can show that she does not adopt narrow bracketing in both sources and her preference can be represented by V^{CN}. This completes the proof for Theorem 1.

The proof of Theorem 2 can be decomposed into four steps. First, we define $\hat{\succsim}$ on $\mathcal{P}$ which satisfies Correlation Neutrality and agrees with $\succsim$ on $\hat{\mathcal{P}}$. Easy to very that $\hat{\succsim}$ satisfies the axioms in Theorem 1 and hence can be represented by either V^{CN} or V^{ANB-CN} or V^{FNB}. If Correlation Neutrality holds, then $\hat{\succsim} = \succsim$ and we are done. From now on, suppose that Correlation Neutrality fails. We say two lotteries P and Q are compatible if their marginal distributions in source 1 have disjoint supports, that is, $P_1(x) \cdot Q_1(x) = 0$ for all $x \in X_1$. Compatibility guarantees that the mixture between P and Q does not alter their correlation structure in terms of conditional distributions of outcomes in source 2. The second step is to show that the standard independence axiom holds subject to this compatibility constraint. Third, we extend $\succsim$ to a richer mixture space that mimics the space of temporal lotteries in Kreps and Porteus (1978). We show that the extended preference satisfies the standard mixture space axioms and admits a linear representation. Finally, exploiting the consistency of the two representations of $\succsim$ restricted to $\hat{\mathcal{P}}$ in the previous steps, we conclude that $\succsim$ can be represented by either V^{EU} or V^{ANB} when Correlation Neutrality fails.

6 Applications

6.1 Portfolio Choices

In this section, we consider a decision maker (DM) faced with uncertainty in multiple income sources. We show that narrow bracketing can explain violations of first order stochastic dominance observed in experiments, and risk aversion over small gambles with background risk. We also develop a rationality ranking of utility functions axiomatized in Theorem 2 based on the readiness with which they yield violations of stochastic dominance. Consistent with the intuition, the more behavior heuristics a utility function encompasses, the more likely it is to yield violations of stochastic dominance and hence the more irrational it is.

The analysis in this section can accommodate various economic settings. For instance, a DM who contemplates the optimal portfolio choice of financial assets needs to take into account the expected returns and volatility of different assets as well as the comovement of their prices; a DM with both labor income and capital income has to consider the probability of getting laid off, the probability of a financial crisis, and the probability of both happening simultaneously; a DM who decides whether to purchase small-scale insurance for her brand new smart phone should bear in mind the background risk of lifetime income and how it might correlate with the risk associated with the smart phone. Unless otherwise specified, we will keep the terminology of portfolio choices.

We assume that the outcome space in each income source is a compact interval $X_1 = X_2 = X := [-\bar{x}, \bar{x}] \in \mathbb{R}$ with $\bar{x} > 0$ sufficiently large. The outcome $x_i \in X$ in each source can be interpreted as the monetary income in source $i \in \{1, 2\}$, and negative values of x_i represent losses.

Recall that $\succsim_{1|r}$ denotes the conditional preference in source 1 given marginal lottery $r \in \mathcal{L}^0(X)$ in source 2. When $r = \delta_0$, we call it the *narrow preference in source 1* and denote it by $\succsim_1$, that is, $p \succsim_1 q$ if and only if $(p, 0) \succsim (q, 0)$. Since the marginal lottery in source 2 is fixed at δ_0, the comparison of lotteries $(p, 0)$ and $(q, 0)$ can be interpreted as the comparison of p and q in source 1 as if source 2 does not exist. Symmetrically, $\succsim_2$ denotes the narrow preference in source 2. We assume that narrow preferences in two sources are identical since both of them represent the decision maker's preference over risky prospects of money.

Symmetry: $\succsim_1 = \succsim_2$.

Along with Monotonicity, Continuity and part (i) of Independence*, Symmetry implies that the two narrow preferences $\succsim_1$ and $\succsim_2$ admit the same expected utility representation with index $u : X \to \mathbb{R}$ continuous and strictly increasing.

Now we formally introduce the notion of first order stochastic dominance. If the DM's income streams in two sources are determined by the lottery $P \in \mathcal{P}$, then the distribution over her final wealth can be denoted by $f[P] \in \mathcal{L}^0(\mathbb{R})$ where for each $z \in \mathbb{R}$,

$$f[P](z) = \sum_{(x,y):x+y=z} P(x, y).$$

For any two distributions $p, q \in \mathcal{L}^0(\mathbb{R})$, we say that p first order stochastically dominates q, denoted by $p \succ_{FOSD} q$, if $p \neq q$ and $\sum_{x \leq z} q(x) \geq \sum_{x \leq z} p(x)$ for all $z \in \mathbb{R}$. Then we define that $\succsim$ satisfies *(first order) stochastic dominance* if and only if for each $P, Q \in \mathcal{P}$, $f[P] \succ_{FOSD} f[Q]$ implies that $P \succ Q$ and $f[P] = f[Q]$ implies that $P \sim Q$.

Intuitively, a rational DM cares about final wealth and prefers more money than less, and hence would choose P over Q if $f[P] \succ_{FOSD} f[Q]$ since P assigns higher probability to higher levels of final wealth. This suggests that the preference of a rational DM should satisfy stochastic dominance. However, experimental evidence shows that many subjects violate stochastic dominance. Consider the following experiment studied in Tversky and Kahneman (1981) and Rabin and Weizsäcker (2009).

Example 1. Suppose you face the following pair of concurrent decisions. All lotteries are independent. First examine both decisions, then indicate your choices. Both choices will be payoff relevant, i.e., the gains and losses will be added to your overall payment.

Decision 1: Choose between:

A. A sure gain of $2.40.

B. A 25 percent chance to gain $10.00, and a 75 percent chance to gain $0.

Decision 2: Choose between:

C. A sure loss of $7.50.

D. A 75 percent chance to lose $10.00, and a 25 percent chance to lose $0.

Across different treatments in Tversky and Kahneman (1981) and Rabin and Weizsäcker (2009), a large fraction of subjects choose A in decision 1 and D in decision 2. Those choices violate stochastic dominance because

$$f[(B,C)] = \frac{3}{4}\delta_{-7.50} + \frac{3}{4}\delta_{2.50} \succ_{FOSD} \frac{3}{4}\delta_{-7.60} + \frac{3}{4}\delta_{2.40} = f[(A,D)].$$

Such violations are stark since the lottery over final wealth induced by B and C is that induced by A and D plus a sure gain of $0.10.

Like other experiments on violations of stochastic dominance, Example 1 involves risk in at least some of the options. This design makes sense because when all options are deterministic, the decision problem is so simple that one can confidently believe that almost all subjects should be able to sum up the monetary prizes across different sources and choose the option delivering the highest final wealth. Hence, we further assume that the DM does not violate stochastic dominance when there is no risk.

Stochastic Dominance without Risk: For each $x, y, x', y' \in X$, $(x, y) \succsim (x', y')$ if and only if $x + y \geq x' + y'$.

Then we have the following direct corollary of Theorem 2.

Corollary 1. *Suppose that $\succsim$ satisfies the axioms stated in Theorem 2 and can be represented by one of V^{EU}, V^{ANB}, V^{CN}, V^{ANB-CN} or V^{FNB}. Then $\succsim$ satisfies Symmetry and Stochastic Dominance without Risk if and only if those utility functions can be rewritten such that for each $P \in \mathcal{P}$,*

$$V^{EU}(P) = \sum_{x,y} u(x+y)P(x,y);$$

$$V^{CN}(P) = \sum_{x,y} u(x+y)P_1(x)P_2(y);$$
$$V^{ANB}(P) = \sum_{x} u(x + CE_u(P_{2|x}))P_1(x);$$
$$V^{ANB-CN}(P) = \sum_{x} u(x + CE_u(P_2))P_1(x);$$
$$V^{FNB}(P) = u(CE_u(P_1) + CE_u(P_2)).$$

where $u : [-2\bar{x}, 2\bar{x}] \to \mathbb{R}$ is continuous and strictly monotonic.

Since u is strictly monotonic, a preference which is represented by V^{FNB} can also be represented by $CE_u(P_1) + CE_u(P_2)$. Hence, according to Corollary 1, a DM with full narrow bracketing evaluates a lottery over two income sources by summing up the *certainty equivalents* of the marginal lotteries in each source. In contrast, an alternative utility function of full narrow bracketing, which has been widely applied in the behavioral and experimental literature, entails the summation of *expected utilities* of the marginal lotteries. Concretely, the utility of $P \in \mathcal{P}$ is

$$\hat{V}^{FNB}(P) = \sum_{x} u(x)P_1(x) + \sum_{y} u(y)P_2(y).$$

$\hat{V}^{FNB}$ can violate stochastic dominance even without risk. To see this intuitively, suppose the DM is choosing between a portfolio P which delivers \$1 in both assets, with a portfolio Q which delivers \$2 in asset one and \$0 in asset two. If the DM is risk averse, then she will strictly prefer portfolio P to portfolio Q, although holding either portfolio delivers total income \$2 for sure. Hence, V^{FNB} should be more reasonable than $\hat{V}^{FNB}$ in general portfolio choice problems. For most applications considered in the literature, replacing $\hat{V}^{FNB}$ with V^{FNB} will not significantly alter the qualitative results.

Now we show that V^{FNB} can explain violations of stochastic dominance in Example 1.

Example 1 (Continued). Suppose that the subject's preference is represented by V^{FNB} with

$$u(x) = \begin{cases} \sqrt{x}, & \text{if } x \geq 0, x \in X, \\ -2\sqrt{-x}, & \text{if } x < 0, x \in X. \end{cases} \tag{1}$$

u is a gain-loss utility index with CRRA risk preference and loss aversion parameter 2, which is a standard parametrization in the behavioral literature. Then the subject will simultaneously choose A in decision 1 and D in decision 2 since

$$CE_v(A) = 2.4 > 0.625 = CE_v(B), CE_v(D) = -5.625 > -7.5 = CE_v(C).$$

Since lotteries are assumed to be independent from each other, with a similar parametrization of u, violations of stochastic dominance can also be explained by V^{ANB} and V^{ANB-CN}.

Before moving on, we briefly discuss the *partial narrow bracketing* model studied in Barberis, Huang, and Thaler (2006) and Rabin and Weizsäcker (2009) and Ellis and Freeman (2020). Without violating Stochastic Dominance without Risk, the model can be adapted to our setting as follows:

$$\begin{aligned} V^{PNB}(P) &= \alpha V^{EU}(P) + (1-\alpha)V^{FNB} \\ &= \alpha \sum_{x,y} u(x+y)P(x,y) + (1-\alpha)u(CE_u(P_1) + CE_u(P_2)), \end{aligned}$$

for each $P \in \mathcal{P}$ where $\alpha \in [0,1]$ and $1-\alpha$ determines the degree of narrow bracketing.

The partial narrow bracketing utility function V^{PNB} does not satisfy the axioms in Theorem 2 and Corollary 1 and is not included from our characterization of narrow bracketing. We justify such exclusion in two ways. From the theoretical point of view, computing the utility level of partial narrow bracketing necessitates computing the utility levels of expected utility and full narrow bracketing separately, and hence is arguably more complex and demanding. This contradicts with our interpretation that narrow bracketing is a simplifying heuristic for the multi-source risk. Empirically, Ellis and Freeman (2020) show that very few subjects are best described by non-trivial partial narrow bracketing with $\alpha \in (0,1)$. This suggests that incorporating such intermediate degrees of narrow bracketing might not significantly increase the model's explanatory power in practice.

6.1.1 Ranking of Rationality

In this section, we develop a ranking of rationality over models in Corollary 1 based on the readiness with which they yield violations of stochastic dominance.

We first define stochastic dominance on a subdomain of lotteries. A *choice environment* is a subset of lotteries $\mathcal{C} \subseteq \mathcal{P}$. A preference $\succsim$ is said to satisfy *stochastic dominance on* $\mathcal{C}$ if for any $P, Q \in \mathcal{C}$, $f[P] \succ_{FOSD} f[Q]$ implies that $P \succ Q$ and $f[P] = f[Q]$ implies that $P \sim Q$. Moreover, we say that the EU model satisfies stochastic dominance on $\mathcal{C}$ if for any $\succsim$ that can be represented by some V^{EU} in Corollary 1, $\succsim$ satisfies stochastic dominance on $\mathcal{C}$. Similar notions for other models including CN, ANB, ANB-CN and FNB can be defined.

For two choice environments $\mathcal{C}$ and $\mathcal{C}'$ with $\mathcal{C}' \subseteq \mathcal{C}$, if $\succsim$ satisfies stochastic dominance on $\mathcal{C}$, then by definition it also satisfies stochastic dominance on $\mathcal{C}'$. Hence we can define a complexity order $\trianglerighteq$ over choice environments such that $\mathcal{C}$ is more complex than $\mathcal{C}'$, i.e., $\mathcal{C} \trianglerighteq \mathcal{C}'$, if and only if $\mathcal{C}' \subsetneq \mathcal{C}$. Our goal is to construct choice environments differing in the complexity levels in order to discriminate among the models characterized in Corollary 1.

The most complex choice environment is the set of all lotteries $\mathcal{P}$. Another natural choice environment is the riskless choice environment $\mathcal{C}_1 := Z$, which is the set of all degenerate lotteries. All five models in Corollary 1 satisfy stochastic dominance on $\mathcal{C}_1$ and it is the simplest environment we consider here. However, $\mathcal{C}_1$ is non-trivial since models featuring summation of expected utilities like $\hat{V}^{FNB}$ violate stochastic dominance on it. We define the following three choice environments whose complexity levels lie between $\mathcal{P}$ and $\mathcal{C}_1$:

- $\mathcal{C}_2 := \mathcal{L}^0(X) \times X$ is the choice environment where there is no risk in source 2;
- $\mathcal{C}_3 := \left\{P \in \mathcal{P} : P_{2|x} \in X,\ \forall x \in X\right\}$ is the choice environment where there is no conditional risk in source 2 give the income in source 1;
- $\mathcal{C}_3' := \hat{\mathcal{P}} = \mathcal{L}^0(X)^2$ is the choice environment where there is no correlation between income in two sources;

Denote $\mathcal{C}_4 := \mathcal{P}$. We have the following complexity ranking of choice environments:

$$\mathcal{C}_4 \trianglerighteq \mathcal{C}_3, \mathcal{C}_3' \trianglerighteq \mathcal{C}_2 \trianglerighteq \mathcal{C}_1$$

where $\mathcal{C}_3$ and $\mathcal{C}_3'$ cannot by ranked by $\trianglerighteq$.

The next proposition shows that the above five complexity environments can be used to distinguish the five models in Corollary 1.

Proposition 2. *(i). FNB satisfies stochastic dominance on $\mathcal{C}_1$, but not on $\mathcal{C}_2$;*

(ii). ANB-CN satisfies stochastic dominance on $\mathcal{C}_2$, but not on $\mathcal{C}_3$ or $\mathcal{C}_3'$;

(iii). ANB satisfies stochastic dominance on $\mathcal{C}_3$, but not on $\mathcal{C}_3'$;

(iv). CN satisfies stochastic dominance on $\mathcal{C}_3'$, but not on $\mathcal{C}_3$;

(v). EU satisfies stochastic dominance on $\mathcal{C}_4$.

We say model A is *more rational* than model B if model A satisfies stochastic dominance on a more complex choice environment than model B. This rationality ranking reflects how readily a model yields violations of stochastic dominance. Proposition 2 reveals the rationality ranking of five models as depicted in Figure 2. In the graph, model A is more rational than model B if and only if there is a path from A to B. Hence, EU is more rational than CN and ANB, which are more rational than ANB-CN, while FNB is the least rational.

The rationality ranking in Figure 2 is consistent with the degree of behavioral heuristics employed in each model as shown in Figure 1. The more and stronger heuristics a model entails, the more it deviates from the expected utility benchmark, the more likely it is to yield violations of stochastic dominance and the more irrational it is. This also explains why CN and ANB cannot be ranked in terms of rationality, because they adopt distinct behavioral heuristics.

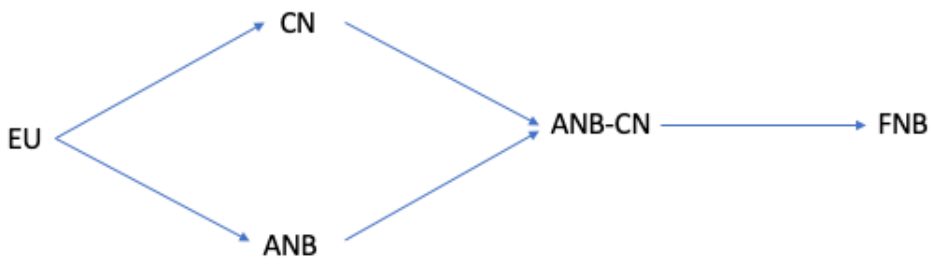

Figure 2: Rationality ranking of models based on violations of stochastic dominance.

6.1.2 Risk Aversion over Small Gambles

In this section, we apply narrow bracketing to problems with background risk. Consider a decision maker who is contemplating whether to purchase small-scale insurance for her new smart phone. Since the risk of a smart phone is negligible compared to the risk in the lifetime wealth, a rational agent should take into account this background risk and be almost risk neutral regarding the risk of a smart phone. The popularity of small-scale insurance in practice reveals the prevalence of narrow bracketing when there is background risk.

Below we briefly show that narrow bracketing can explain risk aversion over small gambles. Consider the calibration result of Rabin (2000) which states that if an EU maximizer turns down 50-50 gambles of losing \$1000 or gaining \$1050 for all initial wealth levels, then she would always turn down 50-50 gambles of losing \$20,000 or gaining any sum. This serves as a critique on the expected utility theory since the DM has plausible risk aversion over small gambles, but is implausibly risk averse over large gambles.

Suppose that the DM's preference is represented by V^{FNB} with the utility index u given by equation (1) instead of V^{EU}. Then we can show that she will also reject the above small gambles, but she is willing to accept a 50-50 gamble of losing \$20,000 or gaining slightly more than \$80,000. Similar results can be derived if source 1 represents the background wealth and source 2 represents the outcome of the gamble, and the DM's preference is represented by an asymmetric narrow bracketing (with or without correlation neglect) utility function.

6.2 Time Preferences

In this section, we study the implications of narrow bracketing and correlation neglect in time preferences. Each outcome profile represents a consumption path in two periods $t = 1, 2$. A DM who finds it complex to evaluate intertemporal risk might resort to heuristics.

For simplicity, we assume that the consumption space in each period is a compact interval $C := [\underline{c}, \overline{c}] \subset \mathbb{R}_+$ and focus on preferences that satisfy the following assumption.

Discounted Utility without Risk: There exists a continuous and strictly increasing function $u : C \to \mathbb{R}$ and $\beta \in (0, 1)$ such that for all $x_1, x_2, y_1, y_2 \in C$, $(x_1, x_2) \succsim (y_1, y_2)$ if and only if $u(x_1) + \beta u(x_2) \geq u(y_1) + \beta u(y_2)$.

This assumption follows from Dillenberger, Gottlieb, and Ortoleva (2020) and it postulates that when there is no risk, preferences can be represented by the summation of the discounted utility of consumption in different periods. This is true for most models of time preferences in the literature.

Corollary 2. *Suppose that $\succsim$ satisfies the axioms stated in Theorem 2 and can be represented by one of V^{EU}, V^{ANB}, V^{CN}, V^{ANB-CN} or V^{FNB}. Then $\succsim$ satisfies Discounted Utility without Risk if and only if the index $w : C^2 \to \mathbb{R}$ in each utility function can be rewritten as*

$$w(x_1, x_2) = \phi(u(x_1) + \beta u(x_2)), \forall (x_1, x_2) \in C^2,$$

where $\phi : [(1 + \beta)u(\underline{c}), (1 + \beta)u(\bar{c})] \to \mathbb{R}$ is continuous and strictly monotonic.

Below we show that different utility functions in Corollary 2 correspond to different commonly used time preferences models in the literature.

Example 2 (EU). V^{EU} takes the form of

$$V^{EU}(P) = \sum_{x_1, x_2} \phi(u(x_1) + \beta u(x_2))P(x_1, x_2),$$

which is the Kihlstrom-Mirman (KM) representation studied in Kihlstrom and Mirman (1974) and Dillenberger, Gottlieb, and Ortoleva (2020). If ϕ is affine, then $\succsim$ can be represented by $\sum_{x_1} u(x_1)P_1(x_1) + \beta \sum_{x_2} u(x_2)P_2(x_2)$, the standard Expected Discounted Utility (EDU) model.[2]

Example 3 (FNB). Suppose $v_1 = v_2 := v$, that is, narrow preferences in both sources are identical, then V^{FNB} takes the form of

$$V^{FNB}(P) = u(CE_v(P_1)) + \beta u(CE_v(P_2)).$$

The DM first narrowly evaluates the marginal risk in each period and then aggregates the certainty equivalents using discounted utility. Note that we can normalize ϕ to be the identity function since w is unique up to a monotonic transformation in V^{FNB}. This functional form is the adaption of the Dynamic Ordinal Certainty Equivalent (DOCE) model (Selden, 1978, Selden and Stux, 1978, Kubler, Selden, and Wei, 2020b) to our setup. However, it emerges from narrow bracketing in our model, instead of normative properties of time preferences.

[2] EDU can also be derived from FNB in Example 3 by assuming $v = u$, and from ANB in Example 4 by assuming $\phi(x) = x$ for all x and $v = u$.

Example 4 (ANB). V^{ANB} takes the form of

$$V^{ANB}(P) = \sum_{x_1} \phi(u(x_1) + \beta u(CE_v(P_{2|x}))P_1(x_1). \tag{2}$$

If we fix the consumption levels for periods 0 and $t > 2$, then this functional form corresponds to a majority of recursive preferences built upon Kreps and Porteus (1978). The DM evaluates risk in different periods recursively. Given the current consumption, she first reduces the conditional future risk to its certainty equivalent; then she evaluates this certainty equivalent and the current consumption using discounted utility; finally she moves one period ahead and repeats the above process. Like DOCE in Example 3, such features also emerge from narrow bracketing, instead of a preference for timing of resolution of uncertainty as in recursive preferences. In Section 6.2.1, we will illustrate this point in detail by comparing ANB in (2) with the Epstein-Zin (EZ) preferences in Epstein and Zin (1989) and Weil (1990).

6.2.1 Epstein-Zin Type Preferences

In this section, we focus on the ANB utility function in (2) and show that it incorporates EZ type preferences. That is, an EZ type utility function emerges because of narrow bracketing. Moreover, we will argue that this new foundation of EZ type preferences might be more suitable for some applications than the original foundation in Epstein and Zin (1989) since it can address the critique in Epstein, Farhi, and Strzalecki (2014).

To facilitate the comparison of ANB and EZ preferences, we assume that the consumption in the first period is deterministic and focus on the restriction of $\succsim$ on $C \times \mathcal{L}^0(C)$.[3] Consider the following specification of ANB

$$V^{ANB}(c_1, p) = \frac{1}{\alpha}\Big\{c_1^\rho + \beta[\mathbb{E}_p(c_2^\alpha)]^{\rho/\alpha}\Big\}^{\alpha/\rho},\ \forall\ (c_1, q) \in C \times \mathcal{L}^0(C). \tag{3}$$

where $0 \neq \rho < 1, 0 \neq \alpha < 1$ and $0 < \beta < 1$. Note that $\gamma := 1 - \alpha \in \mathbb{R}_+\backslash\{1\}$ is the coefficient of relative risk aversion (RRA) and $\psi = \frac{1}{1-\rho} \in \mathbb{R}_+\backslash\{1\}$ is the coefficient of elasticity of intertemporal substitution (EIS). (3) can be derived from (2) by assuming $u(x) = \frac{x^\rho}{\rho}$, $v(x) = \frac{x^\alpha}{\alpha}$ and $\phi(x) = v \circ u^{-1}(x)$. When $\alpha = \rho$, i.e., RRA equals the reciprocal of EIS, (3) reduces to EDU.

Since $f(x) = \frac{x^\alpha}{\alpha}$ is strictly monotonic, f^{-1} is well-defined and is also strictly monotonic. We impose the monotonic transformation f^{-1} on V^{ANB} without changing the preference on $C \times \mathcal{L}^0(C)$:

$$\hat{V}^{ANB}(c_1, p) = \Big\{c_1^\rho + \beta[\mathbb{E}_p(c_2^\alpha)]^{\rho/\alpha}\Big\}^{1/\rho},\ \forall\ (c_1, p) \in C \times \mathcal{L}^0(C). \tag{4}$$

[3] In the appendix, we briefly discuss how to extend the ANB model in an infinite horizon setting.

By comparison, the standard CRRA-CES EZ preference entails the following recursive formulation of continuation utility U_t:

$$U_t = \left\{c_t^\rho + \beta\Big[\mathbb{E}_t\Big(U_{t+1}^\alpha\Big)\Big]^{\rho/\alpha}\right\}^{1/\rho}, \tag{5}$$

Note that the ANB utility function (4) is quite similar to the recursive EZ utility function (5). However, we cannot simply apply (5) to $C \times \mathcal{L}^0(C)$ since EZ is built upon a richer space than the set of lotteries over consumption paths. The richer space involves temporal resolution of uncertainty. Following Kreps and Porteus (1978), we call it the set of *temporal lotteries* and denote by d a generic temporal lottery. The difference between a temporal lottery and a lottery can be illustrated by the following simple example.

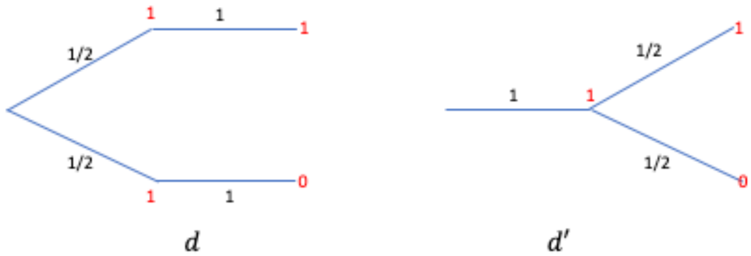

Figure 3: Example of temporal lotteries inducing the same lottery.

Consider the two temporal lotteries d and d' depicted in Figure 3. Both d and d' deliver consumption 1 for sure at date 1 and have probability $1/2$ to deliver either consumption 1 or consumption 0 at date 2; that is, d and d' induce the same lottery over consumption paths. However, they differ in timing of resolution of risk. In temporal lottery d, the consumer knows her future consumption at date 1, while in temporal lottery d', the risk about consumption at date 2 is only resolved at date 2.

EZ captures the consumer's attitude towards the above difference in timing of resolution of risk. Formally, for any lottery $(c_1, p) \in C \times \mathcal{L}^0(C)$, we denote by d^E the temporal lottery that induces lottery (c_1, p) and has early resolution of risk, like d in Figure 3. Similarly, we denote by d^L the temporal lottery that induces lottery (c_1, p) and has late resolution of risk, like d' in Figure 3. According to the EZ functional form in (5), the EZ utility of d^E and d^L are given by

$$V^{EZ}(d^E) = \left\{\mathbb{E}_p\Big[(c_1^\rho + \beta c_2^\rho)^{\alpha/\rho}\Big]\right\}^{1/\alpha},$$
$$V^{EZ}(d^L) = \left\{c_1^\rho + \beta[\mathbb{E}_p(c_2^\alpha)]^{\rho/\alpha}\right\}^{1/\rho}.$$

One can easily observe that if (c_1, p) is induced by d^E and d^L, then $V^{EZ}(d^L) = \hat{V}^{ANB}(c_1, p)$ and $V^{EZ}(d^E)$ agrees with a monotone transformation of $V^{EU}(c, p)$ with the same preference parameters, i.e., $V^{EU}(c,p) = \frac{1}{\alpha}\mathbb{E}_p\Big[(c_1^\rho + \beta c_2^\rho)^{\alpha/\rho}\Big]$. Hence, the ANB utility function $\hat{V}^{ANB}$ can be interpreted as the restriction of the EZ utility function V^{EZ} on temporal lotteries where there is no early resolution of uncertainty.

It is well-known that when $\rho > \alpha$, that is, when RRA $> 1/$EIS, a consumer with the EZ utility function exhibits a preference for early resolution of risk and strictly prefers d to d'. Symmetrically, she has a preference for late resolution of risk when $\rho < \alpha$. As a result, the consumer is indifferent to timing of risk resolution and attaches the same value to d and d' if and only if $\alpha = \rho$, in which case EZ reduces to EDU. Hence the EZ utility is non-trivial and separates the time preference (EIS) and the risk preference (RRA) only if it entails a preference for either early or late resolution of risk.

In contrast, although EZ type preferences can also be incorporated in ANB as in (4), it emerges from narrow bracketing, instead of a non-trivial attitude towards timing of risk resolution. Actually, since d and d' induce the same lottery over consumption paths, a consumer with ANB utility function (4) is naturally indifferent between d and d' and hence exhibits indifference to timing of resolution of risk.

Our alternative foundation of EZ type preferences based on narrow bracketing has the following implications. Theoretically, it shows that one can simultaneously allow for indifference to temporal resolution of risk and separation of time and risk preferences, which is impossible in the original foundation in Epstein and Zin (1991) based on the framework of Kreps and Porteus (1978). It reveals a novel and surprising connection between the EZ type utility function, the most commonly used alternative to EDU, and narrow bracketing, the behavioral heuristic to simplify the evaluation of multi-source risk. Finally, in order to perform a thorough test of the behavioral implications of an EZ-type utility function, our foundation only requires the econometrician to observe the DM's choices in the natural domain of lotteries over consumption paths, while the original foundation entails choices in the more complex domain of temporal lotteries.

The difference of the two utility functions is also empirically relevant and important. Thanks to the separation of time and risk preferences and the convenient functional form, EZ has been widely applied in macroeconomics and finance. For instance, the long-run risks model of Bansal and Yaron (2004) considers a representative agent with CRRA-CES EZ preferences and delivers a unified explanation for several long-standing puzzles of asset markets, including the equity premium puzzle, the excessive asset price volatility, and the large cross-sectional differences in average returns across equity portfolios. Their main empirical results are based on parameters with a RRA of 10 and an IES of 1.5, but have ignored the quantitative implications of the strong preference for early resolution of risk. Through introspection, Epstein, Farhi, and Strzalecki (2014) show that the preference parameters and calibrated endowment process imply that the representative agent in the long-run risks model is willing to give up 25 or 30 percent of her lifetime consumption in order to have all risk about future consumption resolved in the next period. This timing premium is unrealistically high since the risk is about future consumption instead of future income or asset returns, and hence such information has no apparent instrumental value. In other words, the representative agent has no need to reoptimize her contingent consumption plans given early resolution of risk. Meissner and Pfeiffer (2022) show that on average subjects give up 5% of their total consumption to resolve all uncertainty immediately in an experiment.[4] Epstein, Farhi, and Strzalecki (2014) also show that the high timing premium is robust to other models using EZ preferences (such as the rare disasters model in Barro (2009)), different preference parameter values and more general risk preferences. In contrast, in the EZ type preferences based on ANB, the representative agent attaches no value to the non-instrumental information and the timing premium is always zero.

We conclude the section with a final comment. Note that ANB only specifies the preference on lotteries over consumption paths. Although the most natural way to extend ANB to temporal lotteries is assuming indifference to timing of resolution of risk as we did in this paper, one can also extend ANB by assuming a preference for early or late resolution of uncertainty, which we leave for future research. The main message of our alternative foundation of EZ type preferences is that the attitude towards temporal resolution of uncertainty can be isolated from the separation of time and risk preferences.

6.2.2 Effects of Heuristics on Optimal Saving

In this section, we illustrate how narrow bracketing and correlation neglect affect the saving behavior of a consumer who faces income uncertainty and deterministic interest rates through a parsimonious two-period model. One can interpret our analysis as comparative statics for simplifying heuristics, similar to the comparative statics for future income risk or risk attitude in the literature.

Following Selden (1978) and many papers on precautionary savings including Drèze and Modigliani (1972) and Kimball (1990), we start with the standard "certain $\times$ uncertain" setup, where there is only one saving decision (in the first period) and income is uncertain in a single period (in the second period). Concretely, at date 1, the consumer receives income $y_1 > 0$ which is certain and can be allocated between consumption and saving. At date 2, state $i \in \{1, \dots, N\}$ is realized with probability $\pi_i \in (0,1)$ and $\sum_{i=1}^N \pi_i = 1$, which determines the level of date 2 income $y_{2,i} > 0$ with $y_{2,i} \neq y_{2,j}$ for $i \neq j$. The gross return on savings $R > 0$ is independent of the state at date 2. The consumer needs to choose the saving s

[4] Meissner and Pfeiffer (2022) also show a negative correlation between predicted and elicited timing premia in EZ, which questions the structural connection between the preference for early resolution of uncertainty and the preference parameters in EZ.

at date 1 before observing the state of the world. We allow s to be negative, in which case the consumer borrows at date 1 with interest rate R, given that her future income in any realized state can cover the debt. Denote c_1 as the consumption level at date 1 and $c_{2,i}$ as the consumption level at date 2 when state i occurs. The budget constraints are given by

$$\begin{aligned} y_1 - s &= c_1 \geq 0 \\ y_{2,i} + Rs &= c_{2,i} \geq 0 \text{ for all } i = 1, \ldots, N. \end{aligned}$$

Each choice of saving s induces a lottery over 2-period consumption paths P such that P_1 is degenerate at c_1 and $P_2(c_{2,i}) = \pi_i$ for each $i = 1, \ldots, N$. Hence, it suffices to specify the consumer's preference over $C \times \mathcal{L}^0(C) \subset \hat{\mathcal{P}}$. Given the same preference parameters, among the different utility functions characterized in Corollary 2, V^{EU} and V^{CN} coincide on $C \times \mathcal{L}^0(C)$, and V^{ANB}, V^{ANB-CN} and V^{FNB} coincide on $C \times \mathcal{L}^0(C)$. We denote the former by V^{EU} and the later by V^{NB}. [5] This observation suggests that correlation neglect has no bite in this optimal saving problem and we can focus on the effects of narrow bracketing.

For computational simplicity, we further assume that V^{NB} admits a CRRA-CES version with unitary EIS, where we apply (2) to $C \times \mathcal{L}^0(C)$ with $u(x) = \log x$, $v(x) = \frac{x^\alpha}{\alpha}$ and $\phi(x) = v \circ u^{-1}(x)$ with $\alpha < 1$ and $\alpha \neq 0$:

$$V^{NB}(c_1, q) = \frac{1}{\alpha} c_1^\alpha \cdot \Big[\sum_{c_2} c_2^\alpha q(c_2)\Big]^\beta, \ \forall \ (c_1, q) \in C \times \mathcal{L}^0(C).$$

Using the same preference parameters, the expected utility benchmark is V^{EU} where

$$V^{EU}(c_1, q) = \frac{1}{\alpha} c_1^\alpha \cdot \sum_{c_2} c_2^{\beta\alpha} q(c_2), \ \forall \ (c_1, q) \in C \times \mathcal{L}^0(C).$$

Denote by s^{EU} and s^{NB} the optimal levels of savings if the consumer's preference is represented by V^{EU} and V^{NB} respectively. The difference between s^{EU} and s^{NB} reflects the impact of narrow bracketing on the optimal saving behavior, isolated from the preference parameters and correlation neglect.

Lemma 1 **(Saving with narrow bracketing).** *Consider the "certain $\times$ uncertain" saving problem described above. Then $s^{NB} > s^{EU}$ if $\alpha < 0$ and $s^{NB} < s^{EU}$ if $\alpha > 0$.*

The intuition behind Lemma 1 is straightforward. With the same preference parameters, the effective coefficient of RRA for consumption risk in period 2 is $1 - \beta\alpha$ in the EU benchmark, and $1 - \alpha$ in the model with narrow bracketing. Since the discount factor

[5] The observation that ANB and FNB, and hence the models in Kreps and Porteus (1978) and Selden (1978), agree in the "certain $\times$ uncertain" setup also appears in Bommier, Chassagnon, and Le Grand (2012) and Kubler, Selden, and Wei (2020a).

$\beta \in (0,1)$, when $\alpha < 0 = \rho$, we have $1-\alpha > 1-\beta\alpha$, which suggests that narrow bracketing makes the consumer effectively more risk averse regarding the consumption risk in period 2. This intuition remains valid for the general CRRA-CES model in (3) with $\rho > \alpha$. Although it is well-known that the relationship between the optimal saving level and the risk aversion parameter α is non-monotonic in the EZ model,[6] we show that the increased effective risk aversion due to narrow bracketing does lead to higher saving in our setup. Since the preference over deterministic consumption paths are the same in V^{NB} and V^{EU}, our result also implies that narrow bracketing increases precautionary saving when $\alpha < 0$.

Recall that RRA is $1-\alpha$ and EIS is $\frac{1}{1-\rho}$. For $\rho = 0$, RRA $> 1/$EIS if and only if $\alpha < 0$. In order to explain the equity premium puzzle and other puzzles in the financial market using the EZ preference, the empirical asset pricing literature, such as the long-run risks model (Bansal and Yaron, 2004) and the rare disasters model (Barro, 2009), relies on parameters satisfying RRA $> 1/$EIS. Our Lemma 1 provides some intuition why this choice of parameters can work based on the idea that an EZ type utility function emerges because of narrow bracketing. When RRA $> 1/$EIS, a consumer with narrow bracketing is effectively more risk averse regarding future consumption risk compared to a consumer with EU and the same RRA parameter $1-\alpha$. As a result, the EZ type utility function can explain the excessive equity premium with the RRA parameter $1-\alpha$ in a reasonable range.

In order to study the effect of the other heuristic, correlation neglect, on the consumer's optimal saving behavior, we need to modify the previous setup to allow for uncertainty over consumption in the first period. We assume that income is uncertain in the second period, but the consumer knows her future income in the first period and can tailor her saving decision to such information. Although the consumer is indifferent to temporal resolution of uncertainty and does not attach any value to non-instrumental information, she does value such instrumental information about the future income since it can improve the quality of her saving decisions in the first period. For simplicity, we further assume that the income in the second period is either $y_{2,h}$ or $y_{2,l}$, each with probability $1/2$ and $y_{2,h} > y_{2,l} > 0$. We will refer to this as the "uncertain $\times$ conditionally certain" saving problem, since there is no income uncertainty date 2 given the information in date 1.

In this "uncertain $\times$ conditionally certain" setup, the consumer needs to choose the saving level s_i at date 1 conditional on the state $i = h, l$. Denote by $c_{1,i}$ $c_{2,i}$ the consumption levels at date 1 and date 2 respectively when state i occurs. Then the budget constraints are

$$y_1 - s_i = c_{1,i} \geq 0 \text{ for } i = l, h,$$
$$y_{2,i} + Rs_i = c_{2,i} \geq 0 \text{ for } i = l, h.$$

[6]See Kimball and Weil (2009) and Bommier, Chassagnon, and Le Grand (2012). Bommier, Kochov, and Le Grand (2017) illustrates this non-monotonic relationship in a simple example.

Each choice of saving levels s_l and s_h induces a lottery over consumption paths P where $P(c_{1,l}, c_{2,l}) = P(c_{1,h}, c_{2,h}) = 1/2$. If $s_h \neq s_l$, then $c_{1,h} \neq c_{1,l}$ and there is risk about consumption at date 1 but no conditional risk about consumption at date 2. If instead $s_h = s_l$, then $c_{1,h} = c_{1,l}$ and the consumption at date 1 is certain and there is risk about consumption at date 2.

We assume that the consumer either has expected utility or adopts the correlation neglect heuristic, that is, her preference $\succsim$ is represented by one of the following two utility functions

$$V^{EU}(P) = \frac{1}{\alpha} \sum_{c_1, c_2} c_1^{\alpha} \cdot c_2^{\beta\alpha} P(c_1, c_2),$$
$$V^{CN}(P) = \frac{1}{\alpha} \sum_{c_1, c_2} c_1^{\alpha} \cdot c_2^{\beta\alpha} P_1(c_1) P_2(c_2).$$

Denote by s_i^{EU} and s_i^{CN}, $i = l, h$, the optimal levels of savings in two states for the utility function V^{EU} and V^{CN} correspondingly. The different between s_i^{CN} and s_i^{EU} reflects the impact of correlation neglect on the optimal saving behavior in state $i = l, h$, isolated from the preference parameters and narrow bracketing.

Lemma 2 (Saving with correlation neglect). *Consider the "uncertain $\times$ conditionally certain" saving problem described above. The following results hold:*

(i). If $\alpha < 0$, then $s_h^{EU} < s_h^{CN} < s_l^{CN} < s_l^{EU}$;

(ii). If $\alpha > 0$, then $s_h^{CN} < s_h^{EU} < s_l^{EU} < s_l^{CN}$.

The effects of correlation neglect on the optimal saving behavior depends on whether the consumer is correlation averse or correlation seeking in the EU benchmark. Following Bommier (2007), we say the consumer is *correlation averse* if and only if for all $x_1, x_2, y_1, y_2 \in C$ with $x_1 < x_2$ and $y_1 < y_2$, the lottery $\frac{1}{2}(x_1, y_2) + \frac{1}{2}(x_2, y_1)$ is preferred to the lottery $\frac{1}{2}(x_1, y_1) + \frac{1}{2}(x_2, y_2)$. In other words, a correlation averse consumer prefers consumption at different dates to be negatively correlated rather than positive correlated. For V^{EU} considered in this section, the consumer is correlation averse if and only if $\alpha < 0$. Compared to a consumer with correlation neglect, a correlation averse consumer adjusts the current saving levels to compensate for the difference in future income. She saves more when the future income is low and less when the future income is high. A symmetric argument holds for the comparison between correlation seeking and correlation neglect.

7 Conclusion

In this paper, we study narrow bracketing and correlation neglect as two heuristics that decision makers adopt to simplify the evaluation of risk from multiple sources. By relaxing the independence axiom of the expected utility model, we develop a class of models that (i) allow for narrow bracketing or correlation neglect or both and (ii) allow for different degrees of deviations from the benchmark without narrow bracketing. We interpret the different sources as different streams of income to explain experimental evidence on violations of first order stochastic dominance. We propose a rationality ranking of our models based on how likely they are to yield violations of stochastic dominance. Then, we interpret one source as background risk and show that narrow bracketing can explain risk aversion over small gambles. Finally, we interpret the different sources as consumption in different periods to explore the implications of narrow bracketing and correlation neglect on the optimal saving behavior of a consumer. We also show that an Epstein-Zin type utility function can emerge because of narrow bracketing. This new utility function achieves separation between time and risk preferences without inducing an implausibly high timing premium as observed by Epstein, Farhi, and Strzalecki (2014). My results show that both the separation of time and risk preferences and violations of stochastic dominance can result from narrow bracketing and both can be captured by relaxing the independence axiom.

Our analysis suggests several promising directions for future research. In the appendix, we briefly discuss how to extend the asymmetric narrow bracketing model in an infinite horizon setting. We leave for future research the axiomatization of narrow bracketing with more than two sources of risk. In another ongoing project, we consider general models of correlation misperception by modeling the correlation structure among risk in difference sources as copula. Moreover, we maintain the independence axiom within each source and focus on narrow bracketing and correlation neglect as simplifying heuristics for inter-source risk. One extension is to deviate from expected utility within each source and incorporate other commonly used behavioral factors under one-source risk, like cumulative prospect theory and certainty effect.

References

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2018): "Multiattribute Utility Theory, Intertemporal Utility, and Correlation Aversion," *International Economic Review*, 59, 537–555.

Attema, A. E., O. L'haridon, and G. van de Kuilen (2019): "Measuring Multivariate Risk Preferences in the Health Domain," *Journal of Health Economics*, 64, 15–24.

Bansal, R. and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59, 1481–1509.

Barberis, N., M. Huang, and R. H. Thaler (2006): "Individual Preferences, Monetary

Gambles, and Stock Market Participation: A Case for Narrow Framing," *American Economic Review*, 96, 1069–1090.

Barro, R. J. (2009): "Rare Disasters, Asset Prices, and Welfare Costs," *American Economic Review*, 99, 243–64.

Bommier, A. (2007): "Risk Aversion, Intertemporal Elasticity of Substitution and Correlation Aversion," *Economics Bulletin*, 4, 1–8.

Bommier, A., A. Chassagnon, and F. Le Grand (2012): "Comparative Risk Aversion: A Formal Approach with Applications to Saving Behavior," *Journal of Economic Theory*, 147, 1614–1641.

Bommier, A., A. Kochov, and F. Le Grand (2017): "On Monotone Recursive Preferences," *Econometrica*, 85, 1433–1466.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012): "Salience theory of choice under risk," *The Quarterly journal of economics*, 127, 1243–1285.

Camara, M. (2021): "Computationally Tractable Choice," Working Paper.

Chew, S. H. and L. G. Epstein (1991): "Recursive Utility under Uncertainty," in *Equilibrium Theory in Infinite Dimensional Spaces*, Springer, 352–369.

Dillenberger, D., D. Gottlieb, and P. Ortoleva (2020): "Stochastic Impatience and the separation of Time and Risk Preferences," Working Paper.

Drèze, J. H. and F. Modigliani (1972): "Consumption Decisions under Uncertainty," *Journal of Economic Theory*, 5, 308–335.

Ebert, S. and G. van de Kuilen (2016): "Measuring Multivariate Risk Preferences," Working Paper.

Ellis, A. (2021): "Correlation Concern," Working Paper.

Ellis, A. and D. J. Freeman (2020): "Revealing Choice Bracketing," Working Paper.

Ellis, A. and M. Piccione (2017): "Correlation Misperception in Choice," *American Economic Review*, 107, 1264–92.

Enke, B. and F. Zimmermann (2019): "Correlation Neglect in Belief Formation," *The Review of Economic Studies*, 86, 313–332.

Epstein, L. and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–69.

Epstein, L. G., E. Farhi, and T. Strzalecki (2014): "How Much Would You Pay to Resolve Long-Run Risk?" *American Economic Review*, 104, 2680–97.

Epstein, L. G. and S. E. Zin (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99, 263–286.

Eyster, E. and G. Weizsäcker (2016): "Correlation Neglect in Portfolio Choice: Lab Evidence," Working Paper.

Fishburn, P. C. (1982): *The Foundations of Expected Utility.*, Theory & Decision Library.

Kallir, I. and D. Sonsino (2009): "The Neglect of Correlation in Allocation Decisions," *Southern Economic Journal*, 1045–1066.

Kihlstrom, R. E. and L. J. Mirman (1974): "Risk Aversion with Many Commodities," *Journal of Economic Theory*, 8, 361–388.

Kimball, M. and P. Weil (2009): "Precautionary Saving and Consumption Smoothing across Time and Possibilities," *Journal of Money, Credit and Banking*, 41, 245–284.

Kimball, M. S. (1990): "Precautionary Saving in the Small and in the Large," *Econometrica*, 58, 53–73.

Kreps, D. M. and E. L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 185–200.

Kubler, F., L. Selden, and X. Wei (2020a): "Incomplete Market Demand Tests for Kreps-Porteus-Selden Preferences," *Journal of Economic Theory*, 185, 104973.

——— (2020b): "Time Consistency, Temporal Resolution Indifference and the Separation of Time and Risk," Working Paper.

Lanier, J., B. Miao, J. K.-H. Quah, and S. Zhong (2020): "Intertemporal Consumption with Risk: A Revealed Preference Analysis," Working Paper.

Lanzani, G. (2021): "Correlation Made Simple: Applications to Salience and Regret Theory," *The Quarterly Journal of Economics*, qjab041.

Lian, C. (2020): "A Theory of Narrow Thinking," *The Review of Economic Studies.*

Meissner, T. and P. Pfeiffer (2022): "Measuring Preferences over the Temporal Resolution of Consumption Uncertainty," *Journal of Economic Theory*, 200, 105379.

Mu, X., L. Pomatto, P. Strack, and O. Tamuz (2021): "Background Risk and Small-Stakes Risk Aversion," Working Paper.

Rabin, M. (2000): "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68, 1281–1292.

Rabin, M. and G. Weizsäcker (2009): "Narrow Bracketing and Dominated Choices," *American Economic Review*, 99, 1508–43.

Read, D., G. Loewenstein, and M. Rabin (1999): "Choice Bracketing," *Journal of Risk and Uncertainty*, 19, 171–197.

Rees-Jones, A., R. Shorrer, and C. J. Tergiman (2020): "Correlation Neglect in Student-to-School Matching," Working Paper.

Selden, L. (1978): "A New Representation of Preferences over "Certain x Uncertain" Consumption Pairs: The "Ordinal Certainty Equivalent" Hypothesis," *Econometrica*, 1045–1060.

Selden, L. and I. Stux (1978): "Consumption Trees, OCE Utility and the Consumption/Savings Decision," Working Paper.

Thaler, R. (1985): "Mental Accounting and Consumer Choice," *Marketing Science*, 4, 199–214.

Tversky, A. and D. Kahneman (1981): "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453–458.

Vorjohann, P. (2021): "Reference-Dependent Choice Bracketing," Working Paper.

Weil, P. (1990): "Nonexpected Utility in Macroeconomics," *The Quarterly Journal of Economics*, 105, 29–42.

Appendix A: Asymmetric Narrow Bracketing in Infinite Horizon

In this section, we briefly discuss how to extend the asymmetric narrow bracketing model in (3) to one with multiple periods. Assume that the consumption space in each period $t = 1, ..., T$ is a compact interval $C \subset \mathbb{R}_+$, where T can be $+\infty$. The set of deterministic consumption paths is C^T with a generic element $\mathbf{c} = (c_t)_{t=1}^T$. For each consumption path $\mathbf{c} \in C^T$, we denote the subsequence of consumption in the first t periods as $\mathbf{c}^t = (c_\tau)_{\tau=1}^t$.

The preference is defined on the lottery space $\mathcal{P} = \mathcal{L}(C^T)$. Here we allow for lotteries with infinite supports to accommodate applications in finance. For each lottery P, denote $P_{[t]}$ as the marginal lottery in the first t periods, $1 \leq t < T$. For each subsequence of consumption $\mathbf{c}^t$ in the support of $P_{[t]}$, we define $\phi(P|\mathbf{c}^t)$ as the conditional lottery starting from period $t+1$, given that consumption in the first t periods are $\mathbf{c}^t$. When $T < +\infty$, $\phi(P|\mathbf{c}^t) \in \mathcal{L}(C^{T-t})$ and when $T = +\infty$, $\phi(P|\mathbf{c}^t) \in \mathcal{L}(C^\infty)$. Note that for each finite T, $\mathcal{L}(C^{T-t})$ is homeomorphic to a subset of $\mathcal{L}(C^\infty)$ where the consumption levels are always 0 from period $t+1$ on. So we will focus on the case with an infinite horizon.

The following notions are adapted from recursive preferences on temporal lotteries (Chew and Epstein, 1991, Bommier, Kochov, and Le Grand, 2017) to our framework. For each

$V : \mathcal{P} := \mathcal{L}(C^\infty) \to \mathbb{R}$ and $p \in \mathcal{P}$, denote

$$m_V(P)(B) \equiv P_1\Big\{c \in C : V(c, \phi(P|c) \in B)\Big\}, \forall B \in \mathcal{B}(V(\mathcal{P}))$$

where $V(\mathcal{P}) \subset \mathbb{R}$ is the image of V on $\mathcal{P}$ and $\mathcal{B}(V(\mathcal{P}))$ is the set of all Borel subsets of $V(\mathcal{P})$. Then $m_V(P)$ is a probability measure over utilities conditional on the current consumption. Now we define the *recursive preference over lotteries* as $V : \mathcal{P} \to \mathbb{R}$ with

$$\begin{aligned} V(P) &= I(m_V(P)), \\ V(c, q) &= W(c, V(q)), \end{aligned}$$

where $m_V(P)$ is defined as above, $I : \mathcal{L}(\mathbb{R}) \to \mathbb{R}$ is a *certainty equivalent*, that is, I is continuous, increasing with respect to first order stochastic dominance and $I(x) = x$ for each $x \in \mathbb{R}$, $W : C \times \mathbb{R} \to \mathbb{R}$ is continuous and strictly increasing in the second argument. Unlike the recursive preferences Chew and Epstein (1991) and Bommier, Kochov, and Le Grand (2017), V is defined on a different domain and can be discontinuous.

In order to get the CRRA-CES functional form, we can set $I = \phi^{-1} \circ \mathbb{E} \circ \phi$ with $\phi(x) = x^{\alpha/\rho}$ and $W(c, v) = (1-\delta)c^\rho + \delta v$, where $\rho < 1, 0 \neq \alpha < 1$ and $0 < \delta < 1$. The recursive preference is equivalent to the following recursion of value functions (up to a monotonic transformation):

$$V^{ANB}(P) = \frac{1}{\alpha}\mathbb{E}_{P_1}\Big(U_1^\alpha\Big) \tag{6}$$

$$U_t^\rho = (1-\delta)c_t^\rho + \delta\Big[\mathbb{E}_{\phi(P|\mathbf{c}^t)_1}\Big(U_{t+1}^\alpha\Big)\Big]^{\frac{\rho}{\alpha}} \tag{7}$$

where U_t is the value in period t and the expectation is computed with respect to $\phi(P|\mathbf{c}^t)_1$, which is the probability distribution of consumption levels in period $t+1$ conditional on the consumption path in the first t periods $\mathbf{c}^t$.

Then we explore the implications of asymmetric narrow bracketing in a standard asset pricing problem. The consumer is endowed with initial wealth $W_1 > 0$ in period 1 and chooses the consumption level and saving level in each period. Let S denote the finite state space in each period $t \geq 2$ and $\Omega = S^\infty$ denote the space of state sequences. The consumer has a prior belief over Ω. For each $s^\infty \in \Omega$, we denote by s^t the history of states from period 2 to period t for each $t \geq 2$. Let S^t be the set of all histories till period t.

The consumer's preference is represented by V^{ANB} in (6). She chooses a *consumption plan* $\{c_t\}_{t\geq 1}$ to maximizes her utility, where $c_1 \in X$ and $c_t : S^t \to X$ for each $t \geq 2$. Given the history of states s^t, the gross return on wealth from period $t-1$ to period t is $R_{w,t}(s^t) > 0$. Then the wealth dynamics are represented by the following equation:

$$W_{t+1}(s^{t+1}) = R_{w,t+1}(s^{t+1})(W_t(s^t) - c_t(s^t)).$$

We assume that the wealth level always lies in C. We say that a consumption plan $\{c_t\}_{t\geq 1}$ is feasible given initial wealth W_1 if if $c_t(s^t) \leq W_t(s^t)$ for all s^t, t. Each feasible consumption plan $\{c_t\}_{t\geq 1}$ induces a lottery $P \in \mathcal{P}$ where the consumption in the first period is deterministic. Using V^{ANB} in (6), we can define a utility function over feasible consumption plans as $\hat{V}^{ANB}(\{c_t\}_{t\geq 1})$. Then the optimization problem of the consumer is

$$J^{ANB}(W_1) = \sup\Big\{\hat{V}^{ANB}(\{c_t\}_{t\geq 1}) : \{c_t\}_{t\geq 1} \text{ is feasible given } W_1\Big\}.$$

To facilitate the comparison of ANB and EZ, we can similarly consider a consumer with a CRRA-CES EZ recursive utility function $\hat{V}^{EZ}$ over feasible consumption plans and the optimal value $J^{EZ}(W_1)$. We assume that RRA>1/EIS, then the EZ consumer has a preference over early resolution of uncertainty, while the ANB consumer exhibits indifference to temporal resolution of uncertainty. The following proposition shows that the two utility functions lead to the same optimal value.

Proposition 3. *Assume RRA>1/EIS, i.e., $\rho > \alpha$. Then for each $W_1 > 0$, $J^{ANB}(W_1) = J^{EZ}(W_1)$ exists and there exist consumption plans $\{c_t\}_{t\geq 1}$ and $\{c_t^n\}_{t\geq 1, n\geq 1}$ feasible given W_1 such that $c_1^n \to c_1, c_t^n(s^t) \to c_t(s^t)$ as n goes to infinity for each $t \geq 2$, and*

(i). $J^{ANB}(W_1) = \lim_{n\to\infty} \hat{V}^{ANB}(\{c_t^n\}_{t\geq 1}) = \hat{V}^{EZ}(\{c_t\}_{t\geq 1}) = J^{EZ}(W_1)$;

(ii). For each $n \geq 1, t \geq 2$, c_t^n is injective on Ω_t, that is, $c_t^n(s^t) \neq c_t^n(\hat{s}^t)$ if $s^t \neq \hat{s}^t$.

We sketch the proof of Proposition 3. Recall that in Section 6.2.1, we observe that, in the setup with two periods, V^{ANB} agrees with V^{EZ} when risk does not resolve early. This observation remains valid with an infinite horizon. On the one hand, since the EZ consumer has a preference for early resolution of risk, we know that $\hat{V}^{ANB}(\{c_t\}_{t\geq 1}) \leq \hat{V}^{EZ}(\{c_t\}_{t\geq 1})$ for each feasible consumption plan $\{c_t\}_{t\geq 1}$ and $J^{ANB}(W_1) \leq J^{EZ}(W_1)$. On the other hand, for a consumption plan $\{c_t\}_{t\geq 1}$ where c_t is injective on Ω_t for each t, the consumption history contains the same information as state history and there is no early resolution of risk. In this case, $\hat{V}^{ANB}(\{c_t\}_{t\geq 1}) = \hat{V}^{EZ}(\{c_t\}_{t\geq 1})$. By continuity of V^{EZ} and hence $\hat{V}^{EZ}$, we can find a sequence of consumption plans $\{c_t^n\}_{t\geq 1, n\geq 1}$ with injective consumption functions to approximate the optimal value of the EZ consumer. This implies $J^{ANB}(W_1) = J^{EZ}(W_1)$.

A directly corollary of Proposition 3 is that our model of asymmetric narrow bracketing has the same implications in asset pricing as the EZ model when RRA>1/EIS, which is the common parametric assumption in most applications of EZ preferences in finance and macroeconomics. Following the asset pricing literature, let $\gamma := 1-\alpha$ be the RRA, $\psi := \frac{1}{1-\rho}$ be the EIS and $\theta := \frac{\alpha}{\rho}$.

Corollary 3. *Assume RRA>1/EIS, i.e., $\rho > \alpha$. Denote $\mathbb{E}_t$ as the expectation with respect to s^t. Then for each $W_1 > 0$,*

(i). Euler equation:

$$\lim_{n\to\infty} \mathbb{E}_t\Big[\delta^\theta \big(\frac{c_{t+1}^n}{c_t^n}\big)^{-\frac{\theta}{\psi}} R_{w,t+1}^\theta\Big] = \mathbb{E}_t\Big[\delta^\theta \big(\frac{c_{t+1}}{c_t}\big)^{-\frac{\theta}{\psi}} R_{w,t+1}^\theta\Big] = 1;$$

(ii). Asset pricing formula: for each asset i with gross return $R_{i,t}$,

$$\lim_{n\to\infty}\mathbb{E}_t\Big[\delta^\theta(\frac{c^n_{t+1}}{c^n_t})^{-\frac{\theta}{\psi}}R_{w,t+1}^{-(1-\theta)}R_{i,t+1}\Big]=\mathbb{E}_t\Big[\delta^\theta(\frac{c_{t+1}}{c_t})^{-\frac{\theta}{\psi}}R_{w,t+1}^{-(1-\theta)}R_{i,t+1}\Big]=1;$$

We end this section with a brief discussion about the case with RRA<1/EIS, i.e., when the EZ consumer prefers late solution of risk. In this case, $J^{ANB}(W_1)$ might be strictly higher than $J^{EZ}(W_1)$ due to the discontinuity of V^{ANB} on $\mathcal{P}$, which invalidates our proof of Proposition 3. Interestingly, the ANB consumer might have excessive demand for consumption smoothing across different states of the world in the same period. However, the derivation of the optimal value and Euler equation is much less tractable and we leave a formal treatment of this case for future study.

Appendix B: Proofs

Proof of Theorem 1. Necessity. V^{CN}, V^{ANB-CN} and V^{FNB} trivially satisfy Weak Order and Correlation Neutrality. Given Correlation Neutrality, Continuity is equivalent to Continuity 3, which is guaranteed by continuity of w, v_1 and v_2 in the three utility functions. Similarly, Monotonicity of $\succsim$ is guaranteed by monotonicity of w, v_1 and v_2.

Now we verify Independence*. If $\succsim$ is represented by V^{CN}, then for any $P,Q,R,S\in\hat{\mathcal{P}}$ with $P_i=R_i,Q_j=S_j$ for some $i,j\in\{1,2\}$ and $\alpha\in(0,1)$,

$$P\succsim Q,R\sim S\Longleftrightarrow\alpha P+(1-\alpha)R\succsim\alpha Q+(1-\alpha)R.$$

This implies that $\succsim$ satisfies Independence*.

If $\succsim$ is represented by V^{FNB}, then $\succsim$ satisfies part (i) of Independence* since $\succsim_{1|p}$ is represented by EU v_1 and $\succsim_{2|q}$ is represented by EU v_2 for any $p\in\mathcal{L}^0(X_2),q\in\mathcal{L}^0(X_1)$. For part (ii) and (iii), $P_i=R_i,P_{-i}\sim_{-i|x}R_{-i}$ for some $i\in\{1,2\}$ and $x\in X_i$ implies that $P\sim R\sim\alpha P+(1-\alpha)R$ for any $\alpha\in(0,1)$. Similarly, $Q\sim S\sim\alpha Q+(1-\alpha)S$ for any $\alpha\in(0,1)$. Hence $P\succ Q$ suggests that $\alpha P+(1-\alpha)R\succ\alpha Q+(1-\alpha)S$.

If $\succsim$ is represented by V^{ANB}, then $\succsim_{2|q}$ is represented by EU v_2 for any $q\in\mathcal{L}^0(X_1)$ and $P_1=R_1,P_2\sim_{2|x}R_2$ for some $x\in X_1$ implies that $P\sim R$. Hence it suffices to show that $\succsim$ satisfies part (i) and part (ii) of Independence* when $P_2=R_2,Q_2=S_2$, which is guaranteed since for any $P,Q,R,S\in\hat{\mathcal{P}}$ with $P_2=R_2,Q_2=S_2$ and $\alpha\in(0,1)$,

$$P\succsim Q,R\sim S\Longleftrightarrow\alpha P+(1-\alpha)R\succsim\alpha Q+(1-\alpha)R.$$

Sufficiency. Throughout the proof, for any set of lotteries $\mathcal{A}\subseteq\mathcal{P}$, denote $\max_{\succsim}\mathcal{A}=\{P\in\mathcal{A}:P\succsim Q,\forall Q\in\mathcal{A}\}$ whenever it is well-defined. That is, $\max_{\succsim}\mathcal{A}$ is the set of most preferred lotteries in $\mathcal{A}$ under $\succsim$. Finally, for any set $A\subseteq\mathbb{R}$, we denote A^o as its interior.

Assume all axioms hold. First, by Correlation Neutrality, $P\sim(P_1,P_2)$ for all $P\in\mathcal{P}$

and it suffices to focus on the restriction of $\succsim$ on product lotteries $\hat{\mathcal{P}}$. Then Continuity implies that $\succsim$ satisfies strong continuity, that is, for all $Q \in \mathcal{P}$, the sets $\{P \in \mathcal{P} : P \succ Q\}$ and $\{P \in \mathcal{P} : Q \succ P\}$ are open subsets of $\mathcal{P}$. The following lemma guarantees that the conditional preference $\succsim_{i|q}$ admits an EU representation for each $i = 1, 2$ and $q \in \mathcal{L}^0(X_{-i})$.

Lemma 3. *For each $i = 1, 2$ and $q \in \mathcal{L}^0(X_{-i})$, $\succsim_{i|q}$ admits an EU representation with index $v_{i|q}$, which is continuous and unique up to a positive affine transformation. Moreover, if $q \in X_{-i}$, then $v_{i|q}$ can be chosen to be strictly monotonic.*

Proof of Lemma 3 . Fix $i = 1, 2$ and $q \in \mathcal{L}^0(X_{-i})$. By Continuity and Independence*, $\succsim_{i|q}$ admits an EU representation with a continuous utility index $v_{i|q}$ defined on X_i, which is unique up to a positive affine transformation. Moreover, if $q \in X_{-i}$, that is, $q = y$ for some $y \in X_{-i}$, then by Monotonicity, $v_{i|q}$ must be strictly monotonic. □

A direct corollary of Lemma 3 guarantees the existence of "certainty equivalents".

Corollary 4. *For each $P \in \hat{\mathcal{P}}$, there exists $x_1, y_1 \in X_1, x_2, y_2 \in X_2$ such that $P \sim (P_1, x_2) \sim (x_1, P_2) \sim (y_1, y_2)$.*

Proof of Corollary 4 . We will prove the result for $P_1 \notin X_1, P_2 \notin X_2$. The proof for other cases is easier. By Lemma 3, we can find $x_2, a, a' \in X_2$ such that either $\sum_x v_{2|P_1}(x)P_2(x) = v_{2|P_1}(x_2)$ and hence $P \sim (P_1, x_2)$ or $v_{2|P_1}(a) > \sum_x v_{2|P_1}(x)P_2(x) > v_{2|P_1}(a')$. In the latter case, since $v_{2|P_1}$ is continuous and X_2 is a closed interval, there exists $x_2 \in X_2$ where $v_{2|P_1}(x_2) = \sum_x v_{2|P_1}(x)P_2(x)$, which implies $P \sim (P_1, x_2)$. Similarly, we can find $x_1 \in X_1$ with $P \sim (x_1, P_2)$. Now let $y_2 = x_2$. Repeat the above arguments for product lottery (P_1, x_2) and we can find $y_1 \in X_1$ with $(y_1, y_2) \sim (P_1, x_2) \sim P$. □

The next lemma shows that under Continuity, Independence* can be strengthened.

Lemma 4. *(i). For each $P, Q, R \in \hat{\mathcal{P}}$ with $P \succsim R \succsim Q$, $P \succ Q$ and $P_i = Q_i$ for some $i \in \{1, 2\}$, then there exists a unique $\lambda \in [0, 1]$ such that $R \sim \lambda P + (1 - \lambda)Q$.*
(ii). For $P, Q, R, S \in \hat{\mathcal{P}}$, $\alpha \in (0, 1)$, $i, j \in \{1, 2\}$ and $x \in X_i, y \in X_j$, if $P_i = R_i, Q_j = S_j, P_{-i} \sim_{-i|x} R_{-i}$ and $Q_{-j} \sim_{-j|y} S_{-j}$, then

$$P \sim Q, R \sim S \Longrightarrow \alpha P + (1 - \alpha)R \sim \alpha Q + (1 - \alpha)S$$
$$P \succ Q, R \succ S \Longrightarrow \alpha P + (1 - \alpha)R \succ \alpha Q + (1 - \alpha)S$$

(iii). For $P, Q \in \hat{\mathcal{P}}$ and $x \in X_1$, if $R = (P_1, Q_2)$ or (Q_1, P_2) and $P_2 \sim_{2|x} Q_2$, then

$$P \sim Q \Longrightarrow \alpha P + (1 - \alpha)R \sim \alpha Q + (1 - \alpha)R.$$

Proof of Lemma 4. (i). Denote $I = \{\eta \in [0, 1] : R \succ \eta P + (1 - \eta)Q\}$ and $\lambda = \sup I$. λ is well-defined as I is bounded. We claim that $\lambda P + (1 - \lambda)Q \sim R$. If $\lambda P + (1 - \lambda)Q \succ R$, then $\lambda > 0$ and by Continuity, there exists $\varepsilon > 0$ with $(\lambda - \varepsilon)P + (1 - \lambda + \varepsilon)Q \succ R$. This implies $\lambda - \varepsilon \notin I$. Since $P \succ Q$ and $P_i = Q_i$ for some i, by Independence*, for any $\alpha, \beta \in [0, 1]$, $\alpha P + (1 - \alpha)Q \succ \beta P + (1 - \beta)Q$ if any only if $\alpha > \beta$, which implies $[\lambda - \varepsilon, \lambda] \cap I = \emptyset$ and leads to a contradiction with $\lambda = \sup I$. If instead $R \succ \lambda P + (1-\lambda)Q$, then there exists $\varepsilon > 0$ with $R \succ (\lambda+\varepsilon)P + (1-\lambda-\varepsilon)Q$ and hence $\lambda + \varepsilon \in I$, which again contradicts with the definition of λ. Uniqueness of λ is guaranteed by its definition.

(ii). Suppose $P \sim Q, R \sim S$. If $P \sim R$, then the the result trivially holds as $\alpha P + (1-\alpha)R \sim R \sim Q \sim \alpha Q + (1-\alpha)S$. Without loss of generality, assume $P \succ R$. Then $Q \succ S$ and $P \succ \alpha P + (1-\alpha)R \succ R$, $Q \succ \alpha Q + (1-\alpha)S \succ S \sim R$ for any $\alpha \in (0,1)$. Suppose by contradiction that $\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$. By part (i), there exists a unique $\lambda \in (0,1)$ with $\alpha Q + (1-\alpha)S \sim \lambda(\alpha P + (1-\alpha)R) + (1-\lambda)R = \alpha\lambda P + (1-\alpha\lambda)R$. Notice that $Q \sim P \succ \lambda P + (1-\lambda)R$, $S \sim R$, $Q_j = S_j$, $Q_{-j} \sim_{-j|y} S_{-j}$, $(\lambda P + (1-\lambda)R)_i = P_i = R_i$ and $(\lambda P + (1-\lambda)R)_{-i} \sim_{-i|x} R_{-i}$. The last property follows from Lemma 3. Then Independence* implies that

$$\alpha Q + (1-\alpha)S \succ \alpha(\lambda P + (1-\lambda)R) + (1-\alpha)R = \alpha\lambda P + (1-\alpha\lambda)R,$$

which leads to a contradiction. The case for $\alpha P + (1-\alpha)R \prec \alpha Q + (1-\alpha)S$ is symmetric.

Now assume $P \succ Q, R \succ S$. If $P \sim R$, then the result holds as $\alpha P + (1-\alpha)R \sim P \succ \max_{\succsim}\{Q,S\} \succsim \alpha Q + (1-\alpha)S$ for all $\alpha \in (0,1)$. Without loss of generality, suppose $P \succ R$. If $R \succsim Q$, then $\alpha P + (1-\alpha)R \succ R \succsim \max_{\succsim}\{Q,S\} \succsim \alpha Q + (1-\alpha)S$ for all $\alpha \in (0,1)$. If instead $Q \succ R$, then $P \succ Q \succ R \succ S$. By part (i), we can find $\lambda \in (0,1)$ such that $R \sim \lambda Q + (1-\lambda)S := S' \succ S$. Since $S'_j = S_j = Q_j$ and $S'_{-j} = \lambda Q_{-j} + (1-\lambda)S_{-j} \sim_{-j} Q_{-j}$, Independence* implies that any $\alpha \in (0,1)$,

$$\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S' \succ \alpha Q + (1-\alpha)S.$$

(iii). The proof is the same as the first part of the proof for (ii) and omitted. □

A tuple $(P,Q,R,S) \in \hat{\mathcal{P}}^4$ is called *proper* if $P_i = R_i$, $Q_j = S_j$ for some $i,j \in \{1,2\}$ and $P \succsim R, Q \succsim S$. A proper tuple (P,Q,R,S) satisfies the *independence property* if one of the following conditions holds:

- $P \succ Q, R \sim S$ and for all $\alpha \in (0,1)$, $\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$;
- $P \sim Q, R \succ S$ and for all $\alpha \in (0,1)$, $\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$;
- $P \sim Q, R \sim S$ and for all $\alpha \in (0,1)$, $\alpha P + (1-\alpha)R \sim \alpha Q + (1-\alpha)S$;
- $P \succ Q, R \succ S$ and for all $\alpha \in (0,1)$, $\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$.

Now we introduce additional notations that will be used later on. For each $x_i \in X_i$, $i = 1,2$, denote $\Pi^i(x_i|x_{-i})$ as the set of marginal lotteries in source i with certainty equivalent x_i given background outcome x_{-i} in source $-i$. Formally, $\Pi^i(x_i|x_{-i}) = \{p \in \mathcal{L}^0(X_i) : p \sim_{i|x_{-i}} x_i\}$. For any two product lotteries $P,Q \in \hat{\mathcal{P}}$ with $P \succsim Q$, let $[Q,P]$ denote the set of all product lotteries whose utilities lie between P and Q, that is, $[Q,P] = \{S \in \hat{\mathcal{P}} : P \succsim S \succsim Q\}$.

Then, for each $q_i \in \mathcal{L}^0(X_i)$ and $x_i \in X_i$, $i = 1,2$, we define

$$\Gamma_{1,q_1}(x_2|x_1) = \bigcup_{\substack{P,Q \in \{q_1\}\times\Pi^2(x_2|x_1),\\ P \succsim Q}} [Q,P] \ , \quad \Gamma_{2,q_2}(x_1|x_2) = \bigcup_{\substack{P,Q \in \Pi^1(x_1|x_2)\times\{q_2\},\\ P \succsim Q}} [Q,P].$$

Intuitively, $\Gamma_{1,q_1}(x_2|x_1)$ includes all product lotteries whose utilities are bounded by lotteries in $\{q_1\} \times \Pi^2(x_2|x_1)$. The following lemma states that any lottery $\Gamma_{1,q_1}(x_2|x_1)$ is indifferent to some lottery in $\{q_1\} \times \Pi^2(x_2|x_1)$.

Lemma 5. *Fix $q_1 \in \mathcal{L}^0(X_1), q_2 \in \mathcal{L}^0(X_2)$ and $x_1 \in X_1, x_2 \in X_2$.*
(i). For each $P \in \Gamma_{1,q_1}(x_2|x_1)$, there exists $P' \in \{q_1\} \times \Pi^2(x_2|x_1)$ with $P' \sim P$;
(ii). For each $P \in \Gamma_{2,q_2}(x_1|x_2)$, there exists $P' \in \Pi^1(x_1|x_2) \times \{q_2\}$ with $P' \sim P$.

Proof of Lemma 5. (i). By definition, there exist $Q, Q' \in \{q_1\}\times\Pi^2(x_2|x_1)$ with $Q \succsim P \succsim Q'$. As $Q_1 = Q'_1 = q_1$, by Lemma 4, there exists $\lambda \in [0,1]$ with $P \sim \lambda Q + (1-\lambda)Q'$. By Lemma 3, $(\lambda Q + (1-\lambda)Q')_2 \sim_{2|x_1} Q_2 \sim_{2|x_1} Q'_2$. Denote $P' = \lambda Q + (1-\lambda)Q'$. Hence $P' \sim P$, $P' \in \{q_1\} \times \Pi^2(x_2|x_1)$. The proof for (ii) is similar. □

Case 1. There exists $(z_1, z_2) \in X_1 \times X_2$ such that $(p,q) \sim (x,y)$ for all $(x,y) \in X_1 \times X_2$ and $(p,q) \in \Pi^1(x|z_2) \times \Pi^2(y|z_1)$. The following lemma shows that $\succsim$ must admit a FNB representation.

Lemma 6. *In Case 1, $\succsim$ is represented by V^{FNB}.*

Proof of Lemma 6. By assumption, the conditional preferences $\succsim_{1|p}$ and $\succsim_{2|q}$ are independent of $p \in \mathcal{L}^0(X_2)$ and $q \in \mathcal{L}^0(X_1)$. We denote by $\succsim_1$ and $\succsim_2$ the two conditional preferences respectively. By Lemma 3, for $i = 1,2$, we can find a continuous and strictly monotonic v_i as the EU index of $\succsim_i$. Since X_i is a closed interval, the certainty equivalent function CE_{v_i} is well-defined. For any $(p,q) \in \hat{\mathcal{P}}$, we know $(p,q) \sim (CE_{v_1}(p), CE_{v_2}(q))$. We define $\hat{\succsim}$ as the restriction of $\succsim$ on $X_1 \times X_2$. By Continuity, $\hat{\succsim}$ is continuous. Then Debreu's Theorem implies that $\hat{\succsim}$ is represented a continuous utility function w. Monotonicity guarantees that w is strictly monotonic. Therefore w, v_1 and v_2 are continuous and strictly monotonic and for all $P, Q \in \mathcal{P}$,

$$\begin{aligned} P \succsim Q &\iff (CE_{v_1}(P_1), CE_{v_2}(P_2)) \hat{\succsim} (CE_{v_1}(Q_1), CE_{v_2}(Q_2)) \\ &\iff w(CE_{v_1}(P_1), CE_{v_2}(P_2)) \geq w(CE_{v_1}(Q_1), CE_{v_2}(Q_2)) \\ &\iff V^{FNB}(P) \geq V^{FNB}(Q). \end{aligned}$$

□

Case 2: There exists $(z_1, z_2) \in X_1 \times X_2$ such that $(p,q) \sim (p,q')$ for all $(x,y) \in X_1 \times X_2$, $p \in \Pi^1(x|z_2)$ and $q, q' \in \Pi^2(y|z_1)$, and we can find $(x_0, y_0) \in X_1 \times X_2$, $p_0, p'_0 \in \Pi^1(x_0|z_2)$ and $q_0 \in \Pi^2(y_0|z_1)$ with $(p_0, q_0) \succ (p'_0, q_0)$.

By assumption, $\succsim_{2|q}$ is independent of $q \in \mathcal{L}^0(X_1)$ and we can denote it as $\succsim_2$. By Lemma 3, $\succsim_2$ admits a continuous and strictly monotonic EU index v_2. Then for any $(p,q),(p',q') \in \hat{\mathcal{P}}$ with $q \sim_2 q'$, $(p,q) \succsim (p',q')$ if and only if $(p, CE_{v_2}(q)) \succsim (p', CE_{v_2}(q))$. Hence we can focus on the restriction of $\succsim$ on $\mathcal{L}^0(X_1) \times X_2$.

Lemma 7. *In Case 2, there exists $z_0, z_0' \in X_2$ such that for any $x \in X_1^o$, we can find $p_x, p_x' \in \Pi^1(x|z_0')$ with $(p_x, z_0) \succ (p_x', z_0)$ and $q_x, q_x' \in \Pi^1(x|z_0)$ with $(q_x, z_0') \succ (q_x', z_0')$.*

Proof of Lemma 7. By assumption, there exist $z_0, z_0' \in X_2$, $x_0 \in X_1$ and $p_0, p_0' \in \Pi^1(x_0|z_0')$ such that $(p_0, z_0) \succ (p_0', z_0)$. Suppose by contradiction that there exists $x_1 \in X_1^o$ with $(p_1, z_0) \sim (p_1', z_0)$ for all $p_1 \sim_{1|z_0'} p_1' \sim_{1|z_0'} x_1$. By Lemma 3, $\succsim_{1|z_0}$ and $\succsim_{1|z_0'}$ are represented by EU $v_{1|z_0}$ and $v_{1|z_0'}$ respectively. Since $x_1 \in X_1^o$, there exist $\overline{x}, \underline{x} \in X_i^o$ with $\overline{x} > x_1 > \underline{x}$. As $v_{1|z_0}$ and $v_{1|z_0'}$ are unique up to a positive affine transformation, we can set $v_{1|z}(\underline{x}) = v_{1|z'}(\underline{x})$ and $v_{1|z_0}(x_1) = v_{1|z_0'}(x_1)$. For any $x \in X_1$ with $x > x_1$, there exists $\alpha \in (0,1)$ with $\alpha\delta_x + (1-\alpha)\delta_{\underline{x}} \sim_{1|z_0'} x_1$. Then $\alpha v_{1|z_0}(x) + (1-\alpha) v_{1|z_0}(\underline{x}) = v_{1|z_0}(x_1) = v_{1|z_0'}(x_1) = \alpha v_{1|z_0'}(x) + (1-\alpha) v_{1|z_0'}(\underline{x})$. Since $v_{1|z_0}(\underline{x}) = v_{1|z_0'}(\underline{x})$ and $\alpha \in (0,1)$, $v_{1|z_0}(x) = v_{1|z_0'}(x)$ for all $x > x_1$. Specifically, we have $v_{1|z_0}(\overline{x}) = v_{1|z_0'}(\overline{x})$. Repeat the above arguments and we can show $v_{1|z_0}(x) = v_{1|z_0'}(x)$ for all $x < x_1$. Thus $v_{1|z_0} \equiv v_{1|z_0'}$, contradicting with $(p_0, z_0) \succ (p_0', z_0)$ and $(p_0, z_0') \sim (p_0', z_0')$. Hence for any $x \in X_1^o$, we can find $p_x, p_x' \in \Pi^1(x|z_0')$ with $(p_x, z_0) \succ (p_x', z_0)$. The existence of q_x, q_x' can be proved similarly. □

For each $z \in X_2$, denote

$$\Sigma^2(z) := \{y \in X_2 : \exists\ x \in X_1 \text{ and } p, p' \in \Pi^1(x|z) \text{ s.t. } (p,y) \succ (p',y)\}.$$

To interpret $\Sigma^2(z)$, note that $y \in \Sigma^2(z)$ if and only if there exists $q \sim_{2|z} y$ such that $\succsim_{1|q}$ is not identical to $\succsim_{1|z}$. The following lemma follows.

Lemma 8. *In Case 2, (i) $\Sigma^2(z)$ is open in the relative topology of X_2 for each $z \in X_2$; (ii) there exist $z_0, z_0' \in X_2$ such that $\Sigma^2(z_0) \cup \Sigma^2(z_0') = X_2$ and $\Sigma^2(z_0) \cap \Sigma^2(z_0') \neq \emptyset$.*

Proof of Lemma 8. (i). For $z \in X_2$, if $\Sigma^2(z) \neq \emptyset$, then $\Sigma^2(z)$ is open in the relative topology of X_2 due to Continuity.

(ii). By Lemma 7, there exist $z_0, z_0' \in X_2$ with $\succsim_{1|z_0} \neq \succsim_{1|z_0'}$. This implies $z_0 \in \Sigma^2(z_0')$ and $z_0' \in \Sigma^2(z_0)$. For any $z \in X_2$, either $\succsim_{1|z} \neq \succsim_{1|z_0}$ or $\succsim_{1|z} \neq \succsim_{1|z_0'}$, which suggests either $z \in \Sigma^2(z_0)$ or $z \in \Sigma^2(z_0')$. Hence $\Sigma^2(z_0) \cup \Sigma^2(z_0') = X_2$. Since X_2 is a closed interval on $\mathbb{R}$, X_2 is connected and cannot be represented as the union of two disjoint non-empty open subsets. As a result, $\Sigma^2(z_0) \cap \Sigma^2(z_0') \neq \emptyset$. □

Using the same argument in Lemma 7, we have the following corollary.

Corollary 5. *In Case 2, if $z' \in \Sigma^2(z)$ where $z, z' \in X_2$, then for any $x \in X_1^o$, there exist $p_x, p_x' \in \Pi^1(x|z)$ with $(p_x, z') \succ (p_x', z')$.*

Our goal is to prove that a proper tuple of product lotteries (P,Q,R,S) satisfies the independence property whenever $P_2 = R_2$ and $Q_2 = S_2$. The following lemma assures that we can focus on the case where $P \sim Q$, $R \sim S$.

Lemma 9. *Suppose that $(P,Q,R,S) \in \hat{\mathcal{P}}^4$ is a proper tuple with $P_i = R_i, Q_j = S_j$ for some $i,j \in \{1,2\}$. If the independence property holds for any such (P,Q,R,S) with $P \sim Q, R \sim S$, then the independence property holds for any such (P,Q,R,S) with $P \succsim Q, R \succsim S$.*

Proof of Lemma 9. Like Lemma 4, it suffices to consider the case where $P \succsim Q \succ R \succsim S$. By Lemma 3, there exist $\alpha \in (0,1]$ and $\beta \in [0,1)$ such that $P' = \alpha P + (1-\alpha)R \sim Q$ and $S' = \beta Q + (1-\beta)S \sim R$. Then the independence property holds for (P',Q,R,S'), that is, for any $\lambda \in (0,1)$, $\lambda P' + (1-\lambda)R \sim \lambda Q + (1-\lambda)S'$. By Lemma 3 and $P \succsim P'$, $S' \succsim S$, we have

$$\lambda P + (1-\lambda)R \succsim \lambda P' + (1-\lambda)R \sim \lambda Q + (1-\lambda)S' \succsim \lambda Q + (1-\lambda)S.$$

At least one of the above weak preference rankings would be strict if $P \succ Q$ or $R \succ S$. $\square$

The next lemma shows that if the independence property holds on two sets of product lotteries respectively, then it also holds on their union.

Lemma 10. *In Case 2, $(P,Q,R,S) \in (\mathcal{L}^0(X_1) \times X_2)^4$ is a proper tuple where $P \sim Q, R \sim S$, $P_2 = R_2 = y_1, Q_2 = S_2 = y_2$ with $y_1, y_2 \in X_2$. Fix any $T^i \in \hat{\mathcal{P}}$ for $i = 1, ..., 4$ with $T^4 \succ T^2 \succ T^3 \succ T^1$. If the independence property holds for any such (P,Q,R,S) with $\{P,Q,R,S\} \subseteq [T^1,T^2]$ or $\{P,Q,R,S\} \subseteq [T^3,T^4]$, then it also holds for any such (P,Q,R,S) with $\{P,Q,R,S\} \subseteq [T^1,T^4]$.*

Proof of Lemma 10. Without loss of generality, we can assume $P \succ R$ and $T^1 \succsim (x, y_i), i = 1,2$ for some $x \in X_1$, otherwise either the lemma is trivial or we can redefine T^1 without any loss. Moreover, it suffices to focus on the case where $P \sim Q \succ T^2$ and $R \sim S \prec T^3$. Fix any W_1, W_2 with $T^2 \succ W_2 \succ W_1 \succ T^3$. Then we have

$$T^4 \succsim P \sim Q \succ T^2 \succ W_2 \succ W_1 \succ T^3 \succ R \sim S \succsim T^1.$$

By Lemma 3, we can find $\hat{P}, \hat{Q}, \hat{R}, \hat{S}$ such that $\hat{P} \sim \hat{Q} \sim W_2$, $\hat{R} \sim \hat{S} \sim W_1$ and $\hat{P}_2 = \hat{R}_2 = P_2 = y_1$, $\hat{Q}_2 = \hat{S}_2 = Q_2 = y_2$. Notice that $P, Q, \hat{R}, \hat{S} \in [T^3, T^4]$, where the independence property holds. Then there exists $\lambda \in (0,1)$ such that $\lambda P + (1-\lambda)\hat{R} \sim \hat{P} \sim \lambda Q + (1-\lambda)\hat{S}$. Similarly, we can find $\lambda' \in (0,1)$ with $\lambda' \hat{P} + (1-\lambda')R \sim \hat{R} \sim \lambda' \hat{Q} + (1-\lambda')S$.

Actually in the construction of $\hat{R}$ and $\hat{S}$, there exist $\eta_1, \eta_2 \in (0,1)$ with

$$\eta_1 P + (1-\eta_1)R \sim \hat{R} \sim \hat{S} \sim \eta_2 Q + (1-\eta_2)S.$$

We claim that $\eta_1 = \eta_2$. To see this, as $P_2 = R_2 = \hat{P}_2 = \hat{R}_2$, we know

$$\lambda' \hat{P} + (1-\lambda')R \sim \lambda\lambda' P + (1-\lambda)\lambda' \hat{R} + (1-\lambda')R \sim \hat{R}$$

which implies

$$\hat{R} \sim \frac{\lambda\lambda'}{\lambda\lambda' + (1-\lambda')}P + \frac{1-\lambda'}{\lambda\lambda' + (1-\lambda')}R$$

and hence $\eta_1 = \frac{\lambda\lambda'}{\lambda\lambda'+(1-\lambda')}$. Similarly we can show that $\eta_2 = \frac{\lambda\lambda'}{\lambda\lambda'+(1-\lambda')} = \eta_1 := \eta^{w_1}$.

A symmetric argument shows that there exists η^{w_2} with $\eta^{w_1} < \eta^{w_2} < 1$ and

$$\eta^{w_2}P + (1-\eta^{w_2})R \sim \hat{P} \sim \hat{Q} \sim \eta^{w_2}Q + (1-\eta^{w_2})S.$$

Now we consider η with $\eta^{w_1} < \eta < \eta^{w_2}$. Notice that

$$\begin{aligned}\eta P + (1-\eta)R &= \frac{\eta - \eta^{w_1}}{\eta^{w_2} - \eta^{w_1}}[\eta^{w_2}P + (1-\eta^{w_2})R] + \frac{\eta^{w_2} - \eta}{\eta^{w_2} - \eta^{w_1}}[\eta^{w_1}P + (1-\eta^{w_1})R] \\ &\sim \frac{\eta - \eta^{w_1}}{\eta^{w_2} - \eta^{w_1}}\hat{P} + \frac{\eta^{w_2} - \eta}{\eta^{w_2} - \eta^{w_1}}\hat{R}.\end{aligned}$$

Similarly,

$$\eta Q + (1-\eta)S \sim \frac{\eta - \eta^{w_1}}{\eta^{w_2} - \eta^{w_1}}\hat{Q} + \frac{\eta^{w_2} - \eta}{\eta^{w_2} - \eta^{w_1}}\hat{S}.$$

As $\hat{P} \sim \hat{Q}, \hat{R} \sim \hat{S} \in [T^3, T^4]$ and $\hat{P}_2 = \hat{R}_2$, $\hat{Q}_2 = \hat{S}_2$, the independence property holds for $(\hat{P}, \hat{Q}, \hat{R}, \hat{S})$ and hence

$$\eta P + (1-\eta)R \sim \frac{\eta - \eta^{w_1}}{\eta^{w_2} - \eta^{w_1}}\hat{P} + \frac{\eta^{w_2} - \eta}{\eta^{w_2} - \eta^{w_1}}\hat{R} \sim \frac{\eta - \eta^{w_1}}{\eta^{w_2} - \eta^{w_1}}\hat{Q} + \frac{\eta^{w_2} - \eta}{\eta^{w_2} - \eta^{w_1}}\hat{S} \sim \eta Q + (1-\eta)S.$$

Then we check the independence property for $\eta^{w_2} < \eta < 1$.

$$\begin{aligned}\hat{P} \sim \eta^{w_2}P + (1-\eta^{w_2})R &= \frac{\eta^{w_2} - \eta^{w_1}}{\eta - \eta^{w_1}}[\eta P + (1-\eta)R] + \frac{\eta - \eta^{w_2}}{\eta - \eta^{w_1}}[\eta^{w_1}P + (1-\eta^{w_1})R] \\ &\sim \frac{\eta^{w_2} - \eta^{w_1}}{\eta - \eta^{w_1}}[\eta P + (1-\eta)R] + \frac{\eta - \eta^{w_2}}{\eta - \eta^{w_1}}\hat{R}.\end{aligned}$$

Similarly,

$$\hat{Q} \sim \eta^{w_2}Q + (1-\eta^{w_2})S \sim \frac{\eta^{w_2} - \eta^{w_1}}{\eta - \eta^{w_1}}[\eta Q + (1-\eta)S] + \frac{\eta - \eta^{w_2}}{\eta - \eta^{w_1}}\hat{S}.$$

Note that $\eta P + (1-\eta)R, \eta Q + (1-\eta)S, \hat{R}, \hat{S} \in [T^3, T^4]$, and $(\eta P + (1-\eta)R)_2 = \hat{R}_2, (\eta Q + (1-\eta)S)_2 = \hat{S}_2$. By assumption and Lemma 9, the independence property holds for $(\eta P + (1-\eta)R, \eta Q + (1-\eta)S, \hat{R}, \hat{S})$. Whenever $\eta P + (1-\eta)R \not\sim \eta Q + (1-\eta)S$, we know $\hat{P} \not\sim \hat{Q}$, leading to a contradiction. Thus, $\eta P + (1-\eta)R \sim \eta Q + (1-\eta)S$.

The proof for the case with $\eta^{w_1} > \eta > 0$ is symmetric. Hence for all $\eta \in (0,1)$, $\eta P + (1-\eta)R \sim \eta Q + (1-\eta)S$. □

Lemma 11 shows the (P, Q, R, S) with $P_2 = R_2$ and $Q_2 = S_2$ satisfies the independence property if the utilities of P and R are "sufficiently close".

Lemma 11. *In Case 2, suppose $z = z_0$ or z_0'. Then a proper tuple $(P, Q, R, S) \in (\mathcal{L}^0(X_1) \times X_2)^4$ satisfies the independence property if $P \sim Q, R \sim S$, $P_2 = R_2 = y_1, Q_2 = S_2 = y_2$ with with $y_1, y_2 \in \Sigma^2(z)$ and there exist $x_1, x_2 \in X_1$ such that $P, Q, R, S \in \Gamma_{2,y_1}(x_1|z) \cap \Gamma_{2,y_2}(x_2|z)$.*

Proof of Lemma 11. Suppose $(P, Q, R, S) \in (\mathcal{L}^0(X_1) \times X_2)^4$ is a proper tuple that satisfies the conditions stated in the lemma. By Lemma 5, there exist $P_1', R_1' \in \Pi^1(x_1|z)$ and $Q_1', S_1' \in \Pi^1(x_2|z)$ such that $P' := (P_1', P_2) \sim P \sim Q \sim Q' := (Q_1', Q_2)$ and $R' := (R_1', R_2) \sim R \sim S \sim S' := (S_1', S_2)$. By (ii) of Lemma 4, for any $\alpha \in (0, 1)$, $\alpha P' + (1 - \alpha)R' \sim \alpha Q' + (1 - \alpha)S'$. Lemma 3 implies that $\alpha P' + (1-\alpha)R' \sim \alpha P + (1-\alpha)R$ and $\alpha Q' + (1-\alpha)S' \sim \alpha Q + (1-\alpha)S$. By transitivity of $\succsim$, $\alpha P + (1 - \alpha)R \sim \alpha Q + (1 - \alpha)S$. $\square$

Given Lemma 10, we can extend the local property in Lemma 11 to a global one.

Lemma 12. *In Case 2, suppose $z = z_0$ or z_0'. Then a proper tuple $(P, Q, R, S) \in (\mathcal{L}^0(X_1) \times X_2)^4$ satisfies the independence property if $P_2 = R_2 = y_1, Q_2 = S_2 = y_2$ with $y_1, y_2 \in \Sigma^2(z)$.*

Proof of Lemma 12. The independence property holds for (P, Q, R, S) trivially if $y_1 = y_2$. Without loss of generality, we assume $y_1 > y_2$ and $P \succ R$. Fix $\hat{T} \in \mathcal{L}^0(X_1) \times \{y_2\}$ and $a \in X_1$ with $\hat{T} \succ (a, y_1) \succ (a, y_2)$. Take any $T_1 = (z_1, y_2), T^2 = (z_2, y_2)$ with $\hat{T} \succ T^1 \succ T^2 \succ (a, y_1)$. As $y_1, y_2 \in \Sigma^2(z)$, by Lemma 3, we can find $\hat{x}_1$, $\hat{x}_2 \in X_1^o$ and $p_1, q_1 \in \Pi^1(\hat{x}_1|z)$, $p_2, q_2 \in \Pi^1(\hat{x}_2|z)$ such that

$$(q_1, y_1) \sim T^1 \prec (p_1, y_1) \text{ and } (q_2, y_2) \sim T^1 \prec (p_2, y_2).$$

By Lemma 5, we know that

$$[T^1, (p_1, y_1)] \cap [T^1, (p_2, y_2)] \subseteq \Gamma(\hat{x}_1, y_1) \cap \Gamma(\hat{x}_2, y_2).$$

For any $z' \in [z_2, z_1]$, we can choose $\lambda_{z'}$ and $\eta_{z'} \in [0, 1]$ such that

$$\lambda_{z'}(q_1, y_1) + (1 - \lambda_{z'})(a, y_1) \sim (z', y_2) \sim \eta_{z'}(q_2, y_2) + (1 - \eta_{z'})(a, y_2).$$

By Lemma 3, $\lambda_{z'}(p_1, y_1) + (1 - \lambda_{z'})(a, y_1) \succ \lambda_{z'}(q_1, y_1) + (1 - \lambda_{z'})(a, y_1)$ and $\eta_{z'}(p_2, y_2) + (1 - \eta_{z'})(a, y_2) \succ \eta_{z'}(q_2, y_2) + (1 - \eta_{z'})(a, y_2)$. Denote $\hat{x}_1^{z'}$, $\hat{x}_2^{z'}$ for each $z' \in [z_1, z_2]$ with

$$\lambda_{z'} q_1 + (1 - \lambda_{z'})\delta_a \in \Pi^1(\hat{x}_1^{z'}|z), \text{ and } \eta_{z'} q_2 + (1 - \eta_{z'})\delta_a \in \Pi^1(\hat{x}_2^{z'}|z).$$

This leads to

$$[(z', y_2), (\lambda_{z'} p_1 + (1 - \lambda_{z'})\delta_a, y_1)] \cap [(z', y_2), (\eta_{z'} p_2 + (1 - \eta_{z'})\delta_a, y_2)] \subseteq \Gamma_{2,y_1}(\hat{x}_1^{z'}|z) \cap \Gamma_{2,y_2}(\hat{x}_2^{z'}|z).$$

Take the union across all z' between z_1 and z_2, and by Continuity, we have

$$[T^2, T^1] \subseteq \bigcup_{z_2 \le z' \le z_1} \Big(\Gamma_{2,y_1}(\hat{x}_1^{z'}|z) \cap \Gamma_{2,y_2}(\hat{x}_2^{z'}|z) \Big). \tag{8}$$

In order to get an open cover of $[T^2,T^1]$, for $\varepsilon>0$ small enough with $\hat{T}\succ(z_1+\varepsilon,y_2)\succ(z_2-\varepsilon,y_2)\succ(a,y_1)$, we have

$$[T^2,T^1]\subseteq\bigcup_{z_2-\varepsilon\le z'\le z_1+\varepsilon}\Big(\Gamma_{2,y_1}(\hat{x}_1^{z'}|z)\cap\Gamma_{2,y_2}(\hat{x}_2^{z'}|z)\Big).$$

For each $z_2-\varepsilon\le z'\le z_1+\varepsilon$, $\Gamma_{2,y_1}(\hat{x}_1^{z'}|z)\cap\Gamma_{2,y_2}(\hat{x}_2^{z'}|z)$ has a non-empty interior. Hence we can find an open cover of $[T^2,T^1]=[(z_2,y_2),(z_1,y_2)]$ as $\{C^{z'}\}_{z_2-\varepsilon\le z'\le z_1+\varepsilon}$ with $C^{z'}\subset\Gamma_{2,y_1}(\hat{x}_1^{z'}|z)\cap\Gamma_{2,y_2}(\hat{x}_2^{z'}|z)$. Notice that $X_1\times\{y_2\}$ is isomorphic to $X_1\subseteq\mathbb{R}$ and $[T^2,T^1]=[(z_2,y_2),(z_1,y_2)]$ is isomorphic to $[z_2,z_1]$, which is a closed and bounded interval. By Heine–Borel theorem, we can find a finite subcover of $\{C^{z'}\}_{z_2-\varepsilon\le z'\le z_1+\varepsilon}$ for $[T^2,T^1]$. Denote the subcover as $\{C^{z_k}\}_{k=1}^K$ and we can assume that $C^{z_k}\cap C^{z_{k+1}}\neq\emptyset$ for each $k=1,...,K-1$.

Take any proper tuple (P,Q,R,S) with $P\sim Q,R\sim S$, $P_2=R_2=y_1,Q_2=S_2=y_2$ with $y_1>y_2\in\Sigma^2(z)$. By Lemma 11, the independence property holds for (P,Q,R,S) if $P,Q,R,S\in C^{z_k}$ for any $k=1,...,K$. Then Lemma 10 implies that the independence property holds for (P,Q,R,S) if $P,Q,R,S\in[T^2,T^1]\subseteq\bigcup_{k=1}^K C^{z_k}$.

Recall that we can denote $X_1=[\underline{c}_1,\bar{c}_1]$ where $\underline{c}_1\in\mathbb{R}\cup\{-\infty\}$ and $\bar{c}_1\in\mathbb{R}\cup\{+\infty\}$. By arbitrariness of $\hat{T},T^1$, T^2 and $a\in X_2$, the given tuple (P,Q,R,S) satisfies the independence property if $(\bar{c}_1,y_2)\succ P\sim Q\succ R\sim S\succ(\underline{c}_1,y_1)$.

Now suppose $(\bar{c}_1,y_2)\succ P$ and $R\sim(\underline{c}_1,y_1)$, where $\underline{c}_1>-\infty$. By Monotonicity, $R=(\underline{c}_1,y_1)$. Take a sequence of $\{\lambda_n\}_{n\ge1}\subset(0,1)$ with $\lambda_n\to0$. For each n, denote $S^n=(\lambda_nQ_1+(1-\lambda_n)S_1,S_2)$ and by Lemma 3, we can find β_n with $R^n=(\beta_nP_1+(1-\beta_n)R_1,R_2)\sim S^n$. Clearly, $R^n\succ R=(\underline{c}_1,y_1)$ for each n and hence the independence property holds for (P,Q,R^n,S^n), that is, for each $\alpha\in(0,1)$,

$$\alpha P+(1-\alpha)R^n\sim\alpha Q+(1-\alpha)S^n.$$

As n goes to infinity, β_n converges to 0 and hence $S^n\xrightarrow{w}S,R^n\xrightarrow{w}R$. By continuity of $\succsim$ on $\hat{\mathcal{P}}$, we have for each $\alpha\in(0,1)$,

$$\alpha P+(1-\alpha)R\sim\alpha Q+(1-\alpha)S.$$

A similar proof works for the case where $P\sim(\bar{c}_1,y_2)$ and/or $R\sim(\underline{c}_1,y_1)$. Hence, the independence property holds for all (P,Q,R,S) with $P\sim Q,R\sim S$, $P_2=R_2=y_1,Q_2=S_2=y_2$ with $y_1,y_2\in\Sigma^2(z)$ for $z\in\{z_0,z_0'\}$. □

Lemma 13. *In Case 2, a proper tuple $(P,Q,R,S)\in(\mathcal{L}^0(X_1)\times X_2)^4$ satisfies the independence property if $P_2=R_2=y_1,Q_2=S_2=y_2$ with $y_1,y_2\in X_2$.*

Proof of Lemma 13. By Lemma 8, $\Sigma^2(z_0)\cup\Sigma^2(z_0')=X_2$ and $\Sigma^2(z_0)\cap\Sigma^2(z_0')\neq\emptyset$. By Lemma 12, it suffices to consider the case where $y_1\in\Sigma^2(z_0)\backslash\Sigma^2(z_0')$ and $y_2\in\Sigma^2(z_0')\backslash\Sigma^2(z_0)$. Without loss of generality, assume $y_1>y_2$ and $P\sim Q\succ R\sim S$.

Since $\Sigma^2(z_0)$ and $\Sigma^2(z_0')$ are open, there exist $z\in\Sigma^2(z_0)\cap\Sigma^2(z_0')$ with $y_1>z>y_2$.

By Monotonicity and Lemma 3, we can find $P' = (P'_1, z)$ and $R' = (R'_1, z)$ such that $P' \sim P \sim Q, R' \sim R \sim S$. By Lemma 12, the independence property holds for (P, P', R, R') and (Q, P', S, R') respectively. Hence for any $\alpha \in (0, 1)$,

$$\alpha P + (1-\alpha)R \sim \alpha P' + (1-\alpha)R' \sim \alpha Q + (1-\alpha)S.$$

$\square$

We are now ready to show that $\succsim$ can be represented by V^{ANB-CN}.

Recall that we can focus on $\succsim$ restricted to $\mathcal{L}^0(X_1) \times X_2$. For any $(p_1, x), (p_2, x) \in \mathcal{L}^0(X_1) \times X_2$ with $(p_1, x) \succ (p_2, x)$, we claim that there exists some function f representing $\succsim$ on $(\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$ such that f is continuous and linear in the first dimension, that is, for any $(q_1, y), (q_2, y) \in (\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$ and $\alpha \in [0, 1]$, $f(\alpha q_1 + (1-\alpha) q_1, y) = \alpha f(q_1, y) + (1-\alpha) f(q_2, y)$. In addition, f is unique up to a positive affine transformation. We call such f a MAP_1 function.

To prove the claim, notice that by Lemma 4, for any $Q \in [(p_2, x), (p_1, x)]$, there exists a unique $\lambda_Q \in [0, 1]$ such that $Q \sim \lambda_Q(p_1, x) + (1-\lambda_Q)(p_2, x)$. Define $f : [(p_2, x), (p_1, x)] \to [0, 1]$ such that $f(Q) = \lambda_Q$. As $(p_1, x) \succ (p_2, x)$, we have

$$f(Q) \geq f(Q') \Longleftrightarrow Q \sim \lambda_Q(p_1, x) + (1-\lambda_Q)(p_2, x) \succsim \lambda_{Q'}(p_1, x) + (1-\lambda_{Q'})(p_2, x) \sim Q'$$

Hence f represents $\succsim$ on $[(p_2, x), (p_1, x)]$. Continuity of $\succsim$ on $\hat{\mathcal{P}}$ assures that f is continuous. Then we show the linearity of f in the first dimension on $(\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$. Take any $(q_1, y), (q_2, y) \in (\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$. By definition of f and Lemma 3, we have

$$\begin{aligned}(q_1, y) &\sim f(q_1, y)(p_1, x) + (1 - f(q_1, y))(p_2, x),\\ (q_2, y) &\sim f(q_2, y)(p_1, x) + (1 - f(q_2, y))(p_2, x).\end{aligned}$$

Clearly, for any $\alpha \in (0, 1)$, $\alpha(q_1, y) + (1-\alpha)(q_2, y) \in (\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$. By definition of f and Lemma 12,

$$\begin{aligned}\alpha(q_1, y) + (1-\alpha)(q_2, y) \sim &[\alpha f(q_1, y) + (1-\alpha) f(q_2, y)](p_1, x)\\ &+ [1 - \alpha f(q_1, y) - (1-\alpha) f(q_2, y)](p_2, x)\\ \sim &f(\alpha q_1 + (1-\alpha) q_2, y)(p_1, x)\\ &+ (1 - f(\alpha q_1 + (1-\alpha) q_2, y))(p_2, x)\end{aligned}$$

By Lemma 3 given x in source 2, we know $f(\alpha q_1 + (1-\alpha) q_2, y) = \alpha f(q_1, y) + (1-\alpha) f(q_2, y)$. Easy to see that a positive affine transformation of f also represents $\succsim$ on $(\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$.

Now suppose that f, g represent $\succsim$ on $(\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$ and they are continuous and linear in the first dimension. Without loss of generality, let $f(p_2, x) = g(p_2, x), f(p_1, x) = g(p_1, x)$. Recall that for any $Q \in (\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$, there is a unique λ_Q with $Q \sim \lambda_Q(p_1, x) + (1 - \lambda_Q)(p_2, x)$. By linearity of f and g in the first dimension on $(\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$, we have

$$\begin{aligned} f(Q) =& \lambda_Q f(p_1, x) + (1 - \lambda_Q) f(p_2, x) \\ =& \lambda_Q f(p_1, x) + (1 - \lambda_Q) f(p_2, x) \\ =& g(Q) \end{aligned}$$

Hence $f \equiv g$ on $(\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$ and f is unique up to a positive affine transformation.

Now suppose that $\succsim$ admits a MAP_1 representation on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^2, T^1]$ and $(\mathcal{L}^0(X_1) \times X_2) \cap [T^4, T^3]$ where $T^1_2 = T^2_2 = x, T^3_2 = T^4_2 = y \neq x$. If $[T^2, T^1] \cap [T^4, T^3]$ has an empty interior, then it is trivial to show that $\succsim$ admits a MAP_1 representation on $(\mathcal{L}^0(X_1) \times X_2) \cap \Big([T^2, T^1] \cup [T^4, T^3]\Big)$. Now assume without loss of generality that $T^1 \succ T^3 \succ T^2 \succ T^4$. Denote the MAP_1 function on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^2, T^1]$ as f_x and the MAP_1 function on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^4, T^3]$ as f_y. Then f_x must be a positive affine transformation of f_y on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^2, T^3]$. Fix f_x and let $f_y(T^i) = f_x(T^i)$, $i = 2, 3$, then we have $f_x = f_y$ on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^2, T^3]$. Define $\hat{f} = f_x$ on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^2, T^1]$ and $\hat{f} = f_y$ on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^4, T^3]$. Then easy to show that $\hat{f}$ is a MAP_1 function on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^4, T^1]$ and is unique up to a positive affine transformation.

Repeat the above arguments, we can show that if $\succsim$ admits a MAP_1 representation on $(\mathcal{L}^0(X_1) \times X_2) \cap [T^2, T^1]$ and $(\mathcal{L}^0(X_1) \times X_2) \cap [T^4, T^3]$, then it also admits a MAP_1 representation on $(\mathcal{L}^0(X_1) \times X_2) \cap \Big([T^2, T^1] \cup [T^4, T^3]\Big)$ and the utility function is unique up to a positive affine transformation. Recall that $\succsim$ admits a MAP_1 representation on $(\mathcal{L}^0(X_1) \times X_2) \cap [(p_2, x), (p_1, x)]$. By varying $p_1, p_2 \in \mathcal{L}^0(X_1)$ and $x \in X_2$, we can show that there exists a MAP_1 function f^* representing $\succsim$ on $\mathcal{L}^0(X_1) \times X_2$ and it is unique up to a positive affine transformation.

Define $w(x, y) = f^*(x, y)$ for all $(x, y) \in X_1 \times X_2$. By construction of f^*, w is continuous. By Monotonicity, w is strictly monotonic. Recall that $\succsim$ satisfies Correlation Neutrality and for any $(p, q) \in \hat{\mathcal{P}}$, $(p, q) \succsim (p, CE_{v_2}(q))$ for some continuous and strictly monotonic v_2. Hence for any $P, Q \in \mathcal{P}$,

$$\begin{aligned} P \succsim Q &\Longleftrightarrow (P_1, CE_{v_2}(P_2)) \succsim (Q_1, CE_{v_2}(Q_2)) \\ &\Longleftrightarrow \sum_x w(x, CE_{v_2}(P_2)) P_1(x) \geq \sum_y w(y, CE_{v_2}(Q_2)) Q_1(y) \\ &\Longleftrightarrow V^{ANB-CN}(P) \geq V^{ANB-CN}(Q). \end{aligned}$$

Thus, in Case 2, $\succsim$ is represented by V^{ANB-CN}.

Case 3: There exists $(z_1, z_2) \in X_1 \times X_2$ such that $(p, q) \sim (p', q)$ for all $(x, y) \in X_1 \times X_2$, $p, p' \in \Pi^1(x|z_2)$ and $q \in \Pi^2(y|z_1)$, and we can find $(x_0, y_0) \in X_1 \times X_2$, $p_0 \in \Pi^1(x_0|z_2)$ and $q_0, q_0' \in \Pi^2(y_0|z_1)$ with $(p_0, q_0) \succ (p_0, q_0')$.

This is symmetric to Case 2 and hence we can show that $\succsim$ can be represented by

$$\hat{V}(P) = \sum_y w(CE_{v_1}(P_1), y)P_2(y)$$

where v_1 and w are continuous and strictly monotonic. However, we will show that given part (iii) of Independence*, $\succsim$ actually admits a CN representation.

Since $\succsim$ is represented by $\hat{V}$, we can focus on the restriction of $\succsim$ on $(X_1 \times \mathcal{L}^0(X_2))^4$ and a proper tuple $(P,Q,R,S) \in \hat{\mathcal{P}}^4$ satisfies the independence property if $P_1 = R_1, Q_1 = S_1$.

By Lemma 9, we can focus on the tuple with $P \sim Q, R \sim S$. Fix $y_0 \in X_2^o$ and let $P_2 = R_2 = y_0$. Denote $\hat{v}_{1|y_0}(x) = w(x, y_0)$ for each $x \in X_1$. We want to show that $\hat{v}_{1|y_0}$ is a positive affine transformation of v_1.

Define $\Sigma^1(z) := \{x \in X_1 : \exists\, y \in X_2, q, q' \in \Pi^2(y|z) \text{ s.t. } (x,q) \succ (x,q')\}$. As Lemma 8, we know that there exist $z_0, z_0' \in X_1$ such that $\Sigma^1(z_0) \cup \Sigma^1(z_0') = X_1$ and $\Sigma^1(z_0) \cap \Sigma^1(z_0') \neq \emptyset$.

Now consider $Q \sim P$ such that $Q_1 = x_0 \in X_1^0$. Then $x_0 \in \Sigma^1(z)$ for some $z \in \{z_0, z_0'\}$. Then there exist $h(x_0) \in X_1$ with $x_0 < h(x_0)$ and $[(x_0, y_0), (h(x_0), y_0)] \subset \Gamma_{1,x_0}(y_0|z)$. By Lemma 5 and Lemma 4, for any $\alpha \in (0,1)$ and $P \sim Q \in [(x_0,y_0),(h(x_0),y_0)]$, let $R = S = (x_0, y_0)$, we have $\alpha P + (1-\alpha)R = \alpha Q + (1-\alpha)R$. That is, for any $x \in [x_0, h(x_0)]$,

$$\alpha \hat{v}_{1|y_0}(x) + (1-\alpha)\hat{v}_{1|y_0}(x_0) = \hat{v}_{1|y_0}(v_1^{-1}(\alpha v_1(x) + (1-\alpha)v_1(x_0))).$$

Normalize that $\hat{v}_{1|y_0}(x_0) = v_1(x_0)$ and $\hat{v}_{1|y_0}(h(x_0)) = v_1(h(x_0))$ and by continuity and monotonicity of v_1 and $\hat{v}_{1|y_0}$, we know $\hat{v}_{1|y_0}(x) = v_1(x)$ for all $x \in [x_0, h(x_0)]$.

When x_0' is close enough to x_0, $[x_0, h(x_0)]] \cap [x_0', h(x_0')]$ has nonempty interior and we can show that $\hat{v}_{1|y_0}(x) = v_1(x)$ for all $x \in [x_0, h(x_0)] \cup [x_0', h(x_0')]$. Then we vary x_0 across X_1^0 and conclude that $\hat{v}_{1|y_0}(x) = v_1(x)$ for all $x \in X_1^0$. By continuity of $\hat{v}_{1|y_0}$ and v_1, $\hat{v}_{1|y_0}(x) = v_1(x)$ for all $x \in X_1$. Since $y_0 \in X_2^o$ is arbitrarily selected and by continuity of w, for each $y \in X_2$ $\hat{v}_{1|y} = w(\cdot, y)$ is a positive affine transformation of v_1. Hence for each $P \in \mathcal{P}$ $\hat{V}(P)$ can be rewritten as

$$\hat{V}(P) = \sum_y w(CE_{v_1}(P_1), y)P_2(y) = \sum_{x,y} w(x,y)P_1(x)P_2(y).$$

Thus, $\succsim$ admits a CN representation.

Case 4: For any $(z_1, z_2) \in X_1 \times X_2$, there exist $x_1, x_2 \in X_1, y_1, y_2 \in X_2$ and $(p_1, q_1), (p_1, q_1') \in \Pi^1(x_1|z_2) \times \Pi^2(y_1|z_1)$, $(p_2, q_2), (p_2', q_2) \in \Pi^1(x_2|z_2) \times \Pi^2(y_2|z_1)$ such that $(p_1, q_1) \succ (p_1, q_1')$ and $(p_2, q_2) \succ (p_2', q_2)$.

For each $(z_1, z_2) \in X_1 \times X_2$, denote

$$\hat{\Sigma}^1(z_1,z_2) := \{x \in X_1 : \exists\, y \in X_2 \text{ and } p \in \Pi^1(x|z_2), q, q' \in \Pi^2(y|z_1) \text{ s.t. } (p,q) \succ (p,q')\},$$
$$\hat{\Sigma}^2(z_1,z_2) := \{y \in X_2 : \exists\, x \in X_1 \text{ and } p, p' \in \Pi^1(x|z_2), q \in \Pi^2(y|z_1) \text{ s.t. } (p,q) \succ (p',q)\}.$$

Then we know $\hat{\Sigma}^1(z_1,z_2)$ and $\hat{\Sigma}^1(z_1,z_2)$ are open and $\hat{\Sigma}^1(z_1,z_2) \neq \emptyset$ and $\hat{\Sigma}^2(z_1,z_2) \neq \emptyset$. Furthermore the following lemma holds.

Lemma 14. *In Case 4, suppose $(z_1, z_2) \in X_1 \times X_2$. Then for any $(x, y) \in (\hat{\Sigma}^1(z_1, z_2) \cap X_1^o) \times (\hat{\Sigma}^2(z_1, z_2) \cap X_2^o)$, there exist $P, Q, R, S \in \Pi^1(x|z_2) \times \Pi^2(y|z_1)$ such that $P_1 = Q_1$, $R_2 = S_2$, $P \succ Q$ and $R \succ S$.*

Proof of Lemma 14. By definition, there exist $x' \in X_1, y' \in X_2$ and $(p_1, q_1), (p_1, q_1') \in \Pi^1(x|z_2) \times \Pi^2(y'|z_1)$, $(p_2, q_2), (p_2', q_2) \in \Pi^1(x'|z_2) \times \Pi^2(y|z_1)$ such that $(p_1, q_1) \succ (p_1, q_1')$ and $(p_2, q_2) \succ (p_2', q_2)$. By Lemma 3, given p_1 in source 1, as $y \in X_2^o$, we can find $\hat{q}_1, \hat{q}_1' \in \Pi^2(y|z_1)$ with $(p_1, \hat{q}_1) \succ (p_1, \hat{q}_1')$. Similarly, given q_2 in source 2, as $x \in X_1^o$, we can find $\hat{p}_2, \hat{p}_2' \in \Pi^1(x_1|z_2)$ with $(\hat{p}_2, q_2) \succ (\hat{p}_2', q_2)$. This completes the proof. □

For any $x_1 \in \hat{\Sigma}^1(z_1, z_2)$ and $x_2 \in \hat{\Sigma}^2(z_1, z_2)$, we denote

$$\hat{\Pi}^1(x_1|z_1, z_2) = \Big\{p \in \Pi^1(x_1|z_2) : \exists\, x_2' \in X_2, q \in \Pi^2(x_2'|z_1) \text{ s.t. } (p, q) \not\sim (p, x_2')\Big\},$$
$$\hat{\Pi}^2(x_2|z_1, z_2) = \Big\{q \in \Pi^2(x_2|z_1) : \exists\, x_1' \in X_1, p \in \Pi^1(x_1'|z_2) \text{ s.t. } (p, q) \not\sim (x_1', q)\Big\}.$$

By definition $\hat{\Pi}^i(x_i|z_1, z_2) \subseteq \Pi^i(x_i|z_{-i})$ for $i = 1, 2$.

Lemma 15. *For each $i = 1, 2$ and $x_i \in \Sigma^i$, $cl(\hat{\Pi}^i(x_i|z_1, z_2)) = \Pi^i(x_i|z_{-i})$.*

Proof of Lemma 15. We will prove the result for $i = 1$. The proof for $i = 2$ is symmetric. By definition of $x_1 \in \hat{\Sigma}^1(z_1, z_2)$, we can find $p \in \hat{\Pi}^1(x_1|z_1, z_2)$, $x_2' \in X_2$ and $q, q' \in \Pi^2(x_2'|z_1)$ such that $(p, q) \succ (p, q')$. For any $p^o \in \Pi^1(x_1|z_2) \backslash \hat{\Pi}^1(x_1|z_1, z_2)$, we have $(p^o, q) \sim (p^o, q')$. By Lemma 4, the independence property holds for $((p, q), (p, q'), (p^o, q), (p^o, q'))$ as $p, p^o \in \Pi^1(x_1|z_2)$. Then for any $\lambda \in (0, 1)$, $(\lambda p + (1-\lambda)p^o, q) \succ (\lambda p + (1-\lambda)p^o, q')$, which implies $\lambda p + (1-\lambda)p^o \in \hat{\Pi}^1(x_1|z_1, z_2)$. Let $\lambda \to 0$ and we have $p^o \in cl(\hat{\Pi}^1(x_1|z_1, z_2))$. □

Now we introduce the local independence property similar to Lemma 11. Recall that by Lemma 9, it suffices to focus on the tuple (P, Q, R, S) with $P \sim Q$ and $R \sim S$.

Lemma 16. *In Case 4, suppose $(z_1, z_2) \in X_1 \times X_2$. Then a proper tuple (P, Q, R, S) satisfies the independence property if there exist some $i, j \in \{1, 2\}$, $x_i \in \hat{\Sigma}^i(z_1, z_2) \cap X_i^o$, $x_j \in \hat{\Sigma}^j(z_1, z_2) \cap X_j^o$ and $y \in X_{-i}$, $y' \in X_{-j}$ such that $P_i = R_i = p \in \hat{\Pi}^i(x_i|z_1, z_2)$, $Q_j = S_j = q \in \hat{\Pi}^j(x_j|z_1, z_2)$, $P \sim Q$, $R \sim S$ and $P, R \in \Gamma_{i,p}(y|z_i) \cap \Gamma_{j,q}(y'|z_j)$.*

Proof of Lemma 16. By Lemma 5, we can find P', R' with $P_i' = R_i' = p$ and $P_{-i}', R_{-i}' \in \Pi^{-i}(y|z_i)$ such that $P \sim P'$ and $R \sim R'$. By Lemma 3, for any $\lambda \in (0, 1)$, $\lambda P + (1-\lambda)R \sim \lambda P' + (1-\lambda)R'$. Similarly, we can find Q', S' with $Q_j' = S_j' = q$, $Q_{-j}', R_{-j}' \in \Pi^{-j}(y'|z_j)$ and $\lambda Q + (1-\lambda)S \sim \lambda Q' + (1-\lambda)S'$ for any $\lambda \in (0, 1)$. Easy to check that the conditions in (ii) of Lemma 4 hold for the tuple (P', Q', R', S') and hence for any $\lambda \in (0, 1)$,

$$\lambda P + (1-\lambda)R \sim \lambda P' + (1-\lambda)R' \sim \lambda Q' + (1-\lambda)S' \sim \lambda Q + (1-\lambda)S.$$

□

The independence property in Lemma 16 is non-trivial, since by definition of $\hat{\Sigma}^i(z_1, z_2)$ and $\hat{\Pi}^i(x_i|z_1, z_2)$ and Lemma 14, for each $i, j \in \{1, 2\}$, we can find $P, Q, R, S \in \Gamma_{i,p}(y|z_i) \cap \Gamma_{j,q}(y'|z_j)$ such that $P \sim Q \succ R \sim S$ and $P_i = R_i = p, Q_j = S_j = q$. Like in Case 2, our goal is to extend the local property in Lemma 16. The arguments are essentially the same as Lemma 10, Lemma 12 and Lemma 13 in Case 2. Here we just summarize the main steps.

- Step 1. Fix $(z_1, z_2) \in X_1 \times X_2$. Consider a proper tuple (P, Q, R, S) where there exist some $i, j \in \{1, 2\}$, $x_i \in \hat{\Sigma}^i(z_1, z_2) \cap X_i^o$ and $x_j \in \hat{\Sigma}^j(z_1, z_2) \cap X_j^o$ such that $P_i = R_i = p \in \hat{\Pi}^i(x_i|z_1, z_2)$, $Q_j = S_j = q \in \hat{\Pi}^j(x_j|z_1, z_2)$, $P \sim Q, R \sim S$,
  - Step 1.1. Use arguments in Lemma 10, Lemma 12 to show that the independence property holds if $P, R \in \Gamma_{j,q}(y'|z_j)$ for some $y' \in X_{-j}$.
  - Step 1.2. Again use arguments in Lemma 10, Lemma 12 to show that the independence property always holds.
- Step 2. Fix $(z_1, z_2) \in X_1 \times X_2$. By Continuity and Lemma 15, a proper tuple (P, Q, R, S) satisfies the independence property if there exist some $i, j \in \{1, 2\}$, $x_i \in \hat{\Sigma}^i(z_1, z_2)$ and $x_j \in \hat{\Sigma}^j(z_1, z_2)$ such that $P_i = R_i = p \in \Pi^i(x_i|z_{-i})$, $Q_j = S_j = q \in \Pi^j(x_j|z_{-j})$ and $P \sim Q, R \sim S$.
- Step 3. Now we vary $(z_1, z_2) \in X_1 \times X_2$. By a similar argument in Lemma 8, we can find $(z_1', z_2'), (z_1'', z_2'') \in X_1 \times X_2$ such that

$$\begin{aligned}
&\hat{\Sigma}^1(z_1', z_2') \cup \hat{\Sigma}^1(z_1'', z_2') = \hat{\Sigma}^1(z_1', z_2'') \cup \hat{\Sigma}^1(z_1'', z_2'') = X_1, \\
&\hat{\Sigma}^1(z_1', z_2') \cup \hat{\Sigma}^1(z_1', z_2'') = \hat{\Sigma}^1(z_1'', z_2') \cup \hat{\Sigma}^1(z_1'', z_2'') = X_2, \\
&\hat{\Sigma}^1(z_1', z_2') \cap \hat{\Sigma}^1(z_1'', z_2') \neq \emptyset \ , \ \hat{\Sigma}^1(z_1', z_2'') \cap \hat{\Sigma}^1(z_1'', z_2'') \neq \emptyset, \\
&\hat{\Sigma}^1(z_1', z_2') \cap \hat{\Sigma}^1(z_1', z_2'') \neq \emptyset \ , \ \hat{\Sigma}^1(z_1'', z_2') \cap \hat{\Sigma}^1(z_1'', z_2'') \neq \emptyset.
\end{aligned}$$

  Then by the same arguments in Lemma 13, we can show that a proper tuple (P, Q, R, S) satisfies the independence property if there exist some $i, j \in \{1, 2\}$ such that $P_i = R_i$, $Q_j = S_j$ and $P \sim Q, R \sim S$.

Now we will show that $\succsim$ can be represented by some V^{CN}. By Lemma 9 and Step 3 above, for any $P, Q, R, S \in \hat{\mathcal{P}}$, if $P_i = R_i$, $Q_j = S_j$ for some $i, j \in \{1, 2\}$, $R \sim S$ and $\alpha \in (0, 1)$, then

$$P \succsim Q \Longleftrightarrow \alpha P + (1 - \alpha)R \succsim \alpha Q + (1 - \alpha)S.$$

Apply Theorem 1 in Chapter 7.2 (Page 88) of Fishburn (1982) to the set of product lotteries $\hat{\mathcal{P}}$, there exists a continuous and multilinear[7] function V representing $\succsim$ on $\hat{\mathcal{P}}$, which is unique up to a positive affine transformation. By Correlation Neutrality, define V' on $\mathcal{P}$ such that $V'(P) = V(P_1, P_2)$ for each $P \in \mathcal{P}$. Then V' represents $\succsim$ on $\mathcal{P}$.

[7]Suppose $\mathcal{M}_1, \mathcal{M}_2 \subseteq \mathcal{L}^0(\mathbb{R})$ are mixture sets. A function V is multilinear on $\mathcal{M}_1 \times \mathcal{M}_2$ if $V(\lambda p + (1 - \lambda)r, q) = \lambda V(p, q) + (1 - \lambda)V(r, q)$ and $V(p, \lambda q + (1 - \lambda)s) = \lambda V(p, q) + (1 - \lambda)V(p, s)$ for all $\lambda \in (0, 1)$, $p, r \in \mathcal{M}_1$ and $q, s \in \mathcal{M}_2$.

We claim that V' is a CN representation. To see this, define $w(x_1, x_2) = V(x_1, x_2)$. Since V is continuous, w is continuous. Also by Monotonicity, w is continuous and strictly monotonic. Moreover, since V is multilinear and each lottery has a finite support, for each $P \in \mathcal{P}$,

$$V'(P) = V(P_1, P_2) = \sum_{x,y} V(x, y)P_1(x)P_2(y) = \sum_{x,y} w(x, y)P_1(x)P_2(y).$$

Thus in Case 4, $\succsim$ is represented by some $V^{CN} = V'$.

To summarize, if $\succsim$ satisfies Weak Order, Monotonicity, Continuity, Independence* and Correlation Neutrality, then $\succsim$ an be represented by one of the following types of utility functions: V^{CN} (Case 3 and Case 4), V^{ANB-CN} (Case 2) or V^{FNB} (Case 1). □

Proof of Theorem 2.
Necessity. Since Correlation Consistency is weaker than Correlation Neutrality, by Theorem 1, V^{CN}, V^{ANB-CN} and V^{FNB} satisfy all the axioms. Since w is continuous and strictly monotonic, V^{EU} also satisfy those axioms. Hence we can focus on V^{ANB} with

$$V^{ANB}(P) = \sum_{x} w(x, CE_v(P_{2|x}))P_1(x), \text{ for all } P \in \mathcal{P},$$

where w and v are continuous and strictly monotonic. This guarantees that V^{ANB} satisfies Weak Order and Monotonicity. Because V^{ANB} agrees with V^{ANB-CN} on $\hat{\mathcal{P}}$, V^{ANB} satisfies Continuity 3 and Independence*.

To verify Continuity 1, for any $P, Q \in \mathcal{P}$ and $\lambda \in [0, 1]$,

$$\begin{aligned} V^{ANB}(\lambda P + (1-\lambda)Q) =& \lambda \sum_{\substack{x:P_1(x)>0,\\ Q_1(x)=0}} w(x, CE_v(P_{2|x}))P_1(x) \\ &+ (1-\lambda) \sum_{\substack{x:Q_1(x)>0,\\ P_1(x)=0}} w(x, CE_v(Q_{2|x}))Q_1(x) \\ &+ \sum_{\substack{x:Q_1(x)>0,\\ P_1(x)>0}} w(x, CE_v(\alpha P_{2|x} + (1-\alpha)Q_{2|x}))[\lambda P_1(x) + (1-\lambda)Q_1(x)], \end{aligned}$$

where

$$\alpha = \frac{\lambda P_1(x)}{\lambda P_1(x) + (1-\lambda)Q_1(x)}.$$

Then $V^{ANB}(\lambda P + (1-\lambda)Q)$ is continuous in λ and $\succsim$ satisfies Continuity 1.

To verify Continuity 2, for ε_n small enough,

$$V^{ANB}(P_{\varepsilon_n}) = \sum_{x \in X_1} w\Big(\phi_{\varepsilon_n}(x, y)_1, v^{-1}\Big(\sum_{y \in X_2} v(\phi_{\varepsilon_n}(x, y)_2)P_{2|x}(y)\Big)\Big)P_1(x).$$

Recall that $\phi_\varepsilon(x,y) = (x+\varepsilon^1, y+\varepsilon^2)$ if $(x+\varepsilon^1, y+\varepsilon^2) \in Z$; otherwise, $\phi_\varepsilon(x,y) = (x', y')$ such that (x', y') is the element of Z closest to (x,y). This implies ϕ_ε is continuous in ε. By continuity of w and v_2, we know $V^{ANB}(P_{\varepsilon_n})$ converges to $V^{ANB}(P)$ as ε_n converges to $(0,0)$.

Finally we verify Correlation Consistency. For each $P, Q, R, S \in \mathcal{P}$ and $\alpha \in (0,1)$, if $P_i = Q_i$ for $i = 1,2$ and $\text{supp}(P_1) \cap \text{supp}(R_1) = \text{supp}(P_1) \cap \text{supp}(S_1) = \emptyset$, then

$$V^{ANB}(\alpha P + (1-\alpha)R) = \alpha V^{ANB}(P) + (1-\alpha)V^{ANB}(R),$$

$$V^{ANB}(\alpha Q + (1-\alpha)S) = \alpha V^{ANB}(Q) + (1-\alpha)V^{ANB}(S).$$

Hence $P \succ Q, R \sim S$ implies that $\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$.

Sufficiency. First, we utilize Theorem 1 to derive implications of axioms on $\hat{\mathcal{P}}$. Formally, define $\hat{\succsim}$ on $\mathcal{P}$ which satisfies Correlation Neutrality and agrees with $\succsim$ on $\hat{\mathcal{P}}$. Clearly, $\hat{\succsim}$ satisfies Weak Order, Monotonicity, Continuity, Independence* and Correlation Consistency. Then Theorem 1 immediately leads to the following corollary.

Corollary 6. *$\hat{\succsim}$ can be represented by one of the following types of utility functions: V^{CN}, V^{ANB-CN} or V^{FNB}. Moreover, if $\succsim$ satisfies Correlation Neutrality, then $\succsim = \hat{\succsim}$.*

In the rest of the proof, we maintain the assumption that $\succsim$ violates Correlation Neutrality, that is, there exists $P \succ \tilde{P}$ with $P_i = \tilde{P}_i$ for $i = 1,2$.

For any $(p,q) \in \hat{\mathcal{P}}$, denote $M(p,q)$ as the set of lotteries whose marginal lotteries are p and q respectively. For any $P, R \in \mathcal{P}$, we say P and R are *compatible*, or P is *compatible* with R if $\text{supp}(P_1) \cap \text{supp}(R_1) = \emptyset$. Easy to see that if R is compatible with both P and Q, then R is also compatible with $\lambda P + (1-\lambda)Q$ for any $\lambda \in (0,1)$. If P is compatible with Q, then P is compatible with any $Q' \in M(Q_1, Q_2)$.

The first main challenge is that betweenness is not directly implied by the axioms, that is, for $\lambda \in (0,1)$, it is not guaranteed that $P \succ \lambda P + (1-\lambda)\tilde{P} \succ \tilde{P}$. However, we have the following weaker and local version of the betweenness property.

Lemma 17. *For any $Q \succ Q'$, there exists $Q^* = \lambda^* Q + (1-\lambda^*)Q'$ for some $\lambda^* \in [0,1]$ such that for any $\varepsilon > 0$, we can find $\lambda_\varepsilon \in (\lambda^* - \varepsilon, \lambda^* + \varepsilon) \cap [0,1]$ with $Q^* \not\sim \lambda_\varepsilon Q + (1-\lambda_\varepsilon)Q'$.*

Proof of Lemma 17. Suppose the result fails. Then for any $\lambda \in [0,1]$, there exists $\varepsilon_\lambda > 0$ such that for any $\lambda' \in (\lambda - \varepsilon_\lambda, \lambda + \varepsilon_\lambda) \cap [0,1]$, $\lambda Q + (1-\lambda)Q' \sim \lambda' Q + (1-\lambda')Q'$. Notice that $\{(\lambda - \varepsilon_\lambda, \lambda + \varepsilon_\lambda)\}_{\lambda \in [0,1]}$ forms an open cover of the compact set $[0,1]$. We can find a finite subcover of $[0,1]$. By transitivity of $\succsim$, we know that $\lambda Q + (1-\lambda)Q' \sim \lambda' Q + (1-\lambda')Q'$ for all $\lambda, \lambda' \in [0,1]$, which leads to $Q \sim Q'$ and a contradiction. $\square$

The next Lemma shows that if $R \sim S$ and Q is compatible with R, S, then the independence property holds.

Lemma 18. *For any $\alpha \in (0,1)$ and $Q, R, S \in \mathcal{P}$ with $R \sim S$ and Q compatible with R, S,*

$$\alpha Q + (1-\alpha)R \sim \alpha Q + (1-\alpha)S.$$

Proof of Lemma 18. For $P \succ \tilde{P}$ with $\tilde{P} \in M(P_1, P_2)$. Recall that $X_1 = [\underline{c}_1, \bar{c}_1]$. If $P_1(\underline{c}_1) > 0$, then by Continuity 2, we can find $\varepsilon = (\varepsilon^1, 0)$ with $\varepsilon^1 > 0$ small enough such that $P_\varepsilon \succ \tilde{P}_\varepsilon$. Denote $P' = P_\varepsilon$ and $\tilde{P}' = \tilde{P}_\varepsilon$. Then $P' \succ \tilde{P}'$ with $\tilde{P}' \in M(P_1', P_2')$ and $P_1'(\underline{c}_1) = 0$. If $P_1'(\underline{c}_1) = 0$, repeat the above argument by choosing some $\varepsilon^1 < 0$ and we can construct $P'' \succ \tilde{P}''$ with $\tilde{P}'' \in M(P_1'', P_2'')$ and $P_1''(\underline{c}_1) = P_1''(\bar{c}_1) = 0$. Thus, without loss of generality, we assume $P_1(\underline{c}_1) = P_1(\bar{c}_1) = 0$, that is, $\text{supp}(P_1) \in X_1^o$.

Denote $P^* = \lambda^* P + (1 - \lambda^*)\tilde{P}$ as the lottery found in Lemma 17. Clearly, either $P^* \not\sim P$ or $P^* \not\sim \tilde{P}$. P_1, P_2 are not degenerate. By Lemma 17, for any $n > 0$, there exists $\lambda_n \in (\lambda^* - 1/n, \lambda^* + 1/n) \cap [0, 1]$ with $P^* \not\sim \lambda_n P + (1 - \lambda_n)\tilde{P} := P^n$. By completeness, for each n, either $P^n \succ P^*$ or $P^* \succ P^n$. Then we can find a subsequence of $\{P^n\}_{n\geq 1}$, still denoted as $\{P^n\}_{n\geq 1}$, such that either $P^n \succ P^*$ for all n or $P^* \succ P^n$ for all n. Suppose that the former case holds. Take any $R \sim S$ and R, S compatible with P. Correlation Consistency implies that for all $\alpha \in (0, 1)$ and $n \geq 1$, $\alpha P^n + (1 - \alpha)R \succ \alpha P^* + (1 - \alpha)S$. By Continuity 1, as n goes to infinity, we have $\alpha P^* + (1 - \alpha)R \succsim \alpha P^* + (1 - \alpha)S$. This holds for all $R \sim S$ with R, S compatible with P. By symmetry, for all $\alpha \in (0, 1)$ and $R \sim S$ with R, S compatible with P,

$$\alpha P^* + (1 - \alpha)S \sim \alpha P^* + (1 - \alpha)R.$$

The same result holds if $P^* \succ P^n$ for all n. Without loss of generality, we assume that $P^n \succ P^*$ for all n and $P^* \succ \tilde{P}$ from now on.

Fix any Q compatible with P and we know Q is also compatible with $\tilde{P}$, P^* and P^n for each n. By Correlation Consistency, for any $\beta \in (0, 1)$, $\beta P^* + (1 - \beta)Q \succ \beta\tilde{P} + (1 - \beta)Q$ and $\beta P^* + (1 - \beta)Q, \beta\tilde{P} + (1 - \beta)Q \in M(\beta P_1 + (1 - \beta)Q_1, \beta P_2 + (1 - \beta)Q_2)$. Similarly, as $P^n \succ P^*$ for all n, for any $\beta \in (0, 1)$, $\beta P^n + (1 - \beta)Q \succ \beta P^* + (1 - \beta)Q$ and $\beta P^n + (1 - \beta)Q \in M(\beta P_1 + (1 - \beta)Q_1, \beta P_2 + (1 - \beta)Q_2)$. For any $R \sim S$ with R, S compatible with both P, Q, R, S are also compatible with $\beta P^n + (1 - \beta)Q$ and $\beta P^* + (1 - \beta)Q$. With the same arguments as above, we can show that for any $\alpha \in (0, 1)$ and $\beta \in (0, 1)$,

$$\alpha[\beta P^* + (1 - \beta)Q] + (1 - \alpha)R \sim \alpha[\beta P^* + (1 - \beta)Q] + (1 - \alpha)S.$$

This can be rearranged as

$$\beta[\alpha P^* + (1 - \alpha)R] + (1 - \beta)[\alpha Q + (1 - \alpha)R] \sim \beta[\alpha P^* + (1 - \alpha)S] + (1 - \beta)[\alpha Q + (1 - \alpha)S]$$

Again by Continuity 1, let $\beta \to 0^+$ and we have

$$\alpha Q + (1 - \alpha)R \sim \alpha Q + (1 - \alpha)S, \tag{9}$$

for any $\alpha \in (0, 1)$, $R \sim S$, Q compatible with P, R, S and P compatible with Q, R, S.

Fix $P, \tilde{P}$ and Q such that P is compatible with Q, we want to strengthen property (9) by discarding the constraint that R, S are compatible with P. By Continuity 3, as $P \succ \tilde{P}$, we can find $\bar{\varepsilon} > 0$ such that for all $\varepsilon = (\varepsilon^1, 0)$ with $\varepsilon^1 \in (0, \bar{\varepsilon})$, we have $P_\varepsilon \succ \tilde{P}_\varepsilon$. Note that $\tilde{P}_\varepsilon, P_\varepsilon \in M(P_{\varepsilon,1}, P_2)$. Since $\text{supp}(P_1) \cup \text{supp}(Q_1)$ is finite and P is compatible with Q, we can make $\bar{\varepsilon}$ small enough such that for all $\varepsilon^1 \in (0, \bar{\varepsilon})$, $\text{supp}(P_{\varepsilon,1}) \subset X_1^o$, $\text{supp}(P_{\varepsilon,1}) \cap \text{supp}(Q_1) = \emptyset$

and $\tilde{P}_\varepsilon, P_\varepsilon$ are compatible with Q. Then

$$\alpha Q + (1-\alpha)R \sim \alpha Q + (1-\alpha)S,$$

for any $\varepsilon \in (0,\bar{\varepsilon})$, $\alpha \in (0,1)$, $R \sim S$, Q compatible with R, S, P and P_ε compatible with R, S.

Now we show that by varying ε, we can further get rid of the constraint that R, S are compatible with P_ε for some $\varepsilon \in (0,\bar{\varepsilon})$. This is guaranteed by the fact that each lottery in $\mathcal{P}$ has a finite support and $\text{supp}(P_{\varepsilon,1}) \subset X_1^o$ for each $\varepsilon^1 \in (0,\bar{\varepsilon})$. Concretely, for any $R \sim S$ with R, S compatible with Q, we can always find ε^* such that R, S are compatible with P_{ε^*}. Thus,

$$\alpha Q + (1-\alpha)R \sim \alpha Q + (1-\alpha)S,$$

for any $\alpha \in (0,1)$, $R \sim S$ and Q compatible with R, S, P. The same argument can be applied to relax the requirement that Q is compatible with P and we conclude that for any $Q \in \mathcal{P}$,

$$\alpha Q + (1-\alpha)R \sim \alpha Q + (1-\alpha)S,$$

for any $\alpha \in (0,1)$, $R \sim S$, Q compatible with R, S. $\square$

The second main challenge is Lemma 9 only holds for product lotteries and Lemma 18 does not directly imply that $\alpha Q + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$ if $R \succ S$ and Q is compatible with R, S. Moreover, in Lemma 18, we cannot choose $R = Q$ and hence we have not shown yet that $\alpha P + (1-\alpha)$ lies between P and Q even if they are compatible. Our next goal is to deal with these issues.

First, for a fixed $y \in X_2$, we will focus on the set of lotteries whose utilities are strictly bounded by two lotteries in $\mathcal{L}^0(X_1) \times \{y\}$, that is,

$$\Phi_{2,y} = \Big\{P \in \mathcal{P} : \exists\, T, T' \in \mathcal{P} \text{ with } T_2 = T_2' = y \text{ s.t. } T \succ P \succ T'\Big\}.$$

Lemma 19 is a counterpart of Lemma 5 and part (i) of Lemma 4 in $\mathcal{P}$.

Lemma 19. *(i). For each $P, Q, R \in \mathcal{P}$ with $P \succ Q \succ R$, there exists $\lambda \in (0,1)$ such that $\lambda P + (1-\lambda)R \sim Q$.*
(ii). For each $P \in \mathcal{P}$, there exists $(x_1, x_2) \in X_1 \times X_2$ such that $P \sim (x_1, x_2)$. Moreover, if $P \in \Phi_{2,y}$ for some $y \in X_2$, then we can choose $x_2 = y$.

Proof of Lemma 19. (i). Denote $A = \{\alpha \in (0,1) : \alpha P + (1-\alpha)R \succ Q\}$ and $\lambda = inf A$. We claim that $\lambda P + (1-\lambda)R \sim Q$. Suppose by contradiction that $\lambda P + (1-\lambda)R \not\sim Q$. If $\lambda P + (1-\lambda)R \succ Q$, then $\lambda \in A$, which is open by Continuity 1. Hence there exists $\lambda' < \lambda$ with $\lambda' \in A$, which contradicts with the definition of λ. If $\lambda P + (1-\lambda)R \prec Q$, then $\lambda \in \{\alpha \in (0,1) : \alpha P + (1-\alpha)R \prec Q\}$, which is also open. We can find $\varepsilon > 0$ such that $[\lambda, \lambda+\varepsilon) \subseteq (0,1)\backslash A$. Again a contradiction with the definition of λ. Hence $\lambda P + (1-\lambda)R \sim Q$. (Unlike part (i) of Lemma 4, λ is not necessarily unique here.)

(ii). For each $P \in \mathcal{P}$, denote $x_i = \max \operatorname{supp}(P_i), y_i = \min \operatorname{supp}(P_i)$ for $i = 1, 2$. By Monotonicity, $(x_1, x_2) \succsim P \succsim (y_1, y_2)$. Then either $(x_1, x_2) \succsim P \succsim (x_1, y_2)$ or $(x_1, y_2) \succsim P \succsim (y_1, y_2)$. Without loss of generality, suppose the former case holds. Using the same argument in part (i) of Lemma 4, we can find a unique $\lambda \in [0, 1]$ such that $P \sim (x_1, \lambda\delta_{y_1} + (1 - \lambda)\delta_{y_2})$. By Lemma 3, there exists $x_2' \in X_2$ where $P \sim (x_1, \lambda\delta_{y_1} + (1 - \lambda)\delta_{y_2}) \sim (x_1, x_2')$. If further $P \in \Phi_{2,y}$ for some $y \in X_2$, then we can find $p_1, p_1' \in \mathcal{L}^0(X_1)$ with $(p_1, y) \succ P \succ (p_1', y)$. By the same argument, we can find $x' \in X_1$ such that $P \sim (x', y)$. □

The next lemma generalizes Lemma 18 and shows that the independence axiom holds on $\Phi_{2,y}$ subject to the compatibility constraint.

Lemma 20. *For each $y \in X_2$ and $P, Q, R, S \in \Phi_{2,y}$, the following properties hold:*

i). $P \sim Q$ and P is compatible with $Q \Longrightarrow \alpha P + (1 - \alpha)Q \sim P \sim Q$ for all $\alpha \in (0, 1)$;

ii). $P \succ Q$ and P is compatible with $Q \Longrightarrow P \succ \alpha P + (1 - \alpha)Q \succ Q$ for all $\alpha \in (0, 1)$;

iii). $P \succ Q$, $R \sim S$, P is compatible with R and Q is compatible with $S \Longrightarrow \alpha P + (1 - \alpha)R \succ \alpha Q + (1 - \alpha)S$ for all $\alpha \in (0, 1)$;

iv). $P \sim Q$, $R \sim S$, P is compatible with R and Q is compatible with $S \Longrightarrow \alpha P + (1 - \alpha)R \sim \alpha Q + (1 - \alpha)S$ for all $\alpha \in (0, 1)$.

Proof of Lemma 20. We will prove (i) and (ii). (iii) and (iv) can be shown similarly. Suppose $P, Q \in \Phi_{2,y}$ for some $y \in X_2$ and P, Q are compatible. By Lemma 19, there exists $x_P, x_Q \in X_1^o$ such that $P \sim (x_P, y)$ and $Q \sim (x_Q, y)$. By Lemma 3, we can find $\varepsilon > 0$ such that for all $z_P \in [x_P - \varepsilon, x_P], z_Q \in [x_Q - \varepsilon, x_Q]$, there exist $z_P' \geq x_P, z_Q' \geq x_Q$ such that $P \sim (1/2\delta_{z_P} + 1/2\delta_{z_P'}, y)$ and $Q \sim (1/2\delta_{z_Q} + 1/2\delta_{z_Q'}, y)$. Moreover, as z_P, z_Q increases, z_P', z_Q' will be decreasing continuously. Since P, Q are simple, that is, $\operatorname{supp}(P_1) \cup \operatorname{sup}(Q_1)$ is finite, we can construct $z_P^* \neq z_Q^*$, $z_P^{*'} \neq z_Q^{*'}$ and $z_P^*, z_Q^*, z_P^{*'}, z_Q^{*'} \notin \operatorname{supp}(P_1) \cup \operatorname{sup}(Q_1)$. Denote $P' = (1/2\delta_{z_P^*} + 1/2\delta_{z_P^{*'}}, y)$, $Q' = (1/2\delta_{z_Q^*} + 1/2\delta_{z_Q^{*'}}, y)$. Then $P \sim P', Q \sim Q'$ and P, Q, P', Q' are compatible with each other. Apply Lemma 18 twice and we get for any $\alpha \in (0, 1)$,

$$\alpha P + (1 - \alpha)Q \sim \alpha P + (1 - \alpha)Q' \sim \alpha P' + (1 - \alpha)Q'.$$

Again by Lemma 3 given marginal lottery in source 2 as y, we know

$$\begin{aligned} &P \sim Q \Longrightarrow P' \sim Q' \Longrightarrow \alpha P + (1 - \alpha)Q \sim \alpha P' + (1 - \alpha)Q' \sim Q' \sim Q, \\ &P \succ Q \Longrightarrow P' \succ Q' \Longrightarrow P \sim P' \sim \alpha P + (1 - \alpha)Q \sim \alpha P' + (1 - \alpha)Q' \succ Q' \sim Q. \end{aligned}$$

□

Like Lemma 10, we claim that if the independence property holds on $\Phi_{2,y}$ for each y, then it also holds on their union.

Lemma 21. *For each $y \in X_2$ and $P, Q, R, S \in \cup_{y \in X_2} \Phi_{2,y}$, the following properties hold:*

i). $P \sim Q$ and P is compatible with $Q \Longrightarrow \alpha P + (1 - \alpha)Q \sim P \sim Q$ for all $\alpha \in (0, 1)$;

ii). $P \succ Q$ and P is compatible with $Q \implies P \succ \alpha P + (1-\alpha)Q \succ Q$ for all $\alpha \in (0,1)$;

iii). $P \succ Q$, $R \sim S$, P is compatible with R and Q is compatible with $S \implies \alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$ for all $\alpha \in (0,1)$;

iv). $P \sim Q$, $R \sim S$, P is compatible with R and Q is compatible with $S \implies \alpha P + (1-\alpha)R \sim \alpha Q + (1-\alpha)S$ for all $\alpha \in (0,1)$.

Proof of Lemma 21. First, for any $P,Q,R,S \in \cup_{y\in X_2}\Phi_{2,y}$, we claim that there exist finitely many $z_k \in X_2, k=1,...,K$ such that $z_1 < z_2 < ... < z_K$ and $P,Q,R,S \in \cup_{k=1}^K \Phi_{2,z_k}$. Choose $P^1, P^2 \in \{P,Q,R,S\}$ such that $P^1 \succsim P,Q,R,S \succsim P^2$. Suppose that $P^1 \in \Phi_{2,y_1}$ and $P^2 \in \Phi_{2,y_2}$ with $y_1 \geq y_2$. If $y_1 = y_2$, then $P,Q,R,S \in \Phi_{2,y_1}$ and we are done.

Now suppose that $y_1 > y_2$ and by Lemma 3, we can find $t,t' \in \mathcal{L}^0(X_1)$ with $(t,y_1) \succ P^1 \succ (t',y_1), (t,y_2) \succ P^2 \succ (t',y_2)$. Notice that for each $y \in [y_2,y_1]$, $H(y) := \{P \in \mathcal{P} : (t,y) \succ P \succ (t',y)\} \subseteq \Phi_{2,y}$. By Continuity, $\{P \in \mathcal{P} : P^1 \succsim P \succsim P^2\} \subseteq \cup_{y_2\leq y\leq y_1} H(y)$ and for all $y \in [y_2,y_1]$, there exists $\varepsilon_y > 0$ such that $H(y)\cap H(y') \neq \emptyset$ for all $y' \in [y-\varepsilon_y, y+\varepsilon_y]\cap[y_2,y_1]$. By Finite Cover Theorem, we can find finitely many $z_1 < z_2 < ... < z_K \in [y_2,y_1]$ with $z_1 = y_2, z_K = y_1$ and $[y_2,y_1] \subseteq \cup_{k=1}^K [z_k - \varepsilon_{z_k}, z_k + \varepsilon_{z_k}]$. This implies

$$P,Q,R,S \in \{P \in \mathcal{P} : P^1 \succsim P \succsim P^2\} \subseteq \cup_{y_2\leq y\leq y_1} H(y) = \cup_{k=1}^K H(z_k) \subseteq \cup_{k=1}^K \Phi_{2,z_k}.$$

Then we use induction to show that the four properties stated in the lemma hold for $P,Q,R,S \in \cup_{k=1}^K \Phi_{2,z_k}$. The proof idea is similar to the proof of Lemma 10. By Lemma 20, the four properties hold if $P,Q,R,S \in \Phi_{2,z_1}$. Suppose by induction that they also hold if $P,Q,R,S \in \cup_{k=1}^t \Phi_{2,z_k}$ for some $1 \leq t < K$. By our construction of $\{z_k\}$, $\Phi_{2,z_t} \cap \Phi_{2,z_{t+1}}$ is nonempty and we can find $T^1, T^2 \in \Phi_{2,z_t}\cap\Phi_{2,z_{t+1}}$ with $T^1 \succ T^2$. By Lemma 19 and Lemma 3, since P_1,Q_1,R_1,S_1 have finite supports, we can also construct $p_1,p_2,q_1,q_2 \in \mathcal{L}^0(X_1)$ such that $(p_1,z_{t+1}) \sim (q_1,z_t) \sim T^1$, $(p_2,z_{t+1}) \sim (q_2,z_t) \sim T^2$ and $(p_1,z_{t+1}),(q_1,z_t),(p_2,z_{t+1}),(q_2,z_t)$ are compatible with P,Q,R,S.

For property (i), suppose $P \succsim Q$, P is compatible with Q and $P,Q \in \cup_{k=1}^{t+1}\Phi_{2,z_k}$. If $P \sim Q$, then $P,Q \in \Phi_{2,z_k}$ for some $k = 1,...,t+1$ and hence (i) holds by the inductive hypothesis.

Now we check (ii). If $P \succ Q$, then it suffices to consider the case where $P \in \Phi_{2,z_{t+1}} \backslash (\cup_{k=1}^t \Phi_{2,z_k})$ and $Q \in (\cup_{k=1}^t \Phi_{2,z_k})\backslash\Phi_{2,z_{t+1}}$. This implies $P \succ T^1 \succ T^2 \succ Q$. By Lemma 19, there exist $\lambda_1 \neq \lambda_2 \in (0,1)$ such that $T^1 \sim \lambda_1 P + (1-\lambda_1)Q$ and $T^2 \sim \lambda_2 P + (1-\lambda_2)Q$. Then property (ii) holds for $\lambda = \lambda_1, \lambda_2$.

Notice that at the moment we cannot conclude that $\lambda_1 > \lambda_2$. Suppose that $\lambda_i > \lambda_{-i}$ for some $i = 1,2$. By Lemma 19, we can find $P',Q' \in \hat{\mathcal{P}}$ with $Q' \sim Q, P' \sim P$ and $P,P',Q,Q',(p_1,z_{t+1}),(q_1,z_t),(p_2,z_{t+1}),(q_2,z_t)$ compatible with each other. This guarantees

$$T^1 \sim \lambda_1 P + (1-\lambda_1)Q \sim \lambda_1 P' + (1-\lambda_1)Q \sim \lambda_1 P + (1-\lambda_1)Q' \sim \lambda_1 P' + (1-\lambda_1)Q',$$

$$T^2 \sim \lambda_2 P + (1-\lambda_2)Q \sim \lambda_2 P' + (1-\lambda_2)Q \sim \lambda_2 P + (1-\lambda_2)Q' \sim \lambda_2 P' + (1-\lambda_2)Q'.$$

By property (i), for all $\beta,\beta' \in (0,1)$, $\beta P + (1-\beta)P' \sim P, \beta' Q + (1-\beta')Q' \sim Q$. Apply

Lemma 18 twice and we have for each $\lambda, \beta, \beta' \in (0,1)$

$$\lambda P + (1-\lambda)Q \sim \lambda(\beta P + (1-\beta)P') + (1-\lambda)(\beta' Q + (1-\beta')Q'). \tag{10}$$

For any $\lambda \in (\lambda_{-i}, \lambda_i)$, let $\beta = 1$, $\beta' = \frac{\lambda_i - \lambda}{\lambda_i(1-\lambda)}$, and (10) becomes

$$\begin{aligned}\lambda P + (1-\lambda)Q \sim & \frac{\lambda}{\lambda_i}(\lambda_i P + (1-\lambda_i)Q') + (1-\frac{\lambda}{\lambda_i})Q \\ \sim & \frac{\lambda}{\lambda_i}(q_i, z_t) + (1-\frac{\lambda}{\lambda_i})Q\end{aligned}$$

The second indifference comes from the fact that $\lambda_i P + (1-\lambda_i)Q' \sim T^i \sim (q_i, z_t)$ and Lemma 18. Then by the inductive hypothesis on $\cup_{k=1}^{t}\Phi_{2,z_k}$, we have

$$P \succ (q_i, z_t) \succ \lambda P + (1-\lambda)Q \sim \frac{\lambda}{\lambda_i}(q_i, z_t) + (1-\frac{\lambda}{\lambda_i})Q \succ Q.$$

If $\lambda > \lambda_i$, then let $\beta = \frac{\lambda - \lambda_i}{\lambda(1-\lambda_i)}$, $\beta' = 0$ and (10) becomes

$$\begin{aligned}\lambda P + (1-\lambda)Q \sim & \frac{\lambda - \lambda_i}{1-\lambda_i}P + (1-\frac{\lambda-\lambda_i}{1-\lambda_i})(\lambda_i P' + (1-\lambda_i)Q) \\ \sim & \frac{\lambda - \lambda_i}{1-\lambda_i}P + (1-\frac{\lambda-\lambda_i}{1-\lambda_i})(p_i, z_{t+1})\end{aligned}$$

The second indifference comes from the fact that $\lambda_i P' + (1-\lambda_i)Q \sim T^i \sim (p_i, z_{t+1})$ and Lemma 18. Then by Lemma 20 on $\Phi_{2,z_{t+1}}$, we have

$$P \succ \lambda P + (1-\lambda)Q \sim \frac{\lambda - \lambda_i}{1-\lambda_i}P + (1-\frac{\lambda-\lambda_i}{1-\lambda_i})(p_i, z_{t+1}) \succ (p_i, z_{t+1}) \succ Q.$$

A symmetric argument works for $\lambda < \lambda_{-i}$. Hence property (ii) holds on $\cup_{k=1}^{t+1}\Phi_{2,z_k}$.

We claim that for $P, Q \in \cup_{k=1}^{t+1}\Phi_{2,z_k}$ with $P \succ Q$, P compatible with Q and $1 > \lambda_1 > \lambda_2 > 0$, we have $\lambda_1 P + (1-\lambda_1)Q \succ \lambda_2 P + (1-\lambda_2)Q$. To see this, by (10), we can find $P' \sim P$ where P' is compatible with both P and Q such that

$$\lambda_1 P + (1-\lambda_1)Q \sim \frac{\lambda_1 - \lambda_2}{1-\lambda_2}P' + \frac{1-\lambda_1}{1-\lambda_2}[\lambda_2 P + (1-\lambda_2)Q] \succ \lambda_2 P + (1-\lambda_2)Q$$

The second strict ranking follows from property (ii) since $P \sim P' \succ \lambda_2 P + (1-\lambda_2)Q$.

Given this claim, the proof for (iii) and (iv) on $\cup_{k=1}^{t+1}\Phi_{2,z_k}$ is similar to the proof of (ii).

By induction, the four properties hold for $P, Q, R, S \in \cup_{k=1}^{K}\Phi_{2,z_k}$ and hence arbitrary $P, Q, R, S \in \cup_{y \in X_2}\Phi_{2,y}$. □

It is worthwhile to notice that $\cup_{y \in X_2}\Phi_{2,y}$ is the same as $\mathcal{P}$. The next lemmas shows that they only possibly differ in the worst and the best possible lottery.

Lemma 22. $\mathcal{P}\backslash(\cup_{y\in X_2}\Phi_{2,y}) = \{(\underline{c}_1,\underline{c}_2),(\overline{c}_1,\overline{c}_2)\}$.

Proof of Lemma 22. First, for each $P\in\mathcal{P}$ with $(\overline{c}_1,\overline{c}_2)\succ P\succ(\underline{c}_1,\underline{c}_2)$, there exists $Q,Q'\in\cup_{y\in X_2}\Phi_{2,y}$ with $Q\succ P\succ Q'$, which implies $Q\in\cup_{y\in X_2}\Phi_{2,y}$. Hence

$$\mathcal{P} = (\cup_{y\in X_2}\Phi_{2,y})\cup\{P\in\mathcal{P}: P\sim(\underline{c}_1,\underline{c}_2) \text{ or } P\sim(\overline{c}_1,\overline{c}_2)\}.$$

Then it suffices to show that $P\sim(\underline{c}_1,\underline{c}_2)$ if and only if $P=(\underline{c}_1,\underline{c}_2)$, and $P\sim(\overline{c}_1,\overline{c}_2)$ if and only if $P=(\overline{c}_1,\overline{c}_2)$. This is trivial by Axiom M as for any $P\neq(\underline{c}_1,\underline{c}_2),(\overline{c}_1,\overline{c}_2)$, P dominates $(\underline{c}_1,\underline{c}_2)$ and is dominated by $(\overline{c}_1,\overline{c}_2)$. $\square$

Since $P\succ(\underline{c}_1,\underline{c}_2)$ for any $P\neq(\underline{c}_1,\underline{c}_2)$ and $P\prec(\overline{c}_1,\overline{c}_2)$ for any $P\neq(\overline{c}_1,\overline{c}_2)$, we can easily use the arguments in Lemma 20 to show that the independence property holds for $P,Q,R,S\in\Phi_{2,0}\cup\{(\underline{c}_1,\underline{c}_2)\}$ or $P,Q,R,S\in\Phi_{2,\overline{c}_2}\cup\{(\overline{c}_1,\overline{c}_2)\}$.

Using the same proof as in Lemma 20, we can easily show that the independence property holds for $P,Q,R,S\in\Phi_{2,0}\cup\{(\underline{c}_1,\underline{c}_2)\}$ if $\underline{c}_1,\underline{c}_2>-\infty$ or $P,Q,R,S\in\Phi_{2,\overline{c}_2}\cup\{(\overline{c}_1,\overline{c}_2)\}$ if $\overline{c}_1,\overline{c}_2<+\infty$. Hence we can remove the restriction that $P,Q,R,S\in\cup_{y\in X_2}\Phi_{2,y}$ in Lemma 21.

Corollary 7. *The following properties hold:*

i). $P\sim Q$ and P is compatible with $Q \Longrightarrow \alpha P+(1-\alpha)Q\sim P\sim Q$ for all $\alpha\in(0,1)$;

ii). $P\succ Q$ and P is compatible with $Q \Longrightarrow P\succ\alpha P+(1-\alpha)Q\succ Q$ for all $\alpha\in(0,1)$;

iii). $P\succ Q$, $R\sim S$, P is compatible with R and Q is compatible with $S \Longrightarrow \alpha P+(1-\alpha)R\succ\alpha Q+(1-\alpha)S$ for all $\alpha\in(0,1)$;

iv). $P\sim Q$, $R\sim S$, P is compatible with R and Q is compatible with $S \Longrightarrow \alpha P+(1-\alpha)R\sim\alpha Q+(1-\alpha)S$ for all $\alpha\in(0,1)$.

Finally we slightly relax the requirement of compatibility. For each $P,Q\in\mathcal{P}$, we say P and Q are *weakly compatible* if the they satisfy the following conditions:

- $\mathrm{supp}(P_1)\cap\mathrm{supp}(Q_1)\subseteq\{\underline{c}_1,\overline{c}_1\}$;
- If $\underline{c}_1\in\mathrm{supp}(P_1)\cap\mathrm{supp}(Q_1)$, then $P_{2|\underline{c}_1}=Q_{2|\underline{c}_1}=\underline{c}_2$;
- If $\overline{c}_1\in\mathrm{supp}(P_1)\cap\mathrm{supp}(Q_1)$, then $P_{2|\overline{c}_1}=Q_{2|\overline{c}_1}=\overline{c}_2$.

In other words, if P is weakly compatible with Q, then both P_1 and P_2 attach positive probability to $\underline{c}_1$ (or $\overline{c}_1$) only if the conditional lottery in source 2 given $\underline{c}_1$ (or $\overline{c}_1$) is degenerate at $\underline{c}_2$ (or $\overline{c}_2$) for both P and Q.

Lemma 23. *The following properties hold:*

i). $P\sim Q$ and P is weakly compatible with $Q \Longrightarrow \alpha P+(1-\alpha)Q\sim P\sim Q$ for all $\alpha\in(0,1)$;

ii). $P\succ Q$ and P is weakly compatible with $Q \Longrightarrow P\succ\alpha P+(1-\alpha)Q\succ Q$ for all $\alpha\in(0,1)$;

iii). $P \succ Q$, $R \sim S$, P is weakly compatible with R and Q is weakly compatible with S $\implies \alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$ for all $\alpha \in (0,1)$;

iv). $P \sim Q$, $R \sim S$, P is weakly compatible with R and Q is weakly compatible with S $\implies \alpha P + (1-\alpha)R \sim \alpha Q + (1-\alpha)S$ for all $\alpha \in (0,1)$.

Proof of Lemma 23. Suppose P, Q are weakly compatible but not compatible, that is, $\emptyset \neq \mathrm{supp}(P_1) \cap \mathrm{supp}(Q_1) \subseteq \{\underline{c}_1, \bar{c}_1\}$. We claim that we can find $\tilde{P} \sim P$ and $\tilde{Q} \sim Q$ such that $\tilde{P}$ is compatible with $\tilde{Q}$ and for any $\alpha \in (0,1)$, $\alpha P + (1-\alpha)Q \sim \alpha \tilde{P} + (1-\alpha)\tilde{Q}$, unless $P = Q = (\underline{c}_1, \underline{c}_2)$ or $P = Q = (\bar{c}_1, \bar{c}_2)$.

Case 1: If $\mathrm{supp}(P_1) \cap \mathrm{supp}(Q_1) = \{\underline{c}_1, \bar{c}_1\}$, then $P_{2|\underline{c}_1} = Q_{2|\underline{c}_1} = \underline{c}_2$ and $P_{2|\bar{c}_1} = Q_{2|\bar{c}_1} = \bar{c}_2$. First, we suppose that $P_1(\underline{c}_1) + P_1(\bar{c}_1) < 1$ or $Q_1(\underline{c}_1) + Q_1(\bar{c}_1) < 1$. By symmetry, it suffices to focus on the former case. Denote $P^o = \sum_{x \neq \underline{c}_1, \bar{c}_1} \frac{P_1(x)}{1 - P_1(\underline{c}_1) - P_1(\bar{c}_1)}(x, P_{2|x})$. Then

$$P = P_1(\underline{c}_1)(\underline{c}_1, \underline{c}_2) + P_1(\bar{c}_1)(\bar{c}_1, \bar{c}_2) + (1 - P_1(\underline{c}_1) - P_1(\bar{c}_1))P^o.$$

We know P^o and Q are compatible and $\underline{c}_1, \bar{c}_1 \notin \mathrm{supp}(P_1^o)$. We can similarly define Q^o if $Q_1(\underline{c}_1) + Q_1(\bar{c}_1) < 1$. Otherwise, just choose an arbitrary Q^o so long as $\underline{c}_1, \bar{c}_1 \notin \mathrm{supp}(Q_1^o)$.

By Monotonicity, $(\bar{c}_1, \bar{c}_2) \succ P^o \succ (\underline{c}_1, \underline{c}_2)$. Then we can find $\varepsilon_P > 0$ such that $\bar{c}_1 - \varepsilon_P, \varepsilon_P \notin \mathrm{supp}(P_1) \cup \mathrm{supp}(Q_1)$, $\varepsilon_P \neq \bar{c}_1 - \varepsilon_P$ and

$$(\bar{c}_1 - \varepsilon_P, \bar{c}_2) \succ P^o \succ (\underline{c}_1 + \varepsilon_P, \underline{c}_2).$$

By Lemma 19, we can find $\lambda_P \in (0,1)$ such that

$$P^o \sim \lambda_P(\bar{c}_1 - \varepsilon_P, \bar{c}_2) + (1 - \lambda_P)(\underline{c}_1 + \varepsilon_P, \underline{c}_2) := P^{o\prime}.$$

and $P^{o\prime}$ is compatible with P, Q.

By Corollary 7, we know

$$P' := P_1(\underline{c}_1)(\underline{c}_1, \underline{c}_2) + P_1(\bar{c}_1)(\bar{c}_1, \bar{c}_2) + (1 - P_1(\underline{c}_1) - P_1(\bar{c}_1))P^{o\prime} \sim P.$$

Notice that

$$\begin{aligned} P' =& [P_1(\underline{c}_1)(\underline{c}_1, \underline{c}_2) + (1 - P_1(\underline{c}_1) - P_1(\bar{c}_1))(1 - \lambda_P)(\underline{c}_1 + \varepsilon_P, \underline{c}_2)] \\ &+ [P_1(\bar{c}_1)(\bar{c}_1, \bar{c}_2) + (1 - P_1(\underline{c}_1) - P_1(\bar{c}_1))\lambda_P(\bar{c}_1 - \varepsilon_P, \bar{c}_2)]. \end{aligned}$$

By Lemma 3 given $\underline{c}_2$ or $\bar{c}_2$ in source two, we can then find $\bar{p}, \underline{p} \in \mathcal{L}^0(X_1)$ with $(\bar{p}, \bar{c}_2), (\underline{p}, \underline{c}_2), P, P', Q, Q'$ are pairwise compatible and

$$\begin{aligned} (\bar{p}, \bar{c}_2) &\sim \frac{P_1(\bar{c}_1)(\bar{c}_1, \bar{c}_2) + \lambda_P(1 - P_1(\underline{c}_1) - P_1(\bar{c}_1))(\bar{c}_1 - \varepsilon_P, \bar{c}_2)}{P_1(\bar{c}_1) + \lambda_P(1 - P_1(\underline{c}_1) - P_1(\bar{c}_1))}; \\ (\underline{p}, \underline{c}_2) &\sim \frac{P_1(\underline{c}_1)(\underline{c}_1, \underline{c}_2) + (1 - \lambda_P)(1 - P_1(\underline{c}_1) - P_1(\bar{c}_1))(\underline{c}_1 + \varepsilon_P, \underline{c}_2)}{P_1(\underline{c}_1) + (1 - \lambda_P)(1 - P_1(\underline{c}_1) - P_1(\bar{c}_1))}. \end{aligned}$$

It is important to notice that $\overline{p} \neq \overline{c}_1$ and $\underline{p} \neq \underline{c}_1$.

Again by Corollary 7, we have

$$\tilde{P} :=\lambda_P^*(\underline{p}, \underline{c}_2) + (1-\lambda_P^*)(\overline{p}, \overline{c}_2) \sim P' \sim P,$$

where $\lambda_P^* = P_1(\underline{c}_1) + (1-\lambda_P)(1 - P_1(\underline{c}_1) - P_1(\overline{c}_1))$. Easy to see that $\tilde{P}$ is compatible with P and Q.

We then want to show that for each $\alpha \in (0,1)$, $\alpha P + (1-\alpha)Q \sim \alpha\tilde{P} + (1-\alpha)Q$. To see this, notice that for each $\alpha \in (0,1)$,

$$\begin{aligned}\alpha P + (1-\alpha)Q = &(\alpha P_1(\underline{c}_1) + (1-\alpha)Q_1(\underline{c}_1))(\underline{c}_1, \underline{c}_2) + (\alpha P_1(\overline{c}_1) + (1-\alpha)Q_1(\overline{c}_1))(\overline{c}_1, \overline{c}_2)\\ &+ \alpha(1 - P_1(\underline{c}_1) - P_1(\overline{c}_1))P^o + (1-\alpha)(1 - Q_1(\underline{c}_1) - Q_1(\overline{c}_1))Q^o.\end{aligned}$$

Since $P^o \sim P^{o\prime}$, P^o and $P^{o\prime}$ are compatible with Q and P, Corollary 7 implies that

$$\begin{aligned}\alpha P + (1-\alpha)Q \sim&(\alpha P_1(\underline{c}_1) + (1-\alpha)Q_1(\underline{c}_1))(\underline{c}_1, \underline{c}_2) + (\alpha P_1(\overline{c}_1) + (1-\alpha)Q_1(\overline{c}_1))(\overline{c}_1, \overline{c}_2)\\ &+ \alpha(1 - P_1(\underline{c}_1) - P_1(\overline{c}_1))P^{o\prime} + (1-\alpha)(1 - Q_1(\underline{c}_1) - Q_1(\overline{c}_1))Q^o\\ =&\alpha P' + (1-\alpha)Q\\ =&\alpha[P_1(\underline{c}_1)(\underline{c}_1, \underline{c}_2) + (1 - P_1(\underline{c}_1) - P_1(\overline{c}_1))(1-\lambda_P)(\underline{c}_1 + \varepsilon_P, \underline{c}_2)]\\ &+ (1-\alpha)Q_1(\underline{c}_1)(\underline{c}_1, \underline{c}_2)\\ &+ \alpha[P_1(\overline{c}_1)(\overline{c}_1, \overline{c}_2) + (1 - P_1(\underline{c}_1) - P_1(\overline{c}_1))\lambda_P(\overline{c}_1 - \varepsilon_P, \overline{c}_2)]\\ &+ (1-\alpha)Q_1(\overline{c}_1)(\overline{c}_1, \overline{c}_2)\\ &+ (1-\alpha)(1 - Q_1(\underline{c}_1) - Q_1(\overline{c}_1))Q^o.\end{aligned}$$

Notice that the first two terms in the last equation have $\underline{c}_2$ in source 2, while the third and four term have $\overline{c}_2$ in the source 2. Apply Lemma 3 given $\underline{c}_2$ or $\overline{c}_2$ in source 2 and Corollary 7 sequentially, we can derive

$$\begin{aligned}\alpha P + (1-\alpha)Q \sim&\alpha\lambda_P^*(\underline{p}, \underline{c}_2) + (1-\alpha)Q_1(\underline{c}_1)(\underline{c}_1, \underline{c}_2)\\ &+ \alpha(1-\lambda_P^*)(\overline{p}, \overline{c}_2) + (1-\alpha)Q_1(\overline{c}_1)(\overline{c}_1, \overline{c}_2)\\ &+ (1-\alpha)(1 - Q_1(\underline{c}_1) - Q_1(\overline{c}_1))Q^o\\ =&\alpha\tilde{P} + (1-\alpha)Q.\end{aligned}$$

Now we suppose $P_1(\underline{c}_1) + P_1(\overline{c}_1) = Q_1(\underline{c}_1) + Q_1(\overline{c}_1) = 1$. As $\text{supp}(P_1) \cap \text{supp}(Q_1) = \{\underline{c}_1, \overline{c}_1\}$, $P_1(\underline{c}_1), Q_1(\underline{c}_1) \in (0,1)$. By Monotonicity, $(\overline{c}_1, \overline{c}_2) \succ P, Q \succ (\underline{c}_1, \underline{c}_2)$. Then we can find $\tilde{P} \sim P$ such that $\underline{c}_1, \overline{c}_1 \notin \text{supp}(\tilde{P}_1)$ and $\tilde{P}$ is compatible with P and Q. For any $\beta \in (0,1)$, by Corollary 7, $P \sim \beta P + (1-\beta)P' := P^\beta$. Clearly, $P^\beta(\underline{c}_1) + P^\beta(\overline{c}_1) < 1$. We can apply the previous result for P^β and Q, that is, for each β, we can find $\tilde{P}^\beta \sim P^\beta \sim P$ with $\tilde{P}^\beta$ compatible with Q and for each $\alpha \in (0,1)$, $\alpha P^\beta + (1-\alpha)Q \sim \alpha\tilde{P}^\beta + (1-\alpha)Q$. Again by Corollary 7, we can actually choose $\tilde{P}^\beta$ to be the same across all $\beta \in (0,1)$. Denote it as $\tilde{P}$.

Hence, for each $\beta, \alpha \in (0,1)$,

$$\alpha\beta P + \alpha(1-\beta)P' + (1-\alpha)Q \sim \alpha\tilde{P} + (1-\alpha)Q$$

By Continuity 1, let $\beta \to 1$ and we have for each $\alpha \in (0,1)$,

$$\alpha P + (1-\alpha)Q \sim \alpha\tilde{P} + (1-\alpha)Q.$$

Case 2: If $\mathrm{supp}(P_1) \cap \mathrm{supp}(Q_1) = \{\underline{c}_1\}$ or $\{\bar{c}_1\}$, then similar arguments in Case 1 work.

As a summary, we know that either $P = Q = (\underline{c}_1, \underline{c}_2)$ or $P = Q = (\bar{c}_1, \bar{c}_2)$ or there exist $\tilde{P} \sim P$ and $\tilde{Q} \sim Q$ such that $\tilde{P}$ is compatible with $\tilde{Q}$ and for any $\alpha \in (0,1)$, $\alpha P + (1-\alpha)Q \sim \alpha\tilde{P} + (1-\alpha)\tilde{Q}$. Now we are ready to prove the four properties.

For (i) and (ii), if $P = Q = (\underline{c}_1, \underline{c}_2)$ or $P = Q = (\bar{c}_1, \bar{c}_2)$, then the result is trivial. Otherwise, there exist $\tilde{P} \sim P$ and $\tilde{Q} \sim Q$ such that for any $\alpha \in (0,1)$,

$$\begin{aligned} &P \sim Q \Longrightarrow P' \sim Q' \Longrightarrow \alpha P + (1-\alpha)Q \sim \alpha P' + (1-\alpha)Q' \sim P' \sim P, \\ &P \succ Q \Longrightarrow P' \succ Q' \Longrightarrow P \sim P' \succ \alpha P + (1-\alpha)Q \sim \alpha P' + (1-\alpha)Q' \succ Q' \sim Q. \end{aligned}$$

For (iii) and (iv), if $P = R = (\underline{c}_1, \underline{c}_2)$ or $Q = S = (\underline{c}_1, \underline{c}_2)$ or $P = R = (\bar{c}_1, \bar{c}_2)$ or $Q = S = (\bar{c}_1, \bar{c}_2)$, then by Monotonicity, the primitives of (iii) or (iv) hold only if $P = Q = R = S$, or $P = R = S = (\bar{c}_1, \bar{c}_2) \succ Q$ or $P \succ R = S = Q = (\underline{c}_1, \underline{c}_2)$, in which cases the result holds trivially. By excluding those cases, we can construct $\tilde{P} \sim P$, $\tilde{Q} \sim Q$, $\tilde{R} \sim R$ and $\tilde{S} \sim S$ such that $\tilde{P}$ is compatible with $\tilde{R}$, $\tilde{Q}$ is compatible with $\tilde{S}$ and for any $\alpha \in (0,1)$, $\alpha P + (1-\alpha)R \sim \alpha\tilde{P} + (1-\alpha)\tilde{R}$, $\alpha Q + (1-\alpha)S \sim \alpha\tilde{Q} + (1-\alpha)\tilde{S}$. By Corollary 7,

$$P \sim Q, R \sim S \Longrightarrow \alpha P + (1-\alpha)R \sim \alpha\tilde{P} + (1-\alpha)\tilde{R} \sim \alpha\tilde{Q} + (1-\alpha)\tilde{S} \sim \alpha Q + (1-\alpha)S,$$

$$P \succ Q, R \sim S \Longrightarrow \alpha P + (1-\alpha)R \sim \alpha\tilde{P} + (1-\alpha)\tilde{R} \succ \alpha\tilde{Q} + (1-\alpha)\tilde{S} \sim \alpha Q + (1-\alpha)S.$$

This completes the proof. □

Now we can state the main intermediate result in this proof.

Lemma 24. *$\succsim$ is represented by U where for each $P \in \mathcal{P}$,*

$$U(P) = \sum_x \hat{u}_D(x, CE_{v_x}(P_{2|x}))P_1(x) \tag{11}$$

with continuous and strictly monotonic $\hat{u}_D$ and v_x for each $x \in X_1$. $\hat{u}_D$ and v_x for each $x \in X_1$ are unique up to a positive affine transformation.

The utility function (11) differs from V^{ANB} since the preference for conditional lotteries in source 2 is represented by v_x, which can depend on the outcome x in source 1. If v_x is independent of x, then (11) agrees with V^{ANB}. If v_x is a positive affine transformation of $\hat{u}_D(x, \cdot)$, then (11) agrees with V^{EU}.

Proof of Lemma 24. The proof can be decomposed into the following steps. Firstly, we construct a mixture space $\mathcal{D}^*$ which is richer than $\mathcal{P}$ and argue that $\mathcal{P}$ can be mapped bijectively to a subset of the $\mathcal{D}^*$. Secondly, we construct an extension of $\succsim$ from $\mathcal{P}$ to $\succsim^*$ the richer domain $\mathcal{D}^*$. Thirdly, we prove that $\succsim^*$ satisfies the standard EU axioms on $\mathcal{D}^*$ and hence is represented by some EU utility function. Finally, we show that the restriction of the utility function on $\mathcal{P}$ can be written as (11).

Step 1. Denote $\mathcal{D}^* := \mathcal{L}^0(X_1 \times \mathcal{L}^0(X_2))$. $\mathcal{D}^*$ is a mixture space over $X_1 \times \mathcal{L}^0(X_2)$. Following Kreps and Porteus (1978), we call each element $d \in \mathcal{D}^*$ a *temporal lottery*, which involves temporal resolution of uncertainty. We also define

$$\hat{\mathcal{D}}^* := \Big\{ d \in \mathcal{D}^* : d(x,p)d(x,p') = 0, \forall x \in X_1, p \neq p' \in \mathcal{L}^0(X_2) \Big\}.$$

$\hat{\mathcal{D}}^*$ is the set of temporal lotteries where there is no early resolution of uncertainty.

Note that for $i = 1, 2$, X_i with the standard relative topology in $\mathbb{R}$ is separable. By Kreps and Porteus (1978), $\mathcal{P}$ and $\mathcal{D}^*$ with weak topology can be metrizable by the Prokhorov metric. Endow $\hat{\mathcal{P}}$ with the relative topology with respect to the weak topology on $\mathcal{P}$ and $\hat{\mathcal{D}}^*$ with the relative topology with respect to the weak topology on $\mathcal{D}^*$.

Define a mapping $f : \mathcal{P} \to \hat{\mathcal{D}}^*$ as follows: for $P \in \mathcal{P}$, denote $f[P] = d \in \mathcal{D}^*$ such that for any $(x,q) \in X_1 \times \mathcal{L}^0(X_2)$, $f[P](x,q) = P_1(x)$ if $q = P_{2|x}$ and $f[P](x,q) = 0$ if $q \neq P_{2|x}$. Clearly $f[P] \in \hat{\mathcal{D}}^*$ and f is well-defined. Inversely, $f^{-1} : \hat{\mathcal{D}}^* \to \mathcal{P}$ such that $f^{-1}[d](x,y) = \sum_{q \in \mathcal{L}^0(X_2)} d(x,q)q(y)$. This is also well-defined since for each $x \in X_1$ there exists at most one $q \in \mathcal{L}^0(X_2)$ with $d(x,q) > 0$ for any $d \in \hat{\mathcal{D}}^*$. Thus, f is a bijective mapping between $\mathcal{P}$ and $\hat{\mathcal{D}}^*$. It is worth noting that f is not a homeomorphism as f is not continuous, although f^{-1} is continuous.

We define a binary relation $\succsim'$ on $\hat{\mathcal{D}}^*$ such that $d \succsim' d'$ if and only if $f^{-1}[d] \succsim f^{-1}(d')$. $\succ'$ and $\sim'$ are defined correspondingly. The next result directly follows from Lemma 23.

Corollary 8. *Suppose $\lambda_i > 0$ for all i and $\sum_{i=1}^n \lambda_i = 1$. If $d^1 = \sum_{i=1}^n \lambda_i \delta_{(x_i,p_i)}$, $d^2 = \sum_{i=1}^n \lambda_i \delta_{(y_i,q_i)}$ with $x_i \neq x_j, y_i \neq y_j$ for all $i \neq j$ and $(x_i,p_i) \sim (y_i,q_i)$ for all i, then $d^1 \sim' d^2$.*

Proof of Corollary 8. By Lemma 23, as $x_i \neq x_j, y_i \neq y_j$ for all $i \neq j$ and $(x_i,p_i) \sim (y_i,q_i)$ for all i, we have

$$\frac{\lambda_1}{\lambda_1+\lambda_2}(x_1,p_1) + \frac{\lambda_2}{\lambda_1+\lambda_2}(x_2,p_2) \sim \frac{\lambda_1}{\lambda_1+\lambda_2}(y_1,q_1) + \frac{\lambda_2}{\lambda_1+\lambda_2}(y_2,q_2).$$

Then by induction, we can get

$$\sum_{i=1}^n \lambda_i(x_i,p_i) \sim \sum_{i=1}^n \lambda_i(y_i,q_i).$$

Since $d^1 = f^{-1}(\sum_{i=1}^n \lambda_i(x_i,p_i))$, $d^2 = f^{-1}(\sum_{i=1}^n \lambda_i(y_i,q_i))$, we conclude that $d^1 \sim' d^2$. $\square$

Step 2. Then we extend $\succsim'$ on $\hat{\mathcal{D}}^*$ to $\succsim^*$ on $\mathcal{D}^*$.

For any $d \in \mathcal{D}^* \backslash \hat{\mathcal{D}}^*$, denote by $d_1 \in \mathcal{L}^0(X_1)$ the marginal lottery in source 1 and by $d_{2|x} \in \mathcal{L}^0(\mathcal{L}^0(X_2))$ the lottery over marginal lotteries conditional on outcome x in source 1. Denote $\mathrm{supp}(d_1) = \{x_1, ..., x_N\} \subset X_1$ with $x_1 < \cdots < x_N$ and for each $k = 1, ..., N$, $\mathrm{supp}(d_{2|x_k}) = \{p_{k,1}, ..., p_{k,t_k}\} \subset \mathcal{L}^0(X_2)$ with $t_k \geq 1$. Since $d \notin \hat{\mathcal{D}}^*$, there exists some k' with $t_{k'} > 1$. We will map d to some temporal lottery in $\hat{\mathcal{D}}^*$ using the following algorithm.

Stage 1. We first focus on outcome x_1 in source 1. Denote $z_{1,1} = x_1$. If $t_1 = 1$, by Lemma 3, we can find $\hat{z}_{1,1}$ with $(z_{1,1}, \hat{z}_{1,1}) \sim (z_{1,1}, p_{1,1})$. Define $d^1 \in \mathcal{D}^*$ such that $d^1(z_{1,1}, \hat{z}_{1,1}) = d(z_{1,1}, p_{1,1})$, $d^1(z_{1,1}, q) = 0$ for all $q \neq \hat{z}_{1,1}$ and $d^1(x, q) = d(x, q)$ for all $x \neq z_{1,1}$ and $q \in \mathcal{L}^0(X_2)$.

If $t_1 > 1$ and $d(\underline{c}_1, \underline{c}_2) > 0$, that is, $x_1 = \underline{c}_1$ and $\underline{c}_2 \in \mathrm{supp}(d_{2|\underline{c}_1})$, then we reorder lotteries in $\mathrm{supp}(d_{2|\underline{c}_1})$ such that $(x_1, p_{1,1}) \precsim (x_1, p_{1,2}) \precsim \cdots \precsim (x_1, p_{1,t_1})$. By Monotonicity, we know $(x_1, p_{1,1}) = (\underline{c}_1, \underline{c}_2) \prec (x_1, p_{1,i})$ for each $i > 1$. By Continuity 3, we can find $\overline{z}_1 \in (x_1, \frac{x_1+x_2}{2})$ with $(x_1, p_{1,2}) \succ (\overline{z}_1, \underline{c}_2)$. Define $z_{1,j} = z_{1,1} + \frac{j-1}{t_1}(\overline{z}_1 - z_{1,1})$ for all $j = 1, ..., t_1$. Clearly, $z_{1,1} = x_1 < z_{1,2} < \cdots < z_{1,t_1} < \overline{z}_1$. By Monotonicity and Lemma 3, we can find $\hat{z}_{1,j}$ for $j \geq 1$ with $(z_{1,j}, \hat{z}_{1,j}) \sim (x_1, p_{1,j})$. Then we define $d^1 \in \mathcal{D}^*$ such that for any $x \notin \{z_{1,j}\}_{j=1}^{t_1}$, $q \in \mathcal{L}^0(X_2)$, $d^1(x, q) = d(x, q)$, and for each $j = 1, ..., t_1$, $d^1(z_{1,j}, \hat{z}_{1,j}) = d(x_1, p_{1,j})$ and $d^1(z_{1,j}, q) = 0$ for $q \neq \hat{z}_{1,j}$.

If $t_1 > 1$ and $d(\underline{c}_1, \underline{c}_2) = 0$, then we can construct d^1 similarly by choosing $z_{1,1} > \underline{c}_1$.

Stage $i \geq 2$. Suppose that we have constructed d^{i-1}. Now consider x_i in source 1 with $x_i > x_{i-1} \geq x_1 \geq \underline{c}_1$. Denote $z_{i,1} = x_i$. If $t_i = 1$, then by Lemma 3, there exists $\hat{z}_{i,1}$ with $(z_{i,1}, \hat{z}_{i,1}) \sim (z_{i,1}, p_{i,1})$. Define $d^i \in \mathcal{D}^*$ such that $d^i(z_{i,1}, \hat{z}_{i,1}) = d(z_{i,1}, p_{i,1})$, $d^i(z_{i,1}, q) = 0$ for all $q \neq \hat{z}_{i,1}$ and $d^i(x, q) = d^{i-1}(x, q)$ for all $x \neq z_{i,1}$ and $q \in \mathcal{L}^0(X_2)$.

If instead $t_i > 1$, again we can reorder the marginal lotteries such that $(x_i, p_{i,1}) \precsim \cdots \precsim (x_i, p_{i,t_i})$. As $x_i > x_1 \geq \underline{c}_1$, $(x_i, p_{i,1}) \succ (\frac{x_i+x_{i-1}}{2}, \underline{c}_2) \succ (x_1, \underline{c}_2)$ by Monotonicity. Since $p_{i,1} \neq p_{i,2}$, $(x_i, p_{i,2}) \succ (x_i, \underline{c}_2)$. By Continuity 3, we can find $\overline{z}_i \in (x_i, \frac{x_i+x_{i+1}}{2})$ with $(x_i, p_{i,2}) \succ (\overline{z}_i, \underline{c}_2)$. Define $z_{i,j} = z_{i,1} + \frac{j-1}{t_i}(\overline{z}_i - z_{i,1})$ for all $j = 1, ..., t_i$. Clearly, $z_{i,1} = x_i < z_{i,2} < \cdots < z_{i,t_i} < \overline{z}_i$. By Monotonicity and Lemma 3, we can find $\hat{z}_{i,j}$ for $j \geq 1$ with $(z_{i,j}, \hat{z}_{i,j}) \sim (x_i, p_{i,j})$. Then we define $d^i \in \mathcal{D}^*$ such that for any $x \notin \{z_{i,j}\}_{j=1}^{t_i}$, $q \in \mathcal{L}^0(X_2)$, $d^i(x, q) = d^{i-1}(x, q)$, and for each $j = 1, ..., t_i$, $d^1(z_{i,j}, \hat{z}_{i,j}) = d(x_i, p_{i,j})$ and $d^i(z_{i,j}, q) = 0$ for $q \neq \hat{z}_{i,j}$.

The algorithm terminates at $i = N < +\infty$. We know that $\mathrm{supp}(d_1^N) = \bigcup_{k=1}^N \{z_{k,1}, \cdots, z_{k,t_k}\}$. For each $z_{k,i} \in \mathrm{supp}(d_1^N)$ with $1 \leq i \leq t_k$ and $1 \leq k \leq N$, there exists $\hat{z}_{k,i}$ with $(z_{k,i}, \hat{z}_{k,i}) \sim (x_k, p_{k,i})$ and $d^N(z_{k,i}, \hat{z}_{k,i}) = d(x_k, p_{k,i})$. Moreover, $\{z_{k,i}\}$ admits a lexicographic order, that is, $z_{k,i} < z_{k',i'}$ if $k < k'$ or $k = k'$ and $i < i'$. This implies $d^N \in \hat{\mathcal{D}}^*$.

In this way, we constructed a mapping $h : \mathcal{D}^* \backslash \hat{\mathcal{D}}^* \to \hat{\mathcal{D}}^*$ where $h(d) = d^N$. Actually, the above algorithm also applies for $d \in \hat{\mathcal{D}}^*$ and hence we can define h on $\mathcal{D}^*$. Clearly $h(h(d)) = h(d)$ for all $d \in \mathcal{D}^*$. Then we define $\succsim^*$ on $\mathcal{D}^*$ such that $\succsim^*$ agrees with $\succsim'$ on $\hat{\mathcal{D}}^*$ and $d \sim^* h(d)$ for $d \in \mathcal{D}^* \backslash \hat{\mathcal{D}}^*$, that is, for each $d, d' \in \mathcal{D}^*$, $d \succsim^* d'$ if and only if $h(d) \succsim' h(d')$.

To verify that $\succsim^*$ is well-defined, we need to argue that the particular choice of $\{z_{k,i}\}$ does not affect the definition of $\succsim^*$. Consider two constructions h and $\hat{h}$. For each $d = \sum_{k,i} d(x_k, p_{k,i}) \delta_{(x_k, p_{k,i})}$, $h(d) = \sum_{k,i} d(x_k, p_{k,i}) \delta_{(z_{k,i}, \hat{z}_{k,i})}$ and $\hat{h}(d) = \sum_{k,i} d(x_k, p_{k,i}) \delta_{(z'_{k,i}, \hat{z}'_{k,i})}$ such that $z_{k,i} \neq z_{k',i'}$, $z'_{k,i} \neq z'_{k',i'}$ for all $(k,i) \neq (k',i')$ and for all k, i, $(z'_{k,i}, \hat{z}'_{k,i}) \sim (z_{k,i}, \hat{z}_{k,i}) \sim (x_k, p_{k,i})$. By Corollary 8 and the definition of $\succsim'$, $h(d) \sim' \hat{h}(d)$. Hence the definition of $\succsim^*$ is not affected by the specific construction of h. As a direct corollary, $h(d) \sim' d$ for each $d \in \hat{\mathcal{D}}^*$, although it is possible that $h(d) \neq d$.

Step 3. For $d, d' \in \mathcal{D}^* = \mathcal{L}^0(X_1 \times \mathcal{L}^0(X_2))$ and $\alpha \in (0,1)$, we define the α-mixture of d and d' as

$$[\alpha d + (1-\alpha)d'](x,q) = \alpha d(x,q) + (1-\alpha)d'(x,q), \forall\ \alpha \in (0,1), (x,q) \in X_1 \times \mathcal{L}^0(X_2).$$

Then $\mathcal{D}^*$ is a mixture space. We want to show that $\succsim^*$ satisfies the vNM independence axiom and mixture continuity on $\mathcal{D}^*$. We start with an intermediate result for the vNM independence axiom by extending Corollary 8 to $\mathcal{D}^*$.

Lemma 25. *Suppose $\lambda_i > 0$ for all i and $\sum_{i=1}^n \lambda_i = 1$. For $d^1 = \sum_{i=1}^n \lambda_i \delta_{(x_i,p_i)}$, $d^2 = \sum_{i=1}^n \lambda_i \delta_{(y_i,q_i)}$ with $(x_i,p_i) \sim (y_i,q_i)$ for all i, then $d^1 \sim^* d^2$.*

Proof of Lemma 25. We first assume that $(x_i,p_i) \neq (x_j,p_j)$ and $(y_i,q_i) \neq (y_j,q_j)$ for all $i \neq j$. By definition of h, $h(d^1) = \sum_{i=1}^n \lambda_i \delta_{(z_i^1,\hat{z}_i^1)}$ with $z_i^1 \neq z_j^1$ for $i \neq j$ and $(z_i^1, \hat{z}_i^1) \sim (x_i,p_i)$ for each $i = 1,...,n$. Similarly, $h(d^2) = \sum_{i=1}^n \lambda_i \delta_{(z_i^2,\hat{z}_i^2)}$ with $z_i^2 \neq z_j^2$ for $i \neq j$ and $(z_i^2, \hat{z}_i^2) \sim (x_i,p_i)$ for each $i = 1,...,n$. We know that $h(d^1) \sim^* d^1$ and $h(d^2) \sim^* d^2$. By Corollary 8, $h(d^1) \sim^* h(d^2)$ and hence $d^1 \sim^* d^2$.

Now we allow for $(x_i,p_i) = (x_j,p_j)$ or $(y_i,q_i) = (y_j,q_j)$ with some $i \neq j$. If $n = 1$, then the result is trivial. Suppose $n \geq 2$. Relabel the subscripts so that $(x_i,p_i) = (\underline{c}_1, \underline{c}_2)$ for $i \leq k$ for some $k \geq 0$ and $(x_i,p_i) \succ (\underline{c}_1,\underline{c}_2)$ for $i > k$. Since $(x_i,p_i) \sim (y_i,q_i)$, we know that $(y_i,q_i) = (\underline{c}_1,\underline{c}_2)$ for $i \leq k$ and $(y_i,q_i) \succ (\underline{c}_1,\underline{c}_2)$ for $i > k$. Then

$$d^1 = [\sum_{i=1}^k \lambda_i]\delta_{(\underline{c}_1,\underline{c}_2)} + \sum_{i=k+1}^n \lambda_i \delta_{(x_i,p_i)}\ ,\ d^2 = [\sum_{i=1}^k \lambda_i]\delta_{(\underline{c}_1,\underline{c}_2)} + \sum_{i=k+1}^n \lambda_i \delta_{(y_i,q_i)}.$$

This implies that we can assume $(x_i,p_i) \neq (\underline{c}_1,\underline{c}_2)$ for all $i \geq 2$. By a similar argument, we can assume $(x_i,p_i) \neq (\bar{c}_1,\bar{c}_2)$ for all $i < n$.

Without loss of generality, we can further assume $\bar{c}_1 > x_i > \underline{c}_1$ for all $2 \leq i \leq n-1$ since we can construct $\hat{d}^1 \sim d^1$ by replacing $(\underline{c}_1,p_i) \neq (\underline{c}_1,\underline{c}_2)$ with (a,q) and $(\bar{c}_1,p_i) \neq (\bar{c}_1,\bar{c}_2)$ with (b,q') for some $a,b \in (\underline{c}_1,\bar{c}_1)$.

i). Suppose that $(x_1,p_1) \succ (\underline{c}_1,\underline{c}_2)$ and $(x_n,p_n) \prec (\bar{c}_1,\bar{c}_2)$. By relabeling, there exists a partition of $\{1,...,n\}$ as $\Big\{\{1,...,t_1\},\cdots,\{t_{k-1}+1,...,n\}\Big\}$ such that $(x_i,p_i) = (x_j,p_j)$ for all $t_l + 1 \leq i,j \leq t_{l+1}$ with $0 \leq l \leq k-1$ and $t_0 = 0, t_k = n$. For $l = 0$, that is, $1 \leq i \leq t_1$, by Continuity and Lemma 3, we can find $z_i > \underline{c}_1$, $\hat{z}_i \geq \underline{c}_2$ with $(z_i,\hat{z}_i) \sim (x_i,p_i) = (x_1,p_1)$ for all $i = 1,...,t_1$ and $z_i \neq z_j$ for all $i \neq j$. By applying Lemma 23 repeatedly, we derive

$$\sum_{i=1}^{t_1} \frac{\lambda_i}{\sum_{j=1}^{t_1} \lambda_j}(z_i,\hat{z}_i) \sim (x_1,p_1).$$

The same result holds for $l = 1, ..., k-1$.

Recall that $d^1 = \sum_{i=1}^n \lambda_i \delta_{(x_i,p_i)} = \sum_{l=0}^{k-1}(\sum_{i=t_l+1}^{t_{l+1}} \lambda_i)\delta_{(x_i,p_i)}$. By definition of h, we can find $h(d^1) \sim^* d^1$ with $h(d^1) \in \hat{\mathcal{D}}^*$. Denote $h(d^1) = \sum_{l=0}^{k-1} \hat{\lambda}_{l+1}\delta_{(x'_{l+1},\hat{x}'_{l+1})}$, where $\hat{\lambda}_{l+1} = \sum_{i=t_l+1}^{t_{l+1}} \lambda_i$ and $(x'_{l+1}, \hat{x}'_{l+1}) \sim (x_{t_l+1}, p_{t_l+1}) \sim \sum_{i=t_l+1}^{t_{l+1}} \frac{\lambda_i}{\sum_{j=t_l+1}^{t_{l+1}} \lambda_j}(z_i, \hat{z}_i)$ for each l. Denote $R_{l+1} = \sum_{i=t_l+1}^{t_{l+1}} \frac{\lambda_i}{\sum_{j=t_l+1}^{t_{l+1}} \lambda_j}(z_i, \hat{z}_i)$. Note that R_l and $R_{l'}$ are compatible, $(x'_l, \hat{x}'_l)$ and $(x'_{l'}, \hat{x}'_{l'})$ are compatible for all $l \neq l'$. By Lemma 23 and the definition of $\succsim'$,

$$d^1 \sim^* h(d^1) = \sum_{l=1}^{k} \hat{\lambda}_l \delta_{(x'_l,\hat{x}'_l)} \sim' \sum_{l=1}^{k} h(\hat{\lambda}_l R_l) = \sum_{i=1}^{n} \lambda_i \delta_{(z_i,\hat{z}_i)} \in \hat{\mathcal{D}}^*.$$

Similarly, we can find $z'_i, \hat{z}'_i$ for $i = 1, ..., n$ such that $z'_i \neq z'_j$ for all $i \neq j$, $(z'_i, \hat{z}'_i) \sim (y_i, q_i)$ for each i and

$$d^2 \sim^* \sum_{i=1}^{n} \lambda_i \delta_{(z'_i,\hat{z}'_i)} \in \hat{\mathcal{D}}^*.$$

By Corollary 8, we have

$$d^2 \sim^* \sum_{i=1}^{n} \lambda_i \delta_{(z'_i,\hat{z}'_i)} \sim^* \sum_{i=1}^{n} \lambda_i \delta_{(z_i,\hat{z}_i)} \sim^* d^1.$$

(ii). Now we turn to the case where $(x_1, p_1) = (\underline{c}_1, \underline{c}_2)$ or $(x_n, p_n) = (\bar{c}_1, \bar{c}_2)$ or both. This implies $(y_1, q_1) = (\underline{c}_1, \underline{c}_2)$ or $(y_n, q_n) = (\bar{c}_1, \bar{c}_2)$ or both. If $\lambda_1 + \lambda_n = 1$, then the result is trivial as $n = 2$ and we are back to the special case where $(x_i, p_i) \neq (x_j, p_j), (y_i, q_i) \neq (y_j, q_j)$ for all $i \neq j$. By Monotonicity, $(\bar{c}_1, \bar{c}_2) \succ (x_i, p_i) \succ (\underline{c}_1, \underline{c}_2), (\bar{c}_1, \bar{c}_2) \succ (y_i, q_i) \succ (\underline{c}_1, \underline{c}_2)$ for all $2 \leq i \leq n-2$. Define

$$\hat{d}^1 = \frac{1}{1 - \lambda_1 - \lambda_n} \sum_{i=2}^{n} \lambda_i \delta_{(x_i,p_i)} \ , \ \hat{d}^2 = \frac{1}{1 - \lambda_1 - \lambda_n} \sum_{i=2}^{n} \lambda_i \delta_{(y_i,q_i)}.$$

Then we are back to case (i) and $\hat{d}^1 \sim^* \hat{d}^2$. By Lemma 23 and the definition of $\succsim'$, $h(d^1) = \lambda_1 \delta_{(0,0)} + \lambda_n \delta_{(\bar{c}_1,\bar{c}_2)} + (1 - \lambda_1 - \lambda_n)h(\hat{d}^1) \sim' \lambda_1 \delta_{(0,0)} + \lambda_n \delta_{(\bar{c}_1,\bar{c}_2)} + (1 - \lambda_1 - \lambda_n)h(\hat{d}^2) = h(d^2)$. Thus, by definition of $\succsim^*$, we conclude that $d^1 \sim^* d^2$. $\square$

We claim that $\succsim^*$ satisfies the vNM independence axiom, that is, for $d, d', d'' \in \mathcal{D}^*$ and $\alpha \in (0, 1)$,

$$d \succ^* d' \Longrightarrow \alpha d + (1 - \alpha)d'' \succ^* \alpha d' + (1 - \alpha)d''.$$

To see this, fix $d, d', d'' \in \mathcal{D}^*$ with $d \succ^* d'$ and $\alpha \in (0, 1)$. Denote $d = \sum_{i=1}^n \lambda_i \delta_{(x_i,p_i)}$ and $d'' = \sum_{j=1}^m \eta_j \delta_{(y_j,q_j)}$. As $\text{supp}(d_1) \cup \text{supp}(d''_1)$ is finite, for any i such that $(x_i, p_i) \neq (\underline{c}_1, \underline{c}_2)$ and $(\bar{c}_1, \bar{c}_2)$, we can find $(z_i^d, \hat{z}_i^d) \sim (x_i, p_i)$; for any j such that $(y_j, q_j) \neq (\underline{c}_1, \underline{c}_2)$ and $(\bar{c}_1, \bar{c}_2)$, we can find $(z_j^{d''}, \hat{z}_j^{d''}) \sim (y_j, q_j)$. Moreover, we can require that $\bar{c}_1 > z_i^d, z_j^{d''} > \underline{c}_1$, $z_i^d \neq z_j^d, z_i^{d''} \neq z_j^{d''}$ for all $i \neq j$ and $z_i^d \neq z_j^{d''}$ for all i, j. For any i, j with $(x_i, p_i) = (\underline{c}_1, \underline{c}_2)$ or $(y_j, q_j) = (\underline{c}_1, \underline{c}_2)$, denote $z_i^d = z_j^{d''} = \underline{c}_1$ and $\hat{z}_i^d = \hat{z}_j^{d''} = \underline{c}_2$. For any i, j with $(x_i, p_i) = (\bar{c}_1, \bar{c}_2)$

or $(y_j, q_j) = (\bar{c}_1, \bar{c}_2)$, denote $z_i^d = z_j^{d''} = \bar{c}_1$ and $\hat{z}_i^d = \hat{z}_j^{d''} = \bar{c}_2$. Then Lemma 25 assures that

$$\alpha d + (1-\alpha)d'' \sim^* \alpha\hat{d} + (1-\alpha)\hat{d}'',$$

where $\hat{d} = \sum_{i=1}^n \lambda_i \delta_{(z_i^d, \hat{z}_i^d)} \sim^* d$ and $\hat{d}'' = \sum_{j=1}^m \eta_j \delta_{(z_i^{d''}, \hat{z}_i^{d''})} \sim^* d''$. As $\hat{d}, \hat{d}'' \in \hat{\mathcal{D}}^*$, we can denote $P = f^{-1}[\hat{d}]$ and $R = f^{-1}[\hat{d}'']$. Easy to see that P and R are weakly compatible, and $\alpha d + (1-\alpha)d'' \sim^* \alpha f(P) + (1-\alpha)f(R)$ for any $\alpha \in (0,1)$.

Similarly we can construct Q and S such that $d' \sim^* f(Q)$, $d'' \sim^* f(S)$, Q and S are weakly compatible and $\alpha d' + (1-\alpha)d'' \sim^* \alpha f(Q) + (1-\alpha)f(S)$ for any $\alpha \in (0,1)$. Note that $d \succ^* d'$ if and only if $P \succ Q$ and $d'' = d''$ implies that $R \sim S$. By Lemma 23, $\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S$ for any $\alpha \in (0,1)$. It is easy to verify that $f(\alpha P + (1-\alpha R)) = \alpha f(P) + (1-\alpha)f(R)$ and $f(\alpha Q + (1-\alpha S)) = \alpha f(Q) + (1-\alpha)f(S)$ since P and R are weakly compatible and Q and S are weakly compatible. Thus for each $\alpha \in (0,1)$,

$$\begin{aligned}
&\alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)S \\
\Longleftrightarrow & f(\alpha P + (1-\alpha)R) \succ^* f(\alpha Q + (1-\alpha)S) \\
\Longleftrightarrow & \alpha f(P) + (1-\alpha)f(R) \succ^* \alpha f(Q) + (1-\alpha)f(S) \\
\Longleftrightarrow & \alpha d + (1-\alpha)d'' \succ^* \alpha d' + (1-\alpha)d''.
\end{aligned}$$

Hence $\succsim^*$ satisfies the vNM independence axiom on $\mathcal{D}^*$.

Then we show the mixture continuity of $\succsim^*$ on $\mathcal{D}^*$. For any $d, d', d'' \in \mathcal{D}^*$, using the above arguments, we can find $P, Q, R \in \mathcal{P}$ such that $f(P) \sim^* d$, $f(Q) \sim^* d'$, $f(R) \sim^* d''$ and for each $\alpha \in (0,1)$, $\alpha P + (1-\alpha)Q \in \mathcal{P}$, $f(\alpha P + (1-\alpha)Q) = \alpha f(P) + (1-\alpha)f(Q)$ and $\alpha d + (1-\alpha)d' \sim^* \alpha f(P) + (1-\alpha)f(Q)$. Then

$$\begin{aligned}
A &= \Big\{\alpha \in [0,1] : \alpha d + (1-\alpha)d' \succ^* d''\Big\} \\
&= \Big\{\alpha \in [0,1] : \alpha f(P) + (1-\alpha)f(Q) \succ^* f(R)\Big\} \\
&= \Big\{\alpha \in [0,1] : f(\alpha P + (1-\alpha)Q) \succ^* f(R)\Big\} \\
&= \Big\{\alpha \in [0,1] : \alpha P + (1-\alpha)Q \succ R\Big\}.
\end{aligned}$$

By mixture continuity of $\succsim$ on $\mathcal{P}$, we know A is open in $[0,1]$. Similarly, $\Big\{\alpha \in [0,1] : \alpha d + (1-\alpha)d' \prec^* d''\Big\}$ is also open in $[0,1]$. Thus, $\succsim^*$ satisfies mixture continuity on $\mathcal{D}^*$.

By the Mixture Space Theorem, $\succsim^*$ on $\mathcal{D}^*$ admits an EU representation U with a utility index $w_D : X_1 \times \mathcal{L}^0(X_2) \to \mathbb{R}$. That is, for each $d \in \mathcal{D}^*$

$$U(d) = \sum_{x,p} w_D(x,p)d(x,p).$$

Moreover, w_D is unique up to a positive affine transformation.

Step 4. Since $\succsim^*$ extends $\succsim'$ from $\hat{\mathcal{D}}^*$ to $\mathcal{D}^*$, $d \succsim' d'$ if and only if $f^{-1}(d) \succsim f^{-1}(d')$ for all

$d, d' \in \hat{\mathcal{D}}^*$ and f is bijective, $\succsim$ on $\mathcal{P}$ is represented by

$$V(P) = \sum_x P_1(x) w_D(x, P_{2|x}), \forall\ P \in \mathcal{P}.$$

Then we want to show that the above utility function can be rewritten as (11).

First, by Lemma 3, for each $x \in X_1$, $\succsim_{2|x}$ on $\{\delta_x\} \times \mathcal{L}^0(X_2)$ admits an EU representation with some continuous and strictly monotonic utility index v_x. Then there exists a continuous and strictly monotonic function ϕ_x such that for all $p \in \mathcal{L}^0(X_2)$,

$$V(x, p) = w_D(x, p) = \phi_x(CE_{v_x}(p)).$$

Define $\hat{u}_D : X_1 \times X_2 \to \mathbb{R}$ as $\hat{w}(x, y) = \phi_x(y) = w_D(x, y)$ for all $(x, y) \in X_1 \times X_2$. Then the utility function can be rewritten as

$$V(P) = \sum_x \hat{u}_D(x, CE_{v_x}(P_{2|x})) P_1(x), \forall\ P \in \mathcal{P}$$

where v_x is continuous and strictly monotonic for each $x \in X_1$. This is exactly (11). By Monotonicity, we know that $\hat{u}_D$ is strictly monotonic. $\hat{u}_D$ is bounded since $\hat{u}_D(\underline{c}_1, \underline{c}_2) \leq \hat{u}_D(x, y) \leq \hat{u}_D(\bar{c}_1, \bar{c}_2)$ for all $(x, y) \in X_1 \times X_2$. Moreover, $\hat{u}_D$ is unique up to a positive affine transformation.

The final step is to verify that $\hat{u}_D$ is continuous. Suppose by contradiction that $\hat{u}_D$ is not continuous, then we can find $(x, y) \in X_1 \times X_2$ and a sequence $(x_n, y_n) \to (x, y)$ such that $\lim_{n\to\infty} \hat{u}_D(x_n, y_n) \neq \hat{u}_D(x, y)$. As $\hat{u}_D$ is bounded, we can find a subsequence of $\{(x_n, y_n)\}_{n\geq 1}$ (still denoted as $\{(x_n, y_n)\}_{n\geq 1}$) such that $\lim_{n\to\infty} \hat{u}_D(x_n, y_n) = a \neq b = \hat{u}_D(x, y)$. Without loss of generality, assume that $a < b$. Note that for $x_1 \neq x_2$ and $\eta \in (0, 1)$, $V(\eta(x_1, y_1) + (1 - \eta)(x_2, y_2)) = \eta V(x_1, y_1) + (1 - \eta) V(x_2, y_2)$. We can find some $P \in \mathcal{P}$ with $V(P) \in (a, b)$. As $\lim_{n\to\infty} \hat{u}_D(x_n, y_n) = a < V(P)$, for n large enough, we have $\hat{u}_D(x_n, y_n) < V(P) < b$, that is, $(x_n, y_n) \prec P \prec (x, y)$. When n goes to infinity, $(x_n, y_n) \xrightarrow{w} (x, y)$. However, by Continuity, we must have $(x, y) \precsim P$, which leads to a contradiction. Hence $\hat{u}_D$ is continuous. □

Recall that $\hat{\succsim}$ satisfies Correlation Neutrality and agrees with $\succsim$ on $\hat{\mathcal{P}}$. Corollary 6 states that $\hat{\succsim}$ can be represented by one of the following types of utility functions: V^{CN}, V^{ANB-CN} or V^{FNB}. If $\succsim$ satisfies Correlation Neutrality, then $\succsim = \hat{\succsim}$ and it is also represented by one of V^{CN}, V^{ANB-CN} or V^{FNB}. If instead $\succsim$ violates Correlation Neutrality, then Lemma 24 states that $\succsim$ is represented by the utility function (11)

$$U(P) = \sum_x \hat{u}_D(x, CE_{v_x}(P_{2|x})) P_1(x),$$

where $\hat{u}_D$ and v_x for each $x \in X_1$ are continuous, strictly monotonic and unique up to a positive affine transformation. Hence we can exploit the consistency of $\succsim$ and $\hat{\succsim}$ on $\hat{\mathcal{P}}$ to refine the utility function (11). Note that on $\hat{\mathcal{P}}$, (11) becomes

$$U(p,q) = \sum_x \hat{u}_D(x, CE_{v_x}(q))p(x),$$

Case 1: Suppose $\hat{\succsim}$ is represented by V^{CN}. Then for each $(p,q) \in \hat{\mathcal{P}}$,

$$V^{CN}(p,q) = \sum_{x,y} w(x,y)p(x)q(y)$$

where w is continuous, strictly monotonic and unique up to a positive affine transformation. Then it must be the case that u_D is a positive affine transformation of w and v_x is a positive affine transformation of $w(x,\cdot)$ for all $x \in X_2$. Thus $\succsim$ on $\mathcal{P}$ is represented by

$$V^{EU}(P) = \sum_x w(x, CE_{w(x,\cdot)}(P_{2|x}))P_1(x) = \sum_{x,y} w(x,y)P_1(x)P_{2|x}(y) = \sum_{x,y} w(x,y)P(x,y).$$

Case 2: Suppose that $\hat{\succsim}$ is represented by V^{BIB-CN} or V^{FNB}. Then for each $x \in X_1$, $\succsim_{2|x}$ is admits an EU representation with some index v_2, which is independent of x. This suggests that v_x must be a positive affine transformation of v_2 for all $x \in X_1$. Thus $\succsim$ on $\mathcal{P}$ is represented by

$$V^{ANB}(P) = \sum_x \hat{u}_D(x, CE_v(P_{2|x}))P_1(x).$$

To summarize, if $\succsim$ satisfies Weak Order, Monotonicity, Continuity, Independence* and Correlation Consistency, then $\succsim$ an be represented by one of the following types of utility functions: V^{EU}, V^{ANB}, V^{CN} , V^{ANB-CN} or V^{FNB}. □

Proof of Proposition 1. The arguments for uniqueness of utility indices in the representations are already contained in the proofs of Theorem 1 and Theorem 2. □

Proof of Corollary 1. For necessity, Symmetry is guaranteed since both $\succsim_1$ and $\succsim_2$ admit an expected utility representation with utility index u. Stochastic Dominance without Risk is satisfied since for each $x, y, x', y' \in X$, $(x,y) \succsim (x',y')$ if and only if $u(x+y) \geq u(x'+y')$ if and only if $x + y \geq x' + y'$. The last equivalence holds as u is strictly monotonic.

Now we check sufficiency. For each utility function, Stochastic Dominance without Risk implies that $w(x,y)$ can be rewritten as $\phi(x+y)$ for each $(x,y) \in Z$, where ϕ is strictly monotonic and continuous. By Symmetry, the ϕ agrees with u in utility functions V^{ANB} and V^{ANB-CN} and v agrees with u in V^{FNB}. Since w is unique up to a monotonic transformation in V^{FNB}, the five utility functions can be written as in Corollary 1. □

Proof of Proposition 2. We first prove (v). For any u continuous and strictly monotonic and $P \in \mathcal{P}$, $V^{EU}(P) = \sum_{x,y} u(x+y)P(x,y) = \sum_z u(z)f[P](z)$. By monotonicity of u, $V^{EU}(P) > V^{EU}(Q)$ if $f[P] \succ_{FOSD} f[Q]$. Hence EU satisfies stochastic dominance on $\mathcal{C}_4 = \mathcal{P}$.

(i). By Corollary 1, FNB satisfies Stochastic Dominance without Risk and hence it satisfies stochastic dominance on $\mathcal{C}_1$. Now we provide an example that FNB might violate stochastic dominance on $\mathcal{C}_2$. Consider $P = (\frac{1}{2}\delta_{101} + \frac{1}{2}\delta_0, \delta_{100})$ and $Q = (\frac{1}{2}\delta_{200} + \frac{1}{2}\delta_{100}, \delta_0)$. Clearly, $f[P] \succ_{FOSD} f[Q]$ and $P, Q \in \mathcal{C}_2$. Assume that $u(x) = \sqrt{x}$ for $x \in X \cap \mathbb{R}_+$ and $u(x) = -\sqrt{-x}$ for $x \in X \cap \mathbb{R}_-$. Then $V^{FNB}(P) = (101/4 + 100)^{1/2} < V^{FNB}(Q) = 5(1+\sqrt{2})$ and FNB does not satisfy stochastic dominance on $\mathcal{C}_2$.

(ii). For any u continuous and strictly monotonic and $(p, y) \in \mathcal{C}_2$,

$$V^{ANB-CN}(p, y) = \sum_{x \in X} u(x+y)p(x) = V^{EU}(p, y).$$

Then (v) implies that ANB-CN satisfies stochastic dominance on $\mathcal{C}_2$.

To show that ANB-CN might violate stochastic dominance on $\mathcal{C}_3$, consider $P = \frac{1}{2}(0, 100) + \frac{1}{2}(100, 0)$ and $Q = (99, 0)$. Clearly, $f[P] \succ_{FOSD} f[Q]$ and $P, Q \in \mathcal{C}_3$. Assume that $u(x) = \sqrt{x}$ for $x \in X \cap \mathbb{R}_+$ and $u(x) = -\sqrt{-x}$ for $x \in X \cap \mathbb{R}_-$. Then $V^{ANB-CN}(P) = \frac{1}{2}(5 + 5\sqrt{5}) < V^{FNB}(Q) = \sqrt{99}$ and ANB-CN does not satisfy stochastic dominance on $\mathcal{C}_3$.

To show that ANB-CN might violate stochastic dominance on $\mathcal{C}_{3'}$, consider $P = (\delta_{100}, \frac{1}{2}\delta_{101} + \frac{1}{2}\delta_0)$ and $Q = (\delta_0, \frac{1}{2}\delta_{200} + \frac{1}{2}\delta_{100})$. Clearly, $f[P] \succ_{FOSD} f[Q]$ and $P, Q \in \mathcal{C}_{3'}$. Assume that $u(x) = \sqrt{x}$ for $x \in X \cap \mathbb{R}_+$ and $u(x) = -\sqrt{-x}$ for $x \in X \cap \mathbb{R}_-$. Then $V^{FNB}(P) = (101/4 + 100)^{1/2} < V^{FNB}(Q) = 5(1 + \sqrt{2})$ and ANB-CN does not satisfy stochastic dominance on $\mathcal{C}_{3'}$.

(iii). For any u continuous and strictly monotonic and $P \in \mathcal{C}_3$, denote $P_{2|x} = z_x$ for each $P_1(x) > 0$, and we have

$$V^{ANB}(P) = \sum_{x \in X} u(x + z_x)p(x) = V^{EU}(P).$$

Then (v) implies that ANB satisfies stochastic dominance on $\mathcal{C}_3$. The example that ANB does not satisfy stochastic dominance on $\mathcal{C}_{3'}$ is the same as example of ANB-CN in (ii).

(iv). For any u continuous and strictly monotonic and $(p, q) \in \mathcal{C}_{3'}$,

$$V^{CN}(p, q) = \sum_{x,y \in X} u(x+y)p(x)q(y) = V^{EU}(p, q).$$

Then (v) implies that CN satisfies stochastic dominance on $\mathcal{C}_{3'}$. The example that CN does not satisfy stochastic dominance on $\mathcal{C}_3$ is the same as example of ANB-CN in (ii). □

Proof of Corollary 2. The proof is similar to the proof of Corollary 1 and omitted. □

Proof of Lemma 1. s^{EU} satisfies the following first order condition:

$$\frac{1}{y_1 - s} = \beta R \frac{\sum_i \pi_i (y_{2,i} + Rs)^{\beta\alpha - 1}}{\sum_i \pi_i (y_{2,i} + Rs)^{\beta\alpha}}. \tag{12}$$

Similarly, s^{NB} satisfies the following first order condition:

$$\frac{1}{y_1 - s} = \beta R \frac{\sum_i \pi_i (y_{2,i} + Rs)^{\alpha - 1}}{\sum_i \pi_i (y_{2,i} + Rs)^{\alpha}}. \tag{13}$$

For any $\alpha < 1$ and $\alpha \neq 0$, one can easily show that expressions on the right hand side (RHS) of equations (12) and (13) are strictly decreasing in s, while the expressions on the left hand side (LHS) are strictly increasing in s. This implies that both s^{EU} and s^{NB} are unique and $\frac{y_{2,i}}{R} < s^{NB}, s^{EU} < y_1$ for all i. Moreover, if

$$\beta R \frac{\sum_i \pi_i (y_{2,i} + Rs)^{\beta\alpha - 1}}{\sum_i \pi_i (y_{2,i} + Rs)^{\beta\alpha}} < \beta R \frac{\sum_i \pi_i (y_{2,i} + Rs)^{\alpha - 1}}{\sum_i \pi_i (y_{2,i} + Rs)^{\alpha}} \tag{14}$$

for any interior s, i.e., $\frac{y_{2,i}}{R} < s < y_1$ for all i, then we claim that $s^{NB} > s^{EU}$. To see this, suppose by contradiction that $s^{NB} \leq s^{EU}$. Then

$$\begin{aligned}
\frac{1}{y_1 - s^{EU}} &= \beta R \frac{\sum_i \pi_i (y_{2,i} + Rs^{EU})^{\beta\alpha - 1}}{\sum_i \pi_i (y_{2,i} + Rs^{EU})^{\beta\alpha}} \\
&\leq \beta R \frac{\sum_i \pi_i (y_{2,i} + Rs^{NB})^{\beta\alpha - 1}}{\sum_i \pi_i (y_{2,i} + Rs^{NB})^{\beta\alpha}} \\
&< \beta R \frac{\sum_i \pi_i (y_{2,i} + Rs^{NB})^{\alpha - 1}}{\sum_i \pi_i (y_{2,i} + Rs^{NB})^{\alpha}} \\
&= \frac{1}{y_1 - s^{NB}}.
\end{aligned}$$

The first and the last inequalities come from (12) and (13). The second inequality holds since $s^{NB} \leq s^{EU}$ and the RHS of (12) is decreasing in s. The third inequality follows from (14). Hence, we have $s^{EU} < s^{NB}$, which leads to a contradiction.

It remains to be shown that (14) holds for all $\frac{y_{2,i}}{R} < s < y_1$ for all i. Note that (14) is equivalent to

$$\sum_{i,j} \pi_i \pi_j (y_{2,i} + Rs)^{\beta\alpha - 1} (y_{2,j} + Rs)^{\alpha} < \sum_{i,j} \pi_i \pi_j (y_{2,i} + Rs)^{\alpha - 1} (y_{2,j} + Rs)^{\beta\alpha}.$$

Both sides of the inequality consist of N^2 elements, each indexed by (i,j). For $i = j$, the elements on the two sides cancel out. Since $\pi_i > 0$ and $\frac{y_{2,i}}{R} < s < y_1$ for all $i = 1, ..., N$, it suffices to show that for each $i \neq j$,

$$\begin{aligned}
&(y_{2,i} + Rs)^{\beta\alpha - 1} (y_{2,j} + Rs)^{\alpha} + (y_{2,j} + Rs)^{\beta\alpha - 1} (y_{2,i} + Rs)^{\alpha} \\
&< (y_{2,i} + Rs)^{\alpha - 1} (y_{2,j} + Rs)^{\beta\alpha} + (y_{2,j} + Rs)^{\alpha - 1} (y_{2,i} + Rs)^{\beta\alpha}
\end{aligned}$$

$$\Longleftrightarrow (y_{2,j} - y_{2,i})\Big((y_{2,j} + Rs)^{\alpha(1-\beta)} - (y_{2,i} + Rs)^{\alpha(1-\beta))}\Big) < 0.$$

Since $y_{2,j} \neq y_{2,i}$ if $i \neq j$ and $\beta \in (0,1)$, the above inequality holds if $\alpha < 0$. Hence, when $\alpha < 0$, $s^{NB} > s^{EU}$.

Symmetrically, if $\alpha > 0$, then the inequality in (14) is reversed and we have $s^{NB} < s^{EU}$. □

Proof of Lemma 2. Denote by s_i the saving level in state $i = l, h$. s_l^{EU} and s_h^{EU} satisfy the following first order condition:

$$\frac{y_{2,i} + Rs_i}{y_1 - s_i} = \beta R. \tag{15}$$

If the consumer's preference is represented by V^{CN}, then she chooses s_l and s_h to maximize

$$\max_{s_l, s_h} \frac{1}{4\alpha} \sum_{i,j=l.h} (y_1 - s_i)^{\alpha} (y_{2,j} + Rs_j)^{\beta\alpha}$$

s_l^{CN} and s_h^{CN} satisfy the following first order conditions:

$$(y_1 - s_l)^{\alpha-1}\Big[(y_{2,l} + Rs_l)^{\beta\alpha} + (y_{2,h} + Rs_h)^{\beta\alpha}\Big] = \beta R (y_{2,l} + Rs_l)^{\beta\alpha-1}\Big[(y_1 - s_l)^{\alpha} + (y_1 - s_h)^{\alpha}\Big], \tag{16}$$

$$(y_1 - s_h)^{\alpha-1}\Big[(y_{2,l} + Rs_l)^{\beta\alpha} + (y_{2,h} + Rs_h)^{\beta\alpha}\Big] = \beta R (y_{2,h} + Rs_h)^{\beta\alpha-1}\Big[(y_1 - s_l)^{\alpha} + (y_1 - s_h)^{\alpha}\Big]. \tag{17}$$

Divide equation (16) by (17) and we get

$$\frac{(y_1 - s_l)^{\alpha-1}}{(y_1 - s_h)^{\alpha-1}} = \frac{(y_{2,l} + Rs_l)^{\beta\alpha-1}}{(y_{2,h} + Rs_h)^{\beta\alpha-1}}. \tag{18}$$

First, we claim that $s_l^{CN} > s_h^{CN}$. Suppose by contradiction that $s_l^{CN} \leq s_h^{CN}$. Since $\alpha < 1, \beta\alpha < 1$ and $y_{2,l} < y_{2,h}$, easy to see that $y_1 - s_l^{CN} \geq y_1 - s_h^{CN}$ and $y_{2,l} + Rs_l^{CN} < y_{2,h} + Rs_h^{CN}$. Then

$$\frac{(y_1 - s_l^{CN})^{\alpha-1}}{(y_1 - s_h^{CN})^{\alpha-1}} \leq 1 < \frac{(y_{2,l} + Rs_l^{CN})^{\beta\alpha-1}}{(y_{2,h} + Rs_h^{CN})^{\beta\alpha-1}}$$

which contradicts with (18).

Rewrite (16) and (17) as follows:

$$\frac{y_{2,l} + Rs_l^{CN}}{y_1 - s_l^{CN}} = \beta R \frac{1 + \frac{(y_1 - s_h^{CN})^{\alpha}}{(y_1 - s_l^{CN})^{\alpha}}}{1 + \frac{(y_{2,h} + Rs_h^{CN})^{\beta\alpha}}{(y_{2,l} + Rs_l^{CN})^{\beta\alpha}}}, \tag{19}$$

$$\frac{y_{2,h} + Rs_h^{CN}}{y_1 - s_h^{CN}} = \beta R \frac{1 + \frac{(y_1 - s_l^{CN})^{\alpha}}{(y_1 - s_h^{CN})^{\alpha}}}{1 + \frac{(y_{2,l} + Rs_l^{CN})^{\beta\alpha}}{(y_{2,h} + Rs_h^{CN})^{\beta\alpha}}}. \tag{20}$$

We focus on the comparison of s_l^{CN} and s_l^{EU}. The comparison of s_h^{CN} and s_h^{EU} is similar. The expressions on the LHS of (19) and (15) are strictly increasing in s_l. Like the proof of

Lemma 1, we know that $s_l^{CN} > s_l^{EU}$ if and only if

$$\frac{(y_1 - s_h^{CN})^\alpha}{(y_1 - s_l^{CN})^\alpha} > \frac{(y_{2,h} + Rs_h^{CN})^{\beta\alpha}}{(y_{2,l} + Rs_l^{CN})^{\beta\alpha}}.$$

Since s_l^{CN} and s_h^{CN} satisfy (18), the above inequality reduces to

$$\frac{(y_{2,l} + Rs_l^{CN})}{(y_1 - s_l^{CN})} > \frac{(y_{2,h} + Rs_h^{CN})}{(y_1 - s_h^{CN})}. \tag{21}$$

Again by (18),

$$\Big(\frac{y_1 - s_l^{CN}}{y_1 - s_h^{CN}}\Big)^{\alpha(1-\beta)} = \Big(\frac{\frac{y_{2,h}+Rs_h^{CN}}{y_1 - s_h^{CN}}}{\frac{y_{2,l}+Rs_l^{CN}}{y_1 - s_l^{CN}}}\Big)^{1-\beta\alpha}.$$

Since $\alpha < 1$, $\beta \in (0, 1)$ and $s_l^{CN} > s_h^{CN}$, (21) holds if and only if $\alpha > 0$. Hence, $s_l^{CN} > s_l^{EU}$ if $\alpha > 0$ and $s_l^{CN} < s_l^{EU}$ if $\alpha < 0$. Similarly, we can show by (20) that $s_h^{CN} < s_h^{EU}$ if $\alpha > 0$ and $s_h^{CN} > s_h^{EU}$ if $\alpha < 0$. □

Proof of Proposition 3. First, by Theorem 5.1 in Epstein and Zin (1989), we can find an optimal consumption plan $\{c_t^*\}_{t\geq 1}$ for the EZ consumer with $J^{EZ}(W_1) = \hat{V}^{EZ}(\{c_t^*\}_{t\geq 1})$. On the one hand, since $\rho > \alpha$, the EZ consumer has a preference for early resolution of risk. We know that $\hat{V}^{ANB}(\{c_t\}_{t\geq 1}) \leq \hat{V}^{EZ}(\{c_t\}_{t\geq 1})$ for each feasible consumption plan $\{c_t\}_{t\geq 1}$ and $J^{ANB}(W_1) \leq J^{EZ}(W_1)$. Moreover, for a consumption plan $\{c_t\}_{t\geq 1}$ where c_t is injective on Ω_t for each t, the consumption history contains the same information as state history. In this case, $\hat{V}^{ANB}(\{c_t\}_{t\geq 1}) = \hat{V}^{EZ}(\{c_t\}_{t\geq 1})$.

If c_t^* is injective for each $t \geq 2$, we can set $c_t^n \equiv c_t^*$ for each n and the results holds. Otherwise, we want to construct a sequence of consumption plans $\{c_t^n\}_{t\geq 1, n\geq 1}$ with injective consumption functions such that $c_1^n \to c_1^*, c_t^n(s^t) \to c_t^*(s^t)$ as n goes to infinity for each $t \geq 2$. By continuity of V^{EZ} and hence $\hat{V}^{EZ}$, we know

$$J^{ANB}(W_1) \geq \lim_{n\to\infty} \hat{V}^{ANB}(\{c_t^n\}_{t\geq 1}) = \hat{V}^{EZ}(\{c_t^*\}_{t\geq 1}) = J^{EZ}(W_1).$$

Hence $J^{ANB}(W_1) = J^{EZ}(W_1)$. It remains to construct the sequence of consumption plans. This is easy since $\bigcup_{t\geq 2} c_t^*(S^t)$ is countable (as the union of a countable number of countable sets is countable), and the space of consumption is a continuum. □

Proof of Corollary 3. The derivation of the Euler equation and the asset pricing formula for $\{c_t\}_{t\geq 1}$ can be found in Epstein and Zin (1991). The rest follows from the continuity of the two equations and $c_1^n \to c_1, c_t^n(s^t) \to c_t(s^t)$ as n goes to infinity for each $t \geq 2$. □