From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2021

Essays on Rationality: Modeling and Measureme

Mononen, Lasse

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01rx913s98c

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Essays on Rationality: Modeling and Measurement

Lasse Miikka Mononen

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance by
the Department of Economics
Advisor: Pietro Ortoleva

May 2021

© Copyright by Lasse Miikka Mononen, 2021.

All rights reserved.

Abstract

The assumption of rationality is one of the cornerstones of economic analysis. Yet, substantial evidence has documented violations of rationality. This dissertation studies modeling and measurement of rationality.

Chapter 1 studies extending rationality to state dependent utilities in choice under uncertainty where the outcomes depend on the realized state of the world. The typical assumption is that utilities of outcomes do not depend on the realized state and are state independent. Without this simplifying assumption, it is difficult to separately identify utilities and beliefs. We provide novel general foundations for models with state dependent utilities: once we depart from expected utility, it is often possible to uniquely identify utilities and beliefs. Specifically, we show that with general models of non-expected utility under ambiguity we have complete identification of utilities and probabilities under full-dimensional uncertainty.

In Chapter 2, we study how to model cost of rationality or complexity in choice under risk. We consider a decision maker who cancels out common consequences with common probabilities when comparing lotteries. Under continuity assumptions, we show that there are two possibilities. 1) The decision maker is an expected utility maximizer with Shannon's entropy as an additive cost of complexity. 2) The decision maker is a power weighted expected utility maximizer where the probabilities are weighted by a power function. Additionally, we provide novel characterizations for expected utility. We consider a decision maker who evaluates each possible prize independently of the other possible prizes. If the preferences satisfy first order stochastic dominance or joint continuity in prizes and probabilities, then the decision maker is an expected utility maximizer.

Chapter 3 studies measurement of rationality violations. Over the past decades, several measures have been proposed to quantify rationality violations. However, little is known about their foundations and how they compare with each other. We provide novel axiomatizations and simulations for the most used measures of rationality in the standard budget-choice setting. This allows both theoretical and practical comparisons of the ratio-

nality measures. Additionally, the simulations suggest improving the current measures by overweighting small violations of rationality.

Acknowledgements

This dissertation would not have been possible without the effort of many people.

I am very grateful to my advisor Pietro Ortoleva for the support, encouragement, and patience throughout my years at Princeton whose guidance and insights have been invaluable. I am also very grateful to Faruk Gul for the guidance, encouragement, and patience at the beginning of my time at Princeton.

I would like to thank Roland Bénabou, Stephen Morris, Wolfgang Pesendorfer, Can Urgun, and Leeat Yariv for their helpful advice and critical feedback on my research.

I thank Juuso Välimäki for encouraging to apply to graduate school and to aim higher.

I thank Laura Hedden for the excellent administrative support.

Many thanks to my colleagues and friends SeHyoun Ahn, Dre Helmns, Korhan Koçak, Franz Ostrizek, Chrissy Ostrowski, Evgenii Safonov, Denis Shishkin, and Mu Zhang who have made the time at Princeton much better and from whom I learned a great deal.

I gratefully acknowledge financial support from the Department at Economics and Emil Aaltonen Foundation.

Finally, I would like to thank my family for their patience and understanding.

;

Contents

Abstract iii

Acknowledgements v

1 State Dependent Utility and Ambiguity 1

1.1 Introduction 1

1.1.1 An Example of Identification 4

1.2 State Dependent Dual-Self Expected Utility 8

1.2.1 Preliminaries and Notation 9

1.2.2 State Dependent Dual-Self Expected Utility 10

1.2.3 Axioms for Existence 11

1.2.4 Existence 15

1.2.5 Identification 17

1.3 State Dependent Dual-Self Variational Expected Utility 25

1.3.1 Existence, Variational 26

1.3.2 Identification, Variational 27

1.4 Application to Social Choice Theory 31

1.5 Discussion on Identification 35

1.5.1 State Dependent Bewley Expected Utility 36

1.6 Related Literature 38

1.7 Conclusion 40

2 On Preference for Simplicity, Probability Weighting, and Expected Utility 42

2.1 Introduction . . . 42
2.2 Preliminaries and Preference for Simplicity . . . 48
2.2.1 Preliminaries . . . 48
2.2.2 Entropy Modified and Power Weighted Expected Utility . . . 50
2.3 Expected Utility . . . 55
2.3.1 Expected Utility and First Order Stochastic Dominance . . . 57
2.3.2 Expected Utility and Continuity . . . 59
2.4 Extensions on Preference for Simplicity . . . 60
2.4.1 Normalized Power Weighted Expected Utility . . . 60
2.4.2 Variational Expected Utility . . . 62
2.5 Uniqueness of the Representations . . . 64
2.6 Proof Sketches . . . 66
2.6.1 Proof Sketch for Entropy Modified Expected Utility . . . 66
2.6.2 Proof Sketch for Expected Utility with Continuity . . . 68
2.7 Related Literature . . . 68
2.8 Conclusion . . . 70

3 The Foundations of Measuring (Cardinal) Rationality 72

3.1 Introduction . . . 72
3.1.1 Example: The Choice of Measure Matters . . . 74
3.2 Measures of Rationality and Axiomatizations . . . 76
3.2.1 Afriat's Efficiency Index . . . 78
3.2.2 Normalized Money Pump Index . . . 84
3.2.3 Normalized Minimum Cost Index . . . 90
3.2.4 Varian's Goodness of Fit . . . 96
3.2.5 Inverse Varian's Goodness of Fit . . . 98
3.2.6 Axiomatic Summary . . . 99
3.3 Simulation Comparisons . . . 100

3.3.1 Choice Model . . . 101
3.3.2 Ordinal Measure Comparisons . . . 103
3.3.3 Calibration . . . 104
3.3.4 Comparisons of Measures . . . 105
3.3.5 Variations of Measures . . . 108
3.3.6 Computation of Measures and Simulations . . . 116
3.4 Related Literature . . . 118
3.5 Conclusion . . . 118

A Appendix for Chapter 1 **120**

A.1 General Identification . . . 121
A.2 Latent Expected Utility Representation . . . 124
A.2.1 Proof for Latent Expected Utility . . . 124
A.3 Existence of State Dependent Dual-Self Expected Utility and Dual-Self Variational Expected Utility . . . 132
A.3.1 Statewise Weak Affine Representations . . . 132
A.3.2 Preferences Monotonic Between Equally Crisp Acts . . . 134
A.3.3 Transfinite Induction . . . 136
A.3.4 Necessity of the Axioms for Representations . . . 150
A.3.5 State Independent Dual-Self Expected Utility . . . 152
A.4 Dual-Self Identification . . . 158
A.4.1 Dual-Self Variational Representation Additive Between Equally Crisp Acts . . . 159
A.4.2 Behavioral Directional Derivatives . . . 167
A.4.3 Partial Identification . . . 173
A.4.4 Full Identification . . . 179
A.4.5 Identification Characterization with Crisp Acts and Equally Crisp Acts 183
A.4.6 Relative Likelihood and Probability Characterizations . . . 185

A.5 Social Choice Theory: Symmetric State Dependent MaxMin 197
A.6 State Dependent Bewley Representation 202
A.6.1 Bewley Existence Proof 204
A.6.2 Bewley Uniqueness Proof 209
A.6.3 State Independent Bewley Proof 213

B Appendix for Chapter 2 **216**

B.1 Expected Utility 218
B.1.1 Continuity On-Diagonal 221
B.1.2 Monotonicity on Dyadic Lotteries 227
B.1.3 Expected Utility for Dyadic Lotteries 237
B.1.4 Expected Utility under FOSD 242
B.1.5 Expected Utility under Weak Converge Continuity 246
B.2 Basic Results Under Restricted Common Consequence Independence 252
B.2.1 Continuous and Separable Extension of Preferences Conditional on a Co-probability 252
B.2.2 Power Weighted Expected Utility for Preferences Conditional on a Co-probability 269
B.3 Variational Expected Utility 279
B.3.1 Local Power Weighted Expected Utility Representation 279
B.3.2 Transfinite Induction 289
B.4 Entropy Modified Expected Utility 312
B.5 Uniqueness of Variational Expected Utility 317
B.6 Normalized Power Weighted Expected Utility 321
B.7 Proofs for Remarks on Continuity 326

C Appendix for Chapter 3 **332**

C.1 Additional Graphs 333

C.1.1 Absolute Hit Rates 333
C.1.2 Alternative Calibration Method 336
C.2 Afriat's Efficiency Index: Algorithm 340
C.3 Proofs 341
C.3.1 Preliminary Results 341
C.3.2 Afriat's Efficiency Index 344
C.3.3 Normalized Money Pump Index 347
C.3.4 Normalized Minimum Cost Index 351

Bibliography **356**

Chapter 1

State Dependent Utility and Ambiguity[†]

1.1 Introduction

Decision making under uncertainty studies choice behavior when outcomes depend on the realized state of the world. Traditionally, it is assumed that the utilities of outcomes are *state independent* and do not depend on the realized state. This independence simplifies the identification of utilities and beliefs. However, as observed by Aumann (1971) in many situations outcomes and their utilities might be *state dependent*. Known examples are when the state of the world is the health of the decision maker, which is likely to affect the utility of outcomes, such as with health insurance (Arrow, 1971). Assuming state independent utilities in these situations may be inaccurate and lead to wrong identification and predictions.

This paper reconsiders state dependent utilities in the standard framework and shows that, under general conditions, it is possible to separately identify the state dependent utilities and probabilities whenever we have violations of the expected utility hypothesis. Additionally, we provide a novel axiomatic characterization for models with state dependent

[†]Materials from this chapter were presented at Risk, Uncertainty and Decision Conference 2020 and at D-TEA (Decision: Theory, Experiments, and Applications) 2020.

utilities under ambiguity. We illustrate the usefulness of these results with two applications unrelated to choice under uncertainty.

Before moving on to the results, we highlight the importance of identifying utilities and probabilities with a concrete example. Consider a government that wants to change people's behavior with a public health campaign, but people find the change difficult or inconvenient e.g. reduce smoking or increase the use of seat belts. Here the choice of an effective campaign depends critically on if the lack of change reflects a taste-based reason, quitting smoking is difficult, seat belts are uncomfortable, or if it reflects a belief-based reason, only heavy smokers get cancer or only reckless drivers get into accidents. In the first case, an effective campaign would make the change of behavior easier by increasing the availability of nicotine replacement products and making smoking socially less acceptable or redesigning seat belts to be more comfortable and convenient. In the second case, an effective campaign would be an information campaign on the effects of behavior changes and risks associated with the current behavior. Here, it is crucial to separate tastes from beliefs in order to choose an effective campaign.

This paper characterizes and identifies general state dependent models under uncertainty. We focus on standard preferences under ambiguity, that is continuous, monotonic, and risk independent preferences. Our first characterization result shows that if these standard preferences have two unambiguous acts, in a sense that will be made precise, then they admit a *dual-self expected utility* representation (non-convex multiple prior preferences, Chandrasekher et al., 2020) with state dependent utilities.[1]

In this context, we show our first main identification result: when the uncertainty about states of the world is full-dimensional, the probabilities and the utilities are fully identified. That is, the probabilities are uniquely identified and the state dependent utilities are iden-

[1]This result extends Chandrasekher et al.'s (2020) characterization. However, this extension is not entirely straightforward since under state-dependence, constant acts may no longer be unambiguous, so we need to find an alternative way to capture the characterizing properties of the representation. We will infer from the decision maker's behavior which acts are unambiguous and use this to capture the characterizing properties.

tified up to a common positive affine transformation. This shows that the impossibility of identification under expected utility is only a knife-edge case, due to the linearity of expected utility: once we depart from such linearity, we regain identification. Instead, with any full-dimensional uncertainty, we can identify the utilities from the violations of the expected utility hypothesis, as we illustrate with an example in the next section. More in general, we show that the identification of utilities and probabilities between two states is characterized by having uncertainty about the relative likelihood between the states. This general identification result has many important special cases. It characterizes the identification of state dependent utilities for models such as maxmin expected utility (multiple prior preferences, Gilboa and Schmeidler, 1989), Choquet expected utility (Schmeidler, 1989), α-maxmin expected utility (Ghirardato et al., 2004), and invariant biseparable preferences (Ghirardato et al., 2004; Amarante, 2009).

Next, we move on to a more general model under uncertainty. We show that if standard ambiguity preferences have two acts that share the same ambiguity and hedge ambiguity similarly, in a sense that will be made precise later on, then they admit a dual-self *variational* expected utility representation (non-convex variational preferences, Chandrasekher et al., 2020) with state dependent utilities. This generalizes the characterization from Chandrasekher et al. (2020) to state-dependent utilities. In this more general context, we show our second identification result: when the uncertainty about states is full-dimensional, then probabilities and intensities of preferences can be separated. However, in this case, the levels of utilities are not identified.[2] This general result characterizes the identification for state dependent variations of models such as monotone mean-variance preferences (Maccheroni et al., 2009), multiplier preferences (Hansen and Sargent, 2001; Strzalecki, 2011), variational preferences (Maccheroni et al., 2006), monotonic mean-dispersion preferences (Grant and Polak, 2013), and vector expected utility (Siniscalchi, 2009).

[2]Formally, the probabilities are uniquely identified and the changes of state dependent utilities are identified up to a common positive multiplication.

This new foundation for state dependent utilities has applications beyond ambiguity. We illustrate this with an application to social choice theory. Here, we interpret states to be members of a society and study the preferences of the society over distributions of goods. In this context, we show that the fairness of the society and interpersonal utility comparisons can be identified. In a general class of societies, utilitarian societies are only a literal knife-edge case where they are not identified. Instead, with any other society in this general class, the fairness and the interpersonal utilities can be identified.

This paper contributes to the literature studying the identification of state dependent utilities in models of non-expected utilities following Drèze (1987) and Chew and Wang (2020). This provides an alternative to the literature on using additional information (e.g. preferences conditional on signals with Bayesian updating, hypothetical lotteries of state-outcome pairs, or stochastic choice) for the identification of state dependent utilities as in Karni (2011a; 2011b), Karni and Schmeidler (2016), and Lu (2019).

The remainder of the paper proceeds as follows: We begin, in Section 1.1.1, by showing a simple example of the fundamental forces behind the identification result. Next, Section 1.2 studies state dependent dual-self expected utility. We axiomatize its existence in Sections 1.2.3 and 1.2.4 and characterize its identification in Section 1.2.5 with behavioral characterizations for the identification. Section 1.3 extends the existence and the identification results to state dependent dual-self variational expected utility. Section 1.4 discusses applications to social choice theory. Section 1.5 clarifies the underlying assumptions for our identification and discusses the limitations of the identification. Finally, Section 1.6 discusses related literature and Section 1.7 concludes. Proofs for all the results are in the Appendix.

1.1.1 An Example of Identification

We begin with a simple example illustrating that with state dependent expected utility the intensities of preferences and the probabilities cannot be separated. However, this is only an unidentified knife-edge case. In the second part of the example, we show that with state

dependent maxmin expected utility these can be separated and identified from the violations of the expected utility hypothesis.

To make the problem of separating utilities and probabilities concrete, consider an unemployed person, Bob. Unemployed Bob is offered a choice between \$1,000 if he is employed in six months and \$1,000 if he is unemployed in six months. He chooses the unemployment insurance that pays in the case of unemployment. What can we infer from his choice in this situation? Does the choice mean that Bob considers it more likely that he is unemployed in six months than that he would find a job? Or does the choice reflect that without employment Bob could use the additional money to pay for overdue bills and would like to insure against this prospect?

This is the fundamental problem in identifying utilities and probabilities. With state independent utilities we can always infer which event is more likely by comparing the same payoffs in different states as above. However, as soon as we consider the possibility that utilities might be state dependent, the same payoffs in different states are no longer the same in utilities and these two forces become inseparable. Utilities and probabilities are always only observed jointly, and it is not possible to separate them using simple comparisons.

Next, we formalize the above identification problem. Consider Anscombe-Aumann setup with two states of the world, 1 and 2, and outcomes that are lotteries on consequences $\Delta(X)$. Acts consist of consequences for each state, (f_1, f_2). Assume that preferences over acts (f_1, f_2) have a state dependent expected utility representation with an equal probability for both of the states

$$0.5u_1(f_1) + 0.5u_2(f_2)$$

where $u_1, u_2 : \Delta(X) \to \mathbb{R}$ are von Neumann-Morgenstern utilities.

Now these preferences have an alternative expected utility representation with any probability $p \in (0, 1)$ for state 1 since

$$0.5u_1(f_1) + 0.5u_2(f_2) = p\Big(\frac{0.5}{p}u_1(f_1)\Big) + (1-p)\Big(\frac{0.5}{1-p}u_2(f_2)\Big) = p\widetilde{u}_1(f_1) + (1-p)\widetilde{u}_2(f_2),$$

where the terms inside the parentheses define new state dependent utility functions $\widetilde{u}_1, \widetilde{u}_2$.

In this alternative representation following the above example, we have replaced the probability for employment with an intensity of preference for money. This highlights the impossibility of identifying the state dependent expected utility since the intensities of preferences are inseparable from the probabilities.

However, this lack of identification is only due to the fact that we were using expected utility. Instead, we can identify the utilities across the states from the violations of the expected utility hypothesis. Going back to the example, assume this time that Bob is uncertain about the probability of employment and is uncertainty averse focusing on the worst-case probability. Assume that we observe many pairwise choices and there is a violation of the expected utility hypothesis at an unemployment insurance of size \$x in a sense that will be made precise later on. What should we infer from this violation? Since Bob is uncertain about the probability of employment, this violation reflects a change in the worst-case probability. However, with two states of the world, there can be a change in the worst-case probability only if the utility in both states is the same.[3] So we can identify the utilities across states using changes in Bob's uncertainty attitude.

Next, we formalize this identification from changes in the uncertainty attitude. We move on to a state dependent maxmin expected utility over acts (f_1, f_2) defined by two probabilities $p_* < p^*$ for state 1 and affine von Neumann-Morgenstern utilities u_1, u_2

$$\min_{p \in [p_*, p^*]} pu_1(f_1) + (1-p)u_2(f_2).$$

We show that utilities are identifiable across states from non-linearities or violations of Independence axiom[4]. For this, let f_2 be a consequence for state 2 and let x, y be conse-

[3]The worst-case probability always maximizes the probability of the state with a lower utility.

[4]Independence axiom from Anscombe and Aumann (1963) characterizes the linearity of the subjective expected utility. It states that for all acts f, g, h and $\alpha \in (0,1)$,

$$f \succsim g \iff \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h$$

when consequences are gambles and mixtures of acts are defined statewise.

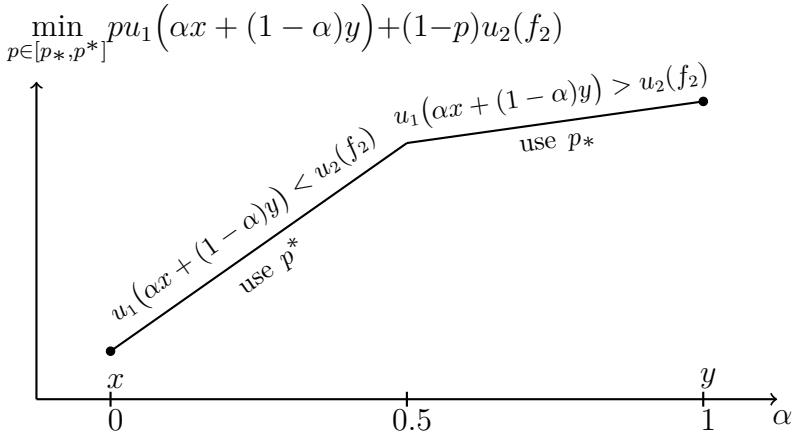

Figure 1.1. An example of identifying when the utilities across states are equal from non-linearities. x and y are such that $u_1(0.5x+0.5y)=u_2(f_2)$ and $u_1(x)<u_2(y)$. The x-axis changes x to y with convex combinations. The y-axis is the state dependent maxmin expected utility for the act $(\alpha x+(1-\alpha)y, f_2)$. At $\alpha=0.5$ utilities across state are equal that is observable as a violation of the expected utility hypothesis.

quences such that $u_1(0.5x+0.5y)=0.5u_1(x)+0.5u_1(y)=u_2(f_2)$ and $u_1(x)<u_1(y)$.[5] Next, we will show that there is a violation of Independence axiom at $(1/2x+1/2y, f_2)$.

We focus on the maxmin expected utility for acts $(\alpha x+(1-\alpha)y, f_2)$ when α changes from 0 to 1 as in Figure 1.1. First, between 0 and 0.5, $\alpha x+(1-\alpha)y$ gives more weight to x than to y and the act $(\alpha x+(1-\alpha)y, f_2)$ gives lower utility in state 1 than in state 2. So the maxmin expected utility uses probability p^*. Thus the maxmin expected utility for the act increases linearly at the rate $p^*\big(u_1(y)-u_1(x)\big)$. Second, between 0.5 and 1, $\alpha x+(1-\alpha)y$ has lower utility in state 2 and so the maxmin expected utility for the act $(\alpha x+(1-\alpha)y, f_2)$ uses probability p_*. This time the expected utility for the act increases linearly at the rate $p_*\big(u_1(y)-u_1(x)\big)$. Since the rate of increase changes at $\alpha=0.5$, there is a non-linearity at that point. This represents a violation of Independence axiom at the act $(0.5x+0.5y, f_2)$ that has the same utility for both states.

Finally, there can be violations of the independence axiom only if the utilities across states are the same. The only points where there can be non-linearities as in Figure 1.1 are points where the used probability changes. However, maxmin expected utility with two

[5]The first inequality follows from u_1 being a von Neumann-Morgenstern utility. If u_1 is unbounded, these x and y always exist.

states always maximizes the probability for the state with lower utility. Thus the only points where there are changes of probabilities are points where the utility order of states changes. Especially, at that point, the utilities across states are exactly the same.

In summary, the acts where the utilities across states are equal are characterized by the violations of Independence axiom and especially they are identifiable. This shows the identification of the utilities across states. Finally, this also shows the identification of probabilities since after identifying the utilities, we can apply the state-independent identification result directly.

This example illustrates that we can behaviorally observe when the utilities across states are equal from the violations of the expected utility hypothesis or Independence axiom. The identification came from the two different probabilities that is from the full-dimensional set of probabilities.

This identification is generalized in our main result, Theorem 1.2, to finitely many states of the world and to non-convex dual-self expected utility. There we show that if there is full-dimensional uncertainty, then we recover the state independent identification: the set of probabilities is unique and the utilities are unique up to a common positive affine transformation.

1.2 State Dependent Dual-Self Expected Utility

In order to study identification in the most general setup, we begin by studying a state dependent version of dual-self expected utility (Chandrasekher et al., 2020). The state-independent version is a general model that includes as special cases maxmin expected utility (Gilboa and Schmeidler, 1989), Choquet expected utility (Schmeidler, 1989), and α-maxmin expected utility (Ghirardato et al., 2004) and that is an alternative representation for invariant biseparable preferences (Ghirardato et al., 2004; Amarante, 2009). Here we characterize the existence and the identification of the state dependent version. Our results

thus encompass those for state dependent versions of all the special cases and alternative representations.

1.2.1 Preliminaries and Notation

Consider the finite Anscombe-Aumann (1963) framework with state dependent consequences. S is a finite state space, for each $s \in S$, X_s is a set of state dependent consequences and $\Delta(X_s)$ is the set of (simple) lotteries on X_s. Acts are mappings from states to state specific consequences and the set of acts is $H = \bigtimes_{s \in S} \Delta(X_s)$.[6] Our primitive is a binary relation $\succsim$ on H. As usual, $\succ$ and $\sim$ denote the asymmetric and symmetric parts of $\succsim$ respectively.

The following notation will be useful. $\Delta(S)$ is the set of probability measures on S. We endow $\Delta(S)$ with the Euclidean topology.[7] $\mathcal{K}(\Delta(S))$ is the set of all closed, convex, and non-empty subsets of $\Delta(S)$ endowed with the Hausdorff topology. For $P \subseteq \Delta(S)$, denote the convex closure of P by $\overline{\text{co}}P$. For $S' \subseteq S$ and $P \subseteq \Delta(S)$, denote the projection of P to S' by $\text{pr}_{S'} P = \{(p_s)_{s \in S'} | p \in P\}$.

For $f \in H, s \in S, x_s \in \Delta(X_s)$, f_s denotes the consequence of the act f in the state s and (x_s, f_{-s}) denotes the act where the consequence in the state s is x_s and in the states $s' \in S \setminus \{s\}$, $f_{s'}$. Mixtures of acts are defined statewise: for all $f, g \in H, \alpha \in [0,1], s \in S$, define $(\alpha f + (1-\alpha)g)_s = \alpha f_s + (1-\alpha) g_s$.

If consequences in some state do not affect the preferences, then the utility for these consequences is unobservable. Hence our focus is on proper states:

Definition A state $s \in S$ is *proper* if there exist $x_s, y_s \in \Delta(X_s)$ and $f \in H$ such that

$$(x_s, f_{-s}) \not\sim (y_s, f_{-s}).$$

The collection of proper states is denoted $S^{\mathcal{P}}$.

We infer the preferences on consequences within each state as follows:

[6] This includes the standard state independent Anscombe-Aumann setting where X_s=X for all s.

[7] This is the $|S|-1$ dimensional Euclidean topology when $\Delta(S)$ is represented as $\{p \in \mathbb{R}_+^{|S|-1} | \sum_{i=1}^{|S|-1} p_i \leq 1\}$.

Definition For each $s \in S$, define $\succsim_s$ on $\Delta(X_s)$ by for all $x_s, y_s \in \Delta(X_s)$,

$$x_s \succsim_s y_s \iff (x_s, f_{-s}) \succsim (y_s, f_{-s}) \text{ for all } f \in H.$$

Additionally, $\succ_s$ is the asymmetric part of $\succsim_s$.

1.2.2 State Dependent Dual-Self Expected Utility

The state independent dual-self expected utility was introduced by Chandrasekher et al. (2020) as a general model for preferences under ambiguity. This model generalizes the maxmin expected utility by allowing for multiple sets of beliefs that are aggregated optimistically. Axiomatically this model corresponds to relaxing the uncertainty aversion of the maxmin expected utility. We will use a state dependent variation of it:

Definition $(u, \mathbb{P})$ is a *state dependent dual-self expected utility for* $\succsim$ if $u = (u_s)_{s \in S}$ and for all $s \in S$, $u_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\mathbb{P} \subseteq \mathcal{K}(\Delta(S))$ is compact and non-empty such that for each $P \in \mathbb{P}, p \in P, s \notin S^{\mathcal{P}}$, $p_s = 0$, and for all $f, g \in H$,

$$f \succsim g \iff \max_{P \in \mathbb{P}} \min_{p \in P} \sum_{s \in S} p_s u_s(f_s) \geq \max_{P \in \mathbb{P}} \min_{p \in P} \sum_{s \in S} p_s u_s(g_s).$$

This representation is the maxmin expected utility when $\mathbb{P}$ is a singleton and the maxmax expected utility when each $P \in \mathbb{P}$ is a singleton. Additionally, since this representation is a non-convex generalization of maxmin expected utility, especially any uncertainty aversion and seeking of the preferences can be represented by the interplay of the sets of beliefs and the collection of sets of beliefs. That is by the interplay of extreme concavity (min) and convexity (max).

As observed in Chandrasekher et al. (2020), the state independent dual-self expected utility is not unique. However, in the state independent case the convex and closed smallest set of probabilities used by the representation is unique which gives us tight dual-self expected utility. We will use this tight representation to study the identification of the state dependent dual-self expected utility.

Definition Let $\succsim \subseteq H \times H$. $(u, \mathbb{P})$ is a *state dependent tight dual-self expected utility for* $\succsim$ if $(u, \mathbb{P})$ is a state dependent dual-self expected utility for $\succsim$ and if $(u, \widetilde{\mathbb{P}})$ is another state dependent dual-self expected utility for $\succsim$, then $\overline{\text{co}} \bigcup_{\widetilde{P} \in \widetilde{\mathbb{P}}} \widetilde{P} \supseteq \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P$.

1.2.3 Axioms for Existence

This section introduces six axioms that characterize the existence of a state dependent dual-self expected utility representation. This axiomatization highlights the generality of the representation by showing that essentially the only standard preferences under ambiguity that do not have a dual-self expected utility representation are such that every act is ambiguous or every act contains different ambiguity. A reader interested only in the identification can skip this and the following existence section.

The first four axioms define standard preferences under ambiguity following Cerreia-Vioglio et al. (2011) in the state dependent context. The first two axioms are standard rationality assumptions that the preferences are a nontrivial weak order that satisfy continuity.

Axiom 1.1 $\succsim$ is complete, transitive, and non-trivial.

Axiom 1.2 For all $f, g, h \in H$, the sets $\{\alpha \in [0,1] | \alpha f + (1-\alpha) g \succsim h\}$ and $\{\alpha \in [0,1] | h \succsim \alpha f + (1-\alpha) g\}$ are closed in $\mathbb{R}$.

The next assumption for standard preferences under ambiguity is Monotonicity.[8] This is the main axiom ruling out state dependent utilities by assuming that the consequences within each state of the world are ranked in the same order. We weaken this axiom to only within state monotonicity which allows for fully separate preferences within each state of the world and for state dependent consequences.

[8]The state independent monotonicity assumes that if $f, g \in H$ are such that for all $s \in S$, $f_s \succsim g_s$, then $f \succsim g$ where f_s and g_s are acts that give the consequences f_s and g_s, respectively, in every state.

Axiom 1.3 For all $s \in S, x_s, y_s \in \Delta(X_s), f, g \in H$,

$$(x_s, f_{-s}) \succ (y_s, f_{-s}) \Longrightarrow (x_s, g_{-s}) \succsim (y_s, g_{-s}).$$

This axiom states that the decision maker's revealed statewise preferences are consistent with monotonicity. If the decision maker reveals to prefer x_s over y_s in some situation, then by monotonicity the decision maker cannot reveal to strictly prefer y_s over x_s in another situation. This axiom is a state dependent version of the standard monotonicity. It is assuming that the decision maker is a consequentialist that only cares about the outcomes received in each state.[9] This axiom guarantees that $\succsim_s$ is complete.

The next assumption for standard preferences under ambiguity is Risk Independence assuming Independence on lotteries.[10] The next axiom assumes a weak version of Independence, ruling out strict preference reversals, for the lotteries within each state of the world. This Weak Independence was introduced in Einy (1989).[11]

Axiom 1.4 For all $s \in S, x_s, y_s, z_s \in \Delta(X_s), \alpha \in (0,1)$,

$$x_s \succ_s y_s \Longrightarrow \alpha x_s + (1-\alpha) z_s \succsim_s \alpha y_s + (1-\alpha) z_s.$$

Under standard ambiguity preferences, the dual-self expected utility is characterized by Certainty Independence which assumes that constant acts are unambiguous. In order to discuss later on preferences that do not satisfy this axiom, we divide this assumption into two parts.

The first part of the Certainty Independence is the existence of a single constant act that is unambiguous and cannot be used for ambiguity hedging. The intuition for this part

[9]Formally, this axiom is a weak version of Savage's (1954) Sure Thing Principle applied only to a single state.

[10]This is usually assumed implicitly through Certainty Independence or Weak Certainty Independence axioms.

[11]In Appendix Appendix A.2, we consider expected utility when Independence is relaxed to this Weak Independence. We show that this gives a latent expected utility representation that is an affine utility u and a continuous and weakly increasing censoring function H such that for all lotteries P, Q on prizes X,

$$x \succsim y \iff H\Big(\sum_x P(x)u(x)\Big) \geq H\Big(\sum_x Q(x)u(x)\Big).$$

Additionally, u is unique up to a positive affine transformation.

is that there is no uncertainty about the utility of constant acts since they give the same consequence with the same utility in every state of the world.

However, in the state dependent setting, it is not known which of the acts are unambiguous. Instead, we will infer from the decision maker's behavior which of the acts are unambiguous. The behavioral interpretation for these acts is that they never hedge ambiguity and Independence axiom is satisfied when mixing with them as defined in Ghirardato et al. (2004).

Definition $h \in H$ is a *crisp act*[12] if for all $f, g \in H, \alpha \in (0,1)$,

$$f \succsim g \iff \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h.$$

The Certainty Independence assumes that all constant acts are crisp. The next axiom relaxes this to assuming that there exists some crisp act.

Axiom 1.5 There exists a crisp act $c \in H$.

Here, it is not necessary that the crisp act is unambiguous. Instead, it suffices that the crisp act is a least ambiguous act. That is the ambiguity of the crisp act is contained also in every other act and so it cannot be used for ambiguity hedging.

Our first part of Certainty Independence assumed the existence of only a single crisp act instead of a continuum of crisp acts as Certainty Independence assumes. Thus our second part of Certainty Independence captures this difference by assuming that all the constant acts contain exactly the same ambiguity since they are unambiguous.

In the state dependent setting, we will need to infer from the decision maker's behavior which acts contain the same ambiguity. Behaviorally this means that the acts always hedge ambiguity similarly and so trading one of the acts for another with the same ambiguity cannot change the ambiguity hedged or the preferences. This is formalized in the following property from Maccheroni et al.'s (2006) Weak Certainty Independence assumption.

[12]Cerreia-Vioglio et al. (2011) use a different definition for crisp acts that does not extend to state dependent setting.

Definition $f, g \in H$ are *equally crisp acts* if for all $h, h' \in H, \alpha \in (0,1)$,

$$\alpha h + (1-\alpha) f \succsim \alpha h' + (1-\alpha) f \iff \alpha h + (1-\alpha) g \succsim \alpha h' + (1-\alpha) g.$$

In this definition, f and g are equally crisp if trading f for g in any mixtures does not affect the preferences.

Under Certainty Independence, all constant acts are equally crisp. We relax this by assuming that there exist some equally crisp acts that are statewise ordered.

Axiom 1.6 There exist $f^*, g^* \in H$ such that f^* and g^* are equally crisp and for all $s \in S^{\mathcal{P}}$, $f^*_s \succ_s g^*_s$.

The intuition for this axiom is that it captures the dispersional or relative nature of ambiguity: Ambiguity of an act comes from it having good outcomes in some states and bad outcomes in some other states and from the uncertainty if the realized state and outcome is good or bad. However, these bad states and outcomes are only bad in the relative sense that the realized outcome could have been better. Now, this axiom is assuming that it is always possible to make all the outcomes better and increase the level of the outcomes without changing the ambiguity of the act by using the two equally crisp acts. Since we are making all the outcomes better, this does not change the relative comparisons of the outcomes. So this axiom is stating that the ambiguity of an act only depends on the relative comparisons of outcomes or the dispersion of outcomes. Finally, this axiom does not restrict how the levels of consequences are changed as long as the changes are strict improvements in all the proper states.

The assumption that the two equally crisp acts are statewise ordered is a weak restriction. Nonordered acts have additional possibilities for different ambiguity from the trade-offs between states. So essentially, the only preferences that do not satisfy this axiom are such that every act contains different ambiguity.

An assumption closer to Certainty Independence than Axioms 1.5 and 1.6 would be to assume that there exist two crisp acts that are statewise ordered. However, this is a stronger assumption than Axioms 1.5 and 1.6 which is formalized in the next remark.

Remark 1.1 If there exist crisp acts f, g such that for all $s \in S^{\mathcal{P}}$, $f_s \succ_s g_s$, then $\succsim$ satisfies Axioms 1.5 and 1.6.

1.2.4 Existence

The previous six axioms characterize the existence of a state dependent dual-self expected utility representation for preferences.

Theorem 1.1 (Existence) The following three conditions are equivalent:

1. $\succsim$ satisfies Axioms 1.1-1.6.
2. There exists $(u, \mathbb{P})$ that is a state dependent tight dual-self expected utility for $\succsim$ such that
$$\bigcap_{s \in S} u_s\Big(\Delta(X_s)\Big) \neq \varnothing.$$
3. There exists $(u, \mathbb{P})$ that is a state dependent tight dual-self expected utility for $\succsim$ and $f \in H$ such that for all $p, q \in \bigcup_{P \in \mathbb{P}} P$, $\sum_{s \in S} p_s u_s(f_s) = \sum_{s \in S} q_s u_s(f_s)$.

This result shows that the Axioms 1.1-1.6 characterize the existence of a state dependent dual-self expected utility with the additional property that there exists an act without uncertainty about its expected utility. A representation with this additional property always has an alternative state dependent tight dual-self expected utility with a utility overlap as shown in the second condition. This additional property follows from Axiom 1.5.

This result shows the generality of the state dependent dual-self expected utility. Assume that the preferences satisfy Axioms 1.1-1.4 that are standard assumptions in the ambiguity literature. Then the preferences have a state dependent dual-self expected utility representation if there exist some well-behaving acts such that one of them is unambiguous and two of them share the same ambiguity and are statewise ordered. This statewise ordering is a weak restriction as discussed above. In other words, the only standard ambiguity preferences that do not have a dual-self representation are such that every act is ambiguous or, essentially, every act contains different ambiguity. We will study in Section 1.3 a generalization that might have uncertainty about every act.

On the other hand, this axiomatization shows how the state dependent dual-self expected utility generalizes the state independent dual-self expected utility by relaxing Monotonicity and Certainty Independence. First, Monotonicity and Risk Independence as a part of Certainty Independence is relaxed to weak statewise Monotonicity and weak Risk Independence on lotteries within each state of the world. These generalizations capture the original intuitions of Monotonicity and Risk Independence that preferences are well-behaving on lotteries and there is no taste uncertainty since the statewise comparisons are done on lotteries within each state.

Second, for the rest of Certainty Independence instead of assuming which acts do not have uncertainty, we infer from the decision maker's behavior which acts are crisp and equally crisp. Our generalization of Certainty Independence is assuming that there exist a crisp act and two statewise ordered equally crisp acts. These existential generalizations can be interpreted as an assumption on the decision maker's behavior or on the structure of ambiguity under consideration as discussed in the previous section.

The next remark formalizes the connection of this result to the state independent case by considering what additional assumptions would be required for state independent utility.

Remark 1.2 Assume a state independent setting where for all $s \in S$, $X_s = X$. Then there exists a (non-trivial) state independent dual-self expected utility for $\succsim$ if and only if the following conditions hold.

1. $\succsim$ satisfies Axioms 1.1-1.6.
2. $\succsim$ satisfies monotonicity with between states comparisons: For all $s, s' \in S, f, g \in H, x, y \in \Delta(X)$,
$$(x_s, f_{-s}) \succ (y_s, f_{-s}) \Longrightarrow (x_{s'}, g_{-s'}) \succsim (y_{s'}, g_{-s'}),$$
3. There exists a constant act $z \in \Delta(X)$ such that z is a crisp act.
4. There exist constant acts $x, y \in \Delta(X)$ such that $x \succ y$ and x and y are equally crisp.

This result shows that the state dependent dual-self expected utility relaxes the monotonicity assumption to statewise monotonicity and does not restrict which acts are crisp or equally crisp. An example of a state dependent dual-self expected utility that satisfies the first two conditions but not necessarily the other ones is if all the state dependent utilities are positive affine transformations of each other.

1.2.5 Identification

This section provides the main result of the paper. We introduce a novel simple axiom stating that the decision maker has full-dimensional uncertainty. This axiom characterizes the full identification of state dependent utilities and probabilities in the state dependent dual-self expected utility. Additionally, we characterize the identification of probabilities for a single state, the identification of a relative likelihood between two states, and the general partial identification of the representation when there does not exist uncertainty about every state of the world.

Full Identification

To make the underlying intuition for our identifying condition clear, we first derive it informally from the idea that the decision maker's uncertainty about the proper states is full-dimensional. Full-dimensional uncertainty means that there is uncertainty about the likelihood ratio between any proper states.[13] Then especially there is uncertainty about every trade-off between proper states. Now consider two acts f and g that have trade-offs between proper states that is the act f is better than g in some states and g is better than f in some other states. Then under full-dimensional uncertainty, there is uncertainty about the trade-offs between f and g and uncertainty about their relative value. This uncertainty about their relative value can be observed as a difference in the ambiguity hedging of f and

[13]Formally full-dimensional uncertainty states that the set of probabilities has a non-empty interior. So especially there is uncertainty about the likelihood ratio between any states.

g when mixing with some act h. This is formalized in the next identification axiom when f and g are indifferent to guarantee observability in the difference for ambiguity hedging.[14]

Axiom 1.7 If $f, g \in H$ are such that $f \sim g$ and there exist $s, s' \in S$ such that $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$, then there exist $h \in H$ and $\alpha \in (0,1)$ such that

$$\alpha h + (1-\alpha) f \not\sim \alpha h + (1-\alpha) g.$$

The main result of this paper is the following uniqueness result for the state dependent dual-self expected utility. The result states that Axiom 1.7 characterizes the separation and identification of the probabilities and state dependent utilities.

Theorem 1.2 (Full Identification) Let $(u, \mathbb{P})$ be a state dependent tight dual-self expected utility for $\succsim$. The following four conditions are equivalent:

1. $\succsim$ satisfies Axiom 1.7.
2. $\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \cup_{P \in \mathbb{P}} P$ has a non-empty interior in $\Delta(S^{\mathcal{P}})$.
3. If $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self expected utility for $\succsim$, then there exist $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}$ such that for all $s \in S^{\mathcal{P}}$,
$$\widetilde{u}_s = \alpha u_s + \beta.$$
4. If $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self expected utility for $\succsim$, then
$$\overline{\mathrm{co}} \bigcup_{\widetilde{P} \in \widetilde{\mathbb{P}}} \widetilde{P} = \overline{\mathrm{co}} \bigcup_{P \in \mathbb{P}} P.$$

The second condition shows that Axiom 1.7's uncertainty about every trade-off between proper states is the behavioral characterization of full-dimensional uncertainty for the proper states.

The third and fourth conditions show that the full-dimensional uncertainty characterizes the identification of the state dependent utilities for the proper states up to a common positive affine transformation and the identification of the convex closure of the probabilities.

[14] This axiom is closely related to the notion of unambiguous preferences from Ghirardato et al. (2004). A decision maker is said to unambiguously prefer f to g if for all acts h and $\alpha \in [0,1]$, $\alpha f + (1-\alpha) h \succsim \alpha g + (1-\alpha) h$. Hence Axiom 1.7 states that if there are trade-offs between two acts across states and they are indifferent, then the acts are not unambiguously indifferent.

That is we have recovered the state independent identification. The equivalency between identifying the utilities and the probabilities shows they can be identified separately but identifying one of them means that the other one is also identified. This theorem with the existence theorem, Theorem 1.1, provides the foundations for state dependent utility under uncertainty. Next, we discuss this result.

First, this identification result has important special cases. This result shows the identification for state dependent maxmin expected utility (Gilboa and Schmeidler, 1989), which will be discussed more in Section 1.4, for state dependent α-maxmin expected utility (Ghirardato et al., 2004), and for state dependent Choquet expected utility (Schmeidler, 1989).

Second, the full-dimensionality of the set of probabilities for the proper states is a generic property for a given utility function.[15] Thus it is not a restrictive condition but rules out important special cases: If $\succsim$ satisfies Independence axiom, then Axiom 1.7 will not be satisfied. This is the observation from Section 1.1.1 that state dependent expected utility is not identified. More generally, if there are unambiguous events, that is each of the probabilities agrees on the probability for an event, then Axiom 1.7 will not be satisfied for these unambiguous events and the utilities and probabilities are not identified across the unambiguous event.

Third, under Axiom 7 the state dependent utilities are identified for all the non-null states and the tight set of probabilities is identified. However, the two-stage structure of probability set collections is not identified.[16] This is symmetrical to the multiplicity of state independent dual-self expected utility where only the largest representing collection of probability sets is identified (Chandrasekher et al., 2020, Proposition S.1.1). Additionally, the utilities for null states are not identifiable since they have a zero probability and do not affect preferences.

[15]Generic property in the topological sense that the closure of all the probability set collections not satisfying this property has an empty interior in the space of compact sets of $\mathcal{K}(\Delta(S))$ that is they are nowhere dense.

[16]This can be easily observed by the following: If $P \in \mathbb{P}$, $P' \supseteq P$, and P' is convex and closed, then $\mathbb{P} \cup \{P'\}$ also gives a dual-self representation since P' is dominated and never used.

Fourth, Section 1.1.1 sketched the proof for this result. The intuition underlying that example generalizes to the state dependent dual-self expected utility and to any finite number of states.

Partial Identification

The previous result showed a characterization for the full identification. However, some states of the world might not be uncertain and in this case, the full identification is not possible. The next two results study the identification of probabilities for a single state and the identification of a relative likelihood between two states.

The next result shows the identification of probabilities for a single state and how full-dimensionality can be extended to a single state by applying Axiom 1.7 only to a single state. Additionally, it provides a formalization for a notion of full-dimensionality of probabilities for a single state.

Proposition 1.3 (Probability Identification) Let $(u, \mathbb{P})$ be a state dependent tight dual-self expected utility for $\succsim$ and $s \in S^{\mathcal{P}}$. The following three conditions are equivalent:

1. If $f, g \in H$ are such that $f \sim g$, $f_s \succ_s g_s$, and there exists $s' \in S$ such that $g_{s'} \succ_{s'} f_{s'}$, then there exist $h \in H$ and $\alpha \in (0,1)$ such that
$$\alpha h + (1-\alpha) f \not\sim \alpha h + (1-\alpha) g.$$
2. $S^{\mathcal{P}} = \{s\}$ or there exist $p, q \in \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P$ such that $p_s \neq q_s$ and for all $\tilde{s} \in S \setminus \{s\}$,
$$\frac{p_{\tilde{s}}}{1-p_s} = \frac{q_{\tilde{s}}}{1-q_s}.$$
3. If $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self expected utility for $\succsim$, then
$$\left\{ \widetilde{p}_s \middle| \widetilde{p} \in \overline{\text{co}} \bigcup_{\widetilde{P} \in \widetilde{\mathbb{P}}} \widetilde{P} \right\} = \left\{ p_s \middle| p \in \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P \right\}.$$

full-dimensional uncertainty about a state s means behaviorally that there is uncertainty about any trade-offs involving state s as stated by the first condition. Our first result shows that a full-dimensional uncertainty about a state s is formalized by the second condition: uncertainty about the probability of the state s while keeping all the probabilities conditional

on the event $S \setminus \{s\}$ as constant. In other words, there exist independent uncertainty about the state s.

Our second result shows that full-dimensional uncertainty about a state characterizes the identification of the probabilities for that state. However, this time the identification of utilities for a single state is not equivalent to identifying the probabilities for that state since utilities are always identified only up to a common positive affine transformation.

This result has an important corollary: if there exists an event E such that Axiom 1.7 is satisfied for the event E: For all $f, g \in H$, $f \sim g$ and $s^* \in E, s \in S$ if $f_{s^*} \succ_{s^*} g_{s^*}$ and $g_s \succ_s f_s$, then there exist $h \in H$ and $\alpha \in (0,1)$ such that

$$\alpha h + (1-\alpha) f \not\sim \alpha h + (1-\alpha) g.$$

Then this result shows that the tight probabilities for the states in E are identified and the utilities in E are identified up to a common positive affine transformation.[17]

Next, we move on to the identification of relative likelihood of two states. When considering a gamble between two states the identification of the probabilities is not important since the gamble depends only on the relative likelihood of the states. The next result provides a behavioral condition for full-dimensional uncertainty about a relative likelihood and shows how that characterizes the identification of utilities and the relative likelihoods between two states. Additionally, it provides a formalization for a notion of full-dimensionality of relative likelihood for two states.

Proposition 1.4 (Relative Likelihood Identification) Let $(u, \mathbb{P})$ be a state dependent tight dual-self expected utility for $\succsim$ and $s, s' \in S^{\mathcal{P}}, s \neq s'$. The following four conditions are equivalent:

1. If $f, g \in H$ are such that and $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$, then there exist $h, h' \in H$ and $\alpha \in (0,1)$ such that

$$\alpha h + (1-\alpha) f \succ \alpha h' + (1-\alpha) f \text{ and } \alpha h + (1-\alpha) g \prec \alpha h' + (1-\alpha) g.$$

[17] The identification of utilities in E is not a direct corollary but follows symmetrically.

2. There exist $p, q \in \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P$ such that $\frac{p_s}{p_{s'}} \neq \frac{q_s}{q_{s'}}$ and for all $\tilde{s} \in S \setminus \{s, s'\}$, $p_{\tilde{s}} = q_{\tilde{s}}$.

3. If $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self expected utility for $\succsim$, then there exist $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}$ such that

$$\widetilde{u}_s = \alpha u_s + \beta \quad \text{and} \quad \widetilde{u}_{s'} = \alpha u_{s'} + \beta.$$

4. If $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self expected utility for $\succsim$, then

$$\left\{ \frac{\widetilde{p}_s}{\widetilde{p}_{s'}} \middle| \widetilde{p} \in \overline{\text{co}} \bigcup_{\widetilde{P} \in \widetilde{\mathbb{P}}} \widetilde{P} \right\} = \left\{ \frac{p_s}{p_{s'}} \middle| p \in \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P \right\}.$$

The first condition is a generalization of Axiom 1.7 when applied to only two states. In Theorem 1.2, Axiom 1.7 is equivalent to the first condition when applied for any states s, s' but there we could use the simpler formulation to capture all relevant trade-offs between states. However, when considering only two states s and s', this is not possible anymore since there can be trade-offs where the two acts are not indifferent. Thus the interpretation for this condition is the same as before that there is uncertainty about any trade-offs between s and s'. However, this time f and g are not indifferent so we need the stronger condition with two convex combinations to observe the uncertainty about the relative value of f and g as a difference in hedging.

Our first result provides a formalization for a notion of full-dimensional uncertainty about the relative likelihood of two states by the second condition: Uncertainty about the relative likelihood while keeping the probabilities for all the other states constant. That is there exist independent uncertainty about the relative likelihood of these states.

The second result shows that this full-dimensional uncertainty about states s and s' characterizes the identification of utilities between these states in the sense that they are identified up to a common positive affine transformation and the identification of the relative likelihoods between these states. As in Theorem 1.2 the identification of utilities and probabilities is equivalent.

Finally, we move on to the general partial identification of the state dependent dual-self expected utility. The next result shows that the value of the representation is unique up to

a positive affine transformation, but there can be additional transformations for the utilities and probabilities as long as they do not affect the value of the representation: First, for the probabilities all statewise multiplicative transformations for which the probabilities remain as probabilities are possible when we do the statewise reciprocal transformations for the utilities. Second, all statewise additive transformations for the utilities that do not affect the expected utility for any probabilities are possible. Finally, all common positive affine transformations are possible for the utilities.

Before stating the next result, we introduce some notation. For all probabilities $p \in \Delta(S)$ and numbers for each state $x \in \mathbb{R}^S_+$, we define multiplications statewise, $xp = (x_s p_s)_{s \in S}$. This induces multiplications for sets of probabilities and collections of probability sets with $x \in \mathbb{R}^S_+$. Especially for $\mathbb{P} \in \mathcal{K}(\Delta(S))$, $x\mathbb{P} = \{\{xp | p \in P\} | P \in \mathbb{P}\}$.

Theorem 1.5 (Partial Identification) Let $(u, \mathbb{P})$ be a state dependent tight dual-self expected utility for $\succsim$, and $\widetilde{u} = (\widetilde{u}_s)_{s \in S}$ be such that for all $s \in S$, $\widetilde{u}_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\widetilde{\mathbb{P}} \subseteq \mathcal{K}(\Delta(S))$ be compact and non-empty. Then $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent dual-self expected utility for $\succsim$ if and only if there exist $x \in \mathbb{R}^S_{++}, y \in \mathbb{R}^S$ such that for all $p, q \in \bigcup_{P \in \mathbb{P}} P$,

$$\sum_{s \in S} x_s p_s = \sum_{s \in S} x_s q_s =: \alpha \text{ and } \sum_{s \in S} y_s p_s = \sum_{s \in S} y_s q_s,$$

for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = \frac{1}{x_s}(u_s + y_s),$$

and for all $f \in H$,

$$\max_{\widetilde{P} \in \widetilde{\mathbb{P}}} \min_{\widetilde{p} \in \widetilde{P}} \sum_{s \in S} \widetilde{p}_s \widetilde{u}_s(f_s) = \max_{P \in \frac{x}{\alpha}\mathbb{P}} \min_{p \in P} \sum_{s \in S} p_s \widetilde{u}_s(f_s).$$

Especially, $(\widetilde{u}, \frac{x}{\alpha}\mathbb{P})$ is a state dependent tight dual-self expected utility for $\succsim$ and if $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a tight representation, then

$$\overline{\text{co}} \bigcup_{\widetilde{P} \in \widetilde{\mathbb{P}}} \widetilde{P} = \overline{\text{co}} \bigcup_{P \in \mathbb{P}} \frac{x}{\alpha} P.$$

This theorem characterizes the general partial identification. For any other state dependent dual-self representation the values of the representation are identified up to a positive affine transformation and the utilities and the tight set of probabilities are partially identi-

fied. In the above formulation, $\frac{1}{\alpha}$ is the common positive multiplicative transformations for utilities and $\frac{x}{\alpha}$ is the multiplicative transformation for the probabilities that keeps them as probabilities. These two multiplicative transformations are combined into a single term x. Similarly, the term y combines all the additive transformations for the utilities.[18]

This theorem shows how the size of the set of probabilities restricts the possible transformations. If there are more probabilities, then there are fewer transformations $x \in \mathbb{R}^S$ such that for all probabilities p and q, $\sum_{s\in S} x_s p_s = \sum_{s\in S} x_s q_s$. Especially these transformations are restricted by how many linearly independent probabilities there are in $\bigcup_{P\in\mathbb{P}} P$.[19]

The rest of this section uses the definitions from Section 1.2.3 to provide an insight into the identification through the connection to the underlying preferences. A reader interested only in the identification can skip the rest of this section.

The next result shows the connection of possible transformations to crisp acts.

Corollary 1.6 Let $(u, \mathbb{P})$ be a state dependent dual-self expected utility for $\succsim$ and $x \in \mathbb{R}^S_{++}, y \in \mathbb{R}^S$. The following two conditions are equivalent:

1. $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent dual-self expected utility for $\succsim$ such that
$$\widetilde{u}_s = \frac{1}{x_s}(u_s + y_s) \text{ for all } s \in S^{\mathcal{P}} \text{ and } 0, 1 \in \bigcap_{s\in S^{\mathcal{P}}} \widetilde{u}_s\Big(\Delta(X_s)\Big).$$
2. There exist crisp acts f^*, g^* such that $u_s(g^*_s) = -y_s$ and $u_s(f^*_s) = -y_s + x_s$.

This corollary with Theorem 1.5 shows the idea of the partial identification. The multiplicity of representations comes from the freedom of choosing any two crisp acts that are statewise ordered to provide the constant utility of zero and one. However, Theorem 1.5 extends the partial identification to situations without utility overlap or crisp acts.

[18] Denoting $p \in \bigcup_{P\in\mathbb{P}} P$, the additive transformations that do not affect the expected utility are $y - \sum_{s\in S} y_s p_s$, where the subtraction is statewise, and the common additive transformations are $\sum_{s\in S} y_s p_s$.

[19] This follows from the following observation. If $x \in \mathbb{R}^S$ and $\alpha \in \mathbb{R}$ are such that for all probabilities p, $\sum_{s\in S} x_s p_s = \alpha$, then we can write $x = \alpha + (x - \alpha)$ where for all probabilities p, $\sum_{s\in S}(x_s - \alpha)p_s = 0$. Thus we can decompose any transformation to a sum of a constant and a vector orthogonal (perpendicular) to the set of probabilities. Since $\mathbb{R}^S$ can be decomposed to the sum of the linear span of $\bigcup_{P\in\mathbb{P}} P$ and its orthogonal complement, this shows the close connection between the size of the set of probabilities and the set of possible transformations.

This corollary provides an insight into our main identification result Theorem 1.2. The full identification is only possible if all crisp acts have a constant utility. This is what Axiom 1.7 guarantees by assuming that for all pairs of crisp acts, one of the crisp acts (weakly) dominates the other one statewise.

1.3 State Dependent Dual-Self Variational Expected Utility

This section studies a more general setup with a state dependent version of dual-self variational expected utility (Chandrasekher et al., 2020). We characterize its existence and identification. The state-independent version is a general model that includes as special cases monotone mean-variance preferences (Maccheroni et al., 2009), multiplier preferences (Hansen and Sargent, 2001; Strzalecki, 2011), variational preferences (Maccheroni et al., 2006), monotonic mean-dispersion preferences (Grant and Polak, 2013), and vector expected utility (Siniscalchi, 2009). None of these representations is a special case of the previous dual-self expected utility. Our identification results in this section encompass those for state dependent versions of all the above special cases.

Dual-self variational expected utility generalizes variational expected utility. In a variational expected utility, the decision maker has an index of ambiguity attitude $c : \Delta(S) \to \mathbb{R} \cup \{\infty\}$ for each probability and the preferences are represented with utility u as

$$f \succsim g \iff \min_{p\in\Delta(S)} \sum_{s\in S} p_s u(f_s) + c(p) \geq \min_{p\in\Delta(S)} \sum_{s\in S} p_s u(g_s) + c(p).$$

This corresponds to the maxmin expected utility if c is always 0 or ∞. Here the effective domain of c denoted by $\operatorname{dom} c = \{p \in \Delta(S) | c(p) \in \mathbb{R}\}$ captures the set of subjective probabilities that the decision maker uses.

The dual-self variational expected utility generalizes this representation by allowing for multiple cost functions that are aggregated optimistically. Axiomatically, this generalizes variational preferences by relaxing Uncertainty Aversion. We will use a state dependent variation of it.

Definition Let $\succsim \subseteq H \times H$. $(u, \mathbb{C})$ is a *state dependent dual-self variational expected utility for* $\succsim$ if $u = (u_s)_{s \in S}$ and for all $s \in S$, $u_s : \Delta(X_s) \to \mathbb{R}$ is affine, $\mathbb{C} \subseteq \{c : \Delta(S) \to \mathbb{R} \cup \{\infty\} \mid c \text{ is convex}\}$ is such that $\max_{c \in \mathbb{C}} \min_{p \in \Delta(S)} c(p) = 0$ and for all $c \in \mathbb{C}$ and $p \in \operatorname{dom} c$, $s \notin S^{\mathcal{P}}$, $p_s = 0$, and for all $f, g \in H$,

$$f \succsim g \iff \max_{c \in \mathbb{C}} \min_{p \in \Delta(S)} \sum_{s \in S} p_s u_s(f_s) + c(p) \geq \max_{c \in \mathbb{C}} \min_{p \in \Delta(S)} \sum_{s \in S} p_s u_s(g_s) + c(p).$$

This representation is the variational expected utility when $\mathbb{C}$ is a singleton. Additionally, since this representation is a non-convex generalization of variational preferences, especially any uncertainty aversion and seeking of the preferences can be represented by the interplay of a cost function (min) and the set of cost functions (max).

Similarly to the dual-self expected utility, the dual-self variational expected utility is not unique. However, the smallest convex and closed effective domain of the cost functions is unique which gives us tight dual-self variational expected utility.

Definition Let $\succsim \subseteq H \times H$. $(u, \mathbb{C})$ is a *state dependent tight dual-self variational expected utility for* $\succsim$, if $(u, \mathbb{C})$ is a state dependent dual-self variational expected utility for $\succsim$ and if $(u, \widetilde{\mathbb{C}})$ is another state dependent dual-self variational expected utility for $\succsim$, then $\overline{\operatorname{co}} \bigcup_{\widetilde{c} \in \widetilde{\mathbb{C}}} \operatorname{dom} \widetilde{c} \supseteq \overline{\operatorname{co}} \bigcup_{c \in \mathbb{C}} \operatorname{dom} c$.

1.3.1 Existence, Variational

We start off with characterizing the existence of the state dependent dual-self variational expected utility. This characterization highlights the generality of the state dependent dual-self variational expected utility by showing that the only standard ambiguity preferences that do not have this representation are such that, essentially, every act contains different ambiguity. This section uses the axioms from Section 1.2.3. A reader only interested in the identification can skip this section.

The next result shows that the existence of state dependent dual-self variational expected utility is characterized by the same axioms as state dependent dual-self expected utility with the exception that there might not be crisp acts and Axiom 1.5 might not be satisfied.

Theorem 1.7 (Existence, Variational) The following two conditions are equivalent:

1. $\succsim$ satisfies Axioms 1.1-1.4, and 1.6.
2. There exists $(u, \mathbb{C})$ that is a state dependent tight dual-self variational expected utility for $\succsim$.

This result shows that a state dependent dual-self variational expected utility is characterized by relatively weak axioms capturing the dispersional nature of ambiguity without imposing any additional restrictions on it. This representation has multiple special cases as discussed above and this theorem shows the state dependent foundations for these models.

This result highlights the generality of the dual-self variational expected utility. Under the standard assumptions under ambiguity, Axioms 1.1-1.4, the only requirement for the state dependent dual-self variational expected utility is the existence of two acts that share the same ambiguity and that they are ordered statewise. As discussed in Section 1.2.3, the statewise ordering is a weak restriction. In other words, essentially the only standard preferences that do not have a state dependent dual-self variational representation are such that all the acts contain different ambiguity.

1.3.2 Identification, Variational

Next, we move on to the identification of the state dependent dual-self variational expected utility. The following identification result shows the main result of this section: Under full-dimensional uncertainty, the intensities of preferences and the probabilities can be separated.

Theorem 1.8 (Full Identification, Variational) Let $(u, \mathbb{C})$ be a state dependent tight dual-self variational expected utility for $\succsim$. The following four conditions are equivalent:

1. $\succsim$ satisfies Axiom 1.7.
2. $\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{c \in \mathbb{C}} \mathrm{dom}\, c$ has a non-empty interior in $\Delta(S^{\mathcal{P}})$.
3. If $(\widetilde{u}, \widetilde{\mathbb{C}})$ is a state dependent tight dual-self variational expected utility for $\succsim$, then there exist $\alpha \in \mathbb{R}_+$ and $B \in \mathbb{R}^S$ such that for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = \alpha u_s + B_s.$$

4. If $(\widetilde{u}, \widetilde{\mathbb{C}})$ is a state dependent tight dual-self variational expected utility for $\succsim$, then

$$\overline{\text{co}} \bigcup_{\widetilde{c} \in \widetilde{\mathbb{C}}} \text{dom}\, \widetilde{c} = \overline{\text{co}} \bigcup_{c \in \mathbb{C}} \text{dom}\, c.$$

This result states that under Axiom 1.7 the changes of state dependent utilities are identified. However, the levels are not identified and instead any additive transformations are possible. This follows from the interchangeability of cost functions and additive constants as shown in the next result. Additionally, the convex closure of the tight set of probabilities is uniquely identified as with the state dependent dual-self expected utility. The symmetric result for the state dependent dual-self expected utility Theorem 1.2 was discussed extensively in Section 1.2.5 and the same discussion applies here. Especially the non-empty interior condition is a weak restriction for state dependent dual-self variational expected utility and Axiom 1.7 is not demanding but rules out important special cases.

This theorem shows that under the more general state dependent tight dual-self variational expected utility and Axiom 1.7, the utility differences between acts are still meaningful. Especially, it is possible to study the utility gains and losses from trading one act for another one. Additionally, states of the world have meaningful probabilities which is crucial for modeling decision making under uncertainty.

Next, we move on to the general partial identification. Before stating the partial identification result, we introduce some notation. For all cost functions $c : \Delta(S) \to \mathbb{R} \cup \{\infty\}$ and numbers for each state $x \in \mathbb{R}^S_+$, we define the product $c \circ x : \Delta(S) \to \mathbb{R} \cup \{\infty\}$ as for all $p \in \Delta(S)$, $c \circ x(p) = c(xp)$ where $c(xp) = \infty$ if $xp \notin \Delta(S)$ and the product xp is the elementwise product as in Section 1.2.5. This induces products of cost sets with $x \in \mathbb{R}^S_+$ as elementwise products. Second, for $B \in \mathbb{R}^S$ define dB, the linear cost function (or a measure) associated with B, by for all $p \in \Delta(S)$, $dB(p) = \sum_{s \in S} B_s p_s$.

Theorem 1.9 (Partial Identification, Variational) Let $(u, \mathbb{C})$ be a state dependent tight dual-self variational expected utility for $\succsim$, and $\widetilde{u} = (\widetilde{u}_s)_{s \in S}$ be such that for all $s \in S$, $\widetilde{u}_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\widetilde{\mathbb{C}} \subseteq \{c : \Delta(S) \to \mathbb{R} \cup \{\infty\} | c \text{ is convex}\}$ be such that

$\max_{\widetilde{c}\in\widetilde{\mathbb{C}}} \min_{p\in\Delta(S)} \widetilde{c}(p) = 0$. Then $(\widetilde{u}, \widetilde{\mathbb{C}})$ is a state dependent dual-self variational expected utility for $\succsim$ if and only if there exist $x \in \mathbb{R}^S_{++}, B \in \mathbb{R}^S$ such that for all $p, q \in \cup_{c\in\mathbb{C}} \operatorname{dom} c$,

$$\sum_{s\in S} x_s p_s = \sum_{s\in S} x_s q_s =: \alpha,$$

for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = \frac{1}{x_s} u_s + B_s,$$

and by denoting $\beta = -\max_{c\in\mathbb{C}} \min_{p\in\Delta(S)} \alpha c(x/\alpha p) - \sum_{s\in S} p_s B_s$ for all $f \in H$,

$$\max_{\widetilde{c}\in\widetilde{\mathbb{C}}} \min_{p\in\Delta(S)} \sum_{s\in S} p_s \widetilde{u}_s(f_s) + \widetilde{c}(p) = \max_{c\in\mathbb{C}} \min_{p\in\Delta(S)} \sum_{s\in S} p_s \widetilde{u}_s(f_s) + \alpha c(x/\alpha p) - \sum_{s\in S} B_s p_s + \beta.$$

Especially $(\widetilde{u}, \alpha(\mathbb{C} \circ x/\alpha) - dB + \beta)$, where multiplication and subtractions are done elementwise for each cost function, is a state dependent tight dual-self variational expected utility for $\succsim$ and if $(\widetilde{u}, \widetilde{\mathbb{C}})$ is a tight representation, then

$$\overline{\operatorname{co}} \bigcup_{\widetilde{c}\in\widetilde{\mathbb{C}}} \operatorname{dom} \widetilde{c} = \overline{\operatorname{co}} \bigcup_{c\in\mathbb{C}} \frac{x}{\alpha} \operatorname{dom} c.$$

This theorem shows that for any other state dependent dual-self variational representation the values of the representation are identified up to a positive affine transformation. All the additional transformations are the ones that do not affect the representation: For the tight set of probabilities, all common statewise multiplicative transformations for which the probabilities still sum to a constant number are allowed when we do the statewise reciprocal transformation for the utilities and scale the probabilities to sum to 1. Additionally, additive transformations for the utilities are allowed when we do the negative of this transformation for all the cost functions and normalize the set of cost functions with β.

Symmetrical results to Propositions 1.3 and 1.4 for the identification of probabilities for a single state or for the identification of relative likelihoods between two states hold under the state dependent dual-self variational expected utility. The only difference is that the additive utility transformations are not identified anymore.

Finally, we connect the state dependent dual-self and dual-self variational expected utilities to each other by linear cost functions.

Proposition 1.10 (Connecting Dual-Self and Dual-Self Variational) The following two conditions are equivalent:

1. There exists $(u, \mathbb{C})$ that is a state dependent dual-self variational expected utility for $\succsim$ and there exists $B \in \mathbb{R}^S$ such that for all $c \in \mathbb{C}, p \in \text{dom}(c)$, $c(p) = \sum_{s \in S} B_s p_s$.
2. There exists $(\widetilde{u}, \mathbb{P})$ that is a state dependent dual-self expected utility for $\succsim$.

This proposition shows that state dependent dual-self variational expected utility with a linear cost function corresponds to state dependent dual-self expected utility. This gives a new insight into the state dependent dual-self expected utility and into unambiguous acts since these linear cost functions are underlying the intuition for unambiguous acts.[20] It is important to contrast this result to the state independent case where only state independent dual-self variational expected utility with a cost function that is constant at 0 correspond to state independent dual-self expected utility.

The rest of this section uses the definitions from Section 1.2.3 to provide an insight into the identification through the connection to the underlying preferences. A reader interested only in the identification can skip the rest of this section.

The next result connects the partial identification from Theorem 1.9 to equally crisp acts.

Corollary 1.11 Let $(u, \mathbb{C})$ be a state dependent dual-self variational expected utility for $\succsim$ and $x \in \mathbb{R}^S_{++}, B \in \mathbb{R}^S$. The following two conditions are equivalent:

1. $(\widetilde{u}, \widetilde{\mathbb{C}})$ is a state dependent dual-self variational expected utility for $\succsim$ such that
$$\widetilde{u}_s = \frac{1}{x_s}(u_s + B_s) \text{ for all } s \in S^{\mathcal{P}} \text{ and } 0, 1 \in \bigcap_{s \in S^{\mathcal{P}}}, \widetilde{u}_s\Big(\Delta(X_s)\Big).$$
2. There exist equally crisp acts f^*, g^* such that $u_s(g^*_s) = -B_s$ and $u_s(f^*_s) = -B_s + x_s$.

This corollary with Theorem 1.9 shows the idea of the partial identification. The multiplicity of representations comes from the freedom of choosing any act g^* to provide the constant utility of zero and choosing f^* such that g^* and f^* are equally crisp and statewise ordered to provide the constant utility of one.

[20] This highlights how mixing with unambiguous or crisp acts changes ambiguity linearly that is the underlying property in the definition of crisp acts in Section 1.2.3.

This highlights the difference to the identification of state dependent dual-self expected utility. Corollary 1.6 showed that the partial identification of dual-self expected utility comes from crisp acts. However, if an act is equally crisp with a crisp act, then both of the acts are crisp acts. So the only difference between these two identifications is that with dual-self expected utility one of the acts has to be crisp, whereas with dual-self variational expected utility there might not exist crisp acts and we are free to choose one of the acts.

Symmetrically to the dual-self expected utility, this corollary provides an insight into the full identification of dual-self variational expected utility, Theorem 1.8. This identification is only possible if the utility differences between all equally crisp acts are constant. This is what Axiom 1.7 guarantees by assuming that in every pair of equally crisp acts, one of the acts (weakly) dominates the other one statewise.

1.4 Application to Social Choice Theory

This section presents an application to social choice theory for state dependent utility models beyond uncertainty by changing the interpretations of the primitives of the models. In this application, we study the preferences of the society over distributions of goods to members of the society. We show the identification of fairness of the society and interpersonal utility comparisons using state dependent maxmin expected utility.

In this context, we will interpret states of the world to be the members of the society, acts to be distributions of goods to each member, and $\succsim$ to represent the preferences of the society. Here, ambiguity reflects the fairness of the distribution. We focus on a state dependent maxmin expected utility (u, P) with a symmetric set of probabilities that is for all permutations of states[21] π, if $(p_s)_{s\in S} \in P$, then $(p_{\pi(s)})_{s\in S} \in P$. This symmetry property captures anonymity of the societal preferences.

In the current context, this symmetric state dependent maxmin expected utility is a generalized Rawlsian social welfare function. Here, the set of probabilities captures the fairness

[21] $\pi : S \to S$ is a permutation if it is a one-to-one (i.e. bijective) function.

of the society. The least unfairness-averse generalized Rawlsian function is the utilitarian social welfare function which is the average utility. This corresponds to a state dependent expected utility with a uniform probability. The most unfairness-averse generalized Rawlsian function is the Rawlsian social welfare function which is the minimum utility. This corresponds to state dependent maxmin expected utility with a full set of probabilities $\Delta(S)$.

The main result of this section characterizes the existence of a generalized Rawlsian social welfare function and its identification. Especially we show that the utilitarian social welfare function is an unidentified knife-edge case of a generalized Rawlsian social welfare function: All representations except this knife-edge case are fully identified. First, we interpret shortly the axioms from Section 1.2.3 in the current context and introduce a new anonymity axiom.

The society has continuous preferences (Axioms 1.1 and 1.2) and is benevolent (Axiom 1.3) caring about the well-being of each person that is the society's preferences respect individual Pareto improvements. Additionally, the society consists of rational expected utility maximizers (Axiom 1.4) and cares about the fairness of the distribution (Axiom 1.8).[22]

Axiom 1.5 assumes that there exists a fair distribution that does not hedge the fairness of any other distribution. Axiom 1.6 assumes that the unfairness is over the variation of well-being across the society. That is fairness depends only on what each person receives relative to others.

The next axiom is Uncertainty Aversion from Gilboa and Schmeidler (1989). This captures that the society is inequality averse.

Axiom 1.8 For all $f, g \in H, \alpha \in [0,1]$, if $f \succsim g$, then $\alpha f + (1-\alpha)g \succsim g$.

Finally, we have a new anonymity axiom. This axiom states that there exist a better and a worse fair distribution such that the society is indifferent on the identity of the members who receive the better or the worse fair distribution.

[22]This assumption can be also interpreted directly as an aversion to ambiguity under a Rawlsian interpretation: The preferences of the society are determined by a person behind the veil of ignorance. However, the person deciding the preferences is uncertain about their identity in the society. This uncertainty about the identity gives a rational for the uncertainty aversion.

Axiom 1.9 There exist crisp acts f^*, g^* such that $f^* \succ g^*$ and for all $\gamma \in \Delta(S)$ and permutations $\pi : S \to S$,

$$(\gamma_s f_s^* + (1-\gamma_s) g_s^*)_{s\in S} \sim (\gamma_{\pi(s)} f_s^* + (1-\gamma_{\pi(s)}) g_s^*)_{s\in S}.$$

In this axiom, γ is a lottery across the members of the society and one of the members is awarded the better fair consequence f_s^* and everybody else receives the worse fair consequence g_s^*. That is γ captures the distribution of the better and worse consequences in the society. Then the identity of the winning member does not matter for societal preferences but only the distribution of the consequences matters.

These axioms characterize the generalized Rawlsian social welfare function. That is the symmetric state dependent maxmin expected utility. As a notation, a state dependent maxmin expected utility is a state dependent dual-self expected utility with a single probability set $(u, \{P\})$ and we omit the inner brackets in the following.

Theorem 1.12 (Symmetric State Dependent MaxMin) $\succsim$ satisfies Axioms 1.1-1.6, 1.8, and 1.9 if and only if (1) or (2) of the following conditions holds:

1. For all $p \in \operatorname{int} \Delta(S)$, there exists $u = (u_s)_{s\in S}$ such that for all $s \in S$, $u_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\operatorname{int} u_s(\Delta(X_s)) \neq \varnothing$ and for all $f, g \in H$,
$$f \succsim g \iff \sum_{s\in S} p_s u_s(f_s) \geq \sum_{s\in S} p_s u_s(g_s).$$
2. There exists (u, P) that is a state dependent maxmin expected utility for $\succsim$ such that $\operatorname{int} P \neq \varnothing$, for all $p \in P$, permutations $\pi : S \to S$, $(p_{\pi(s)})_{s\in S} \in P$, and
$$\operatorname{int} \bigcap_{s\in S} u_s\Big(\Delta(X_s)\Big) \neq \varnothing.$$

Additionally, if (2) holds and $(\widetilde{u}, \widetilde{P})$ is another state dependent maxmin expected utility for $\succsim$, then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that

$$\widetilde{P} = P \text{ and } \widetilde{u}_s = \alpha u_s + \beta \text{ for all } s \in S.$$

First, this result highlights that utilitarianism is a knife-edge case of generalized Rawlsian social welfare functions where identification is not possible. There are only two possibilities:

Either the preferences have a fully non-identified utilitarian representation or a fully identified generalized Rawlsian social welfare function with a non-singleton set of probabilities. In the ambiguity context, this result also highlights how state dependent expected utility is a non-identified knife-edge case of state dependent maxmin expected utility.

Second, this result provides the identification for the fairness of the society and interpersonal utility comparisons when the societal preferences are not utilitarian. The identification of fairness allows us to compare the fairness of different societies to each other independently of the members that form the society.[23] Additionally, the identification of interpersonal utilities allows us to do comparisons within the society and compare the welfare of its members. As an important special case, this result shows that the Rawlsian social welfare function is well identified if there is utility overlap.

Third, Theorem 1.2 generalizes this identification to non-symmetric generalized Rawlsian social welfare functions. In this context, the full-dimensionality axiom, Axiom 1.7, assumes that any redistribution that involves redistributing from one person to another makes the distribution of well-being less fair in some situations. Essentially this assumes that starting from a fair (crisp) distribution, redistribution decreases the fairness of the distribution. Under this full-dimensionality condition, we achieve the same identification of the fairness of the society and interpersonal utilities as above in the more general case without anonymity of the social welfare function. Especially, we can identify if the social welfare function is anonymous based on Axiom 1.9.

Remark Applying state dependent dual-self variational expected utility in this context would give the identification of fairness of the society independently of the members and the identification of interpersonal utility gain comparisons. This is a useful identification for example in the context of redistribution. In studying redistribution, the focus is on people who gain and who lose from the redistribution but not necessarily on the absolute levels of utilities that are not identified here.

[23]Under the Rawlsian interpretation, the weights that society gives to its members are the probability assessments of the person behind the veil of ignorance for their identity.

1.5 Discussion on Identification

The example in Section 1.1.1 showed how the identification in our models comes from attributing the changes in uncertainty attitude to changes in the probabilities. Next, we formalize this underlying assumption and consider state dependent Bewley preferences that highlight the differences between state dependent and state independent identifications.

First, we consider the state dependent dual-self expected utility. Here, the identification comes from the assumption that unambiguous acts do not have uncertainty in the representation. One of the characterizing axioms for this representation, Axiom 1.5, was that there exists an act that does not hedge ambiguity. That is, it is behaviorally revealed to be unambiguous as a crisp act. In the representation, these behaviorally unambiguous acts do not have uncertainty and they are the only acts that can have a constant utility.[24] Especially, under full-dimensional uncertainty, every behaviorally unambiguous act has a constant utility. Our identification results follow from this observation.

The underlying assumption here is that the decision maker considers some acts as unambiguous: The condition that an act is behaviorally revealed to be unambiguous as a crisp act is only a necessary condition for the decision maker to consider it as unambiguous. As illustrated in Proposition 1.10, there exist alternative dual-self variational representations for these preferences with linear costs functions where these behaviorally unambiguous acts are not unambiguous in the representation.

Next, we move on to the state dependent dual-self variational expected utility that has a similar underlying assumption. Here, the identification uses a pair of acts that hedge ambiguity always similarly. That is they are behaviorally revealed to share the same ambiguity as equally crisp acts.[25] In the representation, there is no uncertainty about the utility difference between these equally crisp acts and especially under full-dimensionality, these acts have a

[24]Formally, this is shown in Corollary 1.6.

[25]The existence of this pair of equally crisp is one of the characterizing properties for this representation and formalized in Axiom 1.6.

constant utility difference.[26] In this case, the underlying assumption is that the decision maker considers some acts as sharing the same ambiguity. Without this underlying assumption, these preferences have more general representations where the acts that behaviorally share the same ambiguity do not share the same ambiguity in the representation. Especially, the probabilities and the intensities of utilities would not be separated and identified.

These underlying assumptions have a counterpart in the state independent representations. In the state independent case, the underlying assumption is that constant acts, giving the same consequence in every state, are unambiguous for every decision maker (Ghirardato and Marinacci, 2002). This gives the identification since the ambiguity of other acts can be measured relative to these known unambiguous constant acts. Our underlying assumptions relax this.

Finally, consider the possibility that the decision maker has also uncertainty about tastes. Then the above example's change in uncertainty attitude could also be attributed to changes in tastes weakening the identification. In this case, acts that do not hedge ambiguity would not necessarily be unambiguous since the lack of hedging could also reflect certainty about tastes or utility for these acts. Karni (2020) has recently illustrated this lack of identification with rank dependent probabilities when state dependent utilities are also rank dependent.

1.5.1 State Dependent Bewley Expected Utility

Next, we move on to state dependent Bewley expected utility that highlights the difference between state independent and state dependent identifications. With state independent utility, the decision maker's uncertainty about states can be inferred from the Independence satisfying core of the preferences.[27] These are incomplete preferences that have a multi-prior Bewley representation (Bewley, 2002) as defined below. However, this identification

[26] Formally, this is shown in Corollary 1.11.

[27] Formally, this is the unambiguous preference $\succsim^*$ that is defined as for all acts $f, g \in H$,

$$f \succsim^* g \iff \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h \text{ for all } h \in H, \alpha \in (0,1].$$

Ghirardato et al. (2004) show that $\succsim^*$ is the largest Independence satisfying restriction of $\succsim$ and captures the beliefs and tastes of the decision maker.

of uncertainty does not extend to state dependent setting as we show next with a state dependent Bewley representation.

Definition (u, P) is a *state dependent Bewley expected utility for* $\succsim$ if $u = (u_s)_{s\in S}$ and for all $s \in S$, $u_s : \Delta(X_s) \to \mathbb{R}$ is affine and $P \subseteq \Delta(S)$ is closed, convex, and non-empty such that for each $p \in P, s \notin S^{\mathcal{P}}$, $p_s = 0$ and for all $f, g \in H$,

$$f \succsim g \iff \sum_{s\in S} p_s u_s(f_s) \geq \sum_{s\in S} p_s u_s(g_s) \text{ for all } p \in P.$$

In this representation, the set of probabilities P captures the decision maker's uncertainty. However, here, the decision maker does not aggregate the uncertainty for complete preferences. For example, in the state independent case, maxmin expected utility corresponds to cautious aggregation of the uncertainty using the minimum probability (Gilboa et al., 2010).

The next result shows the lack of identification for the state dependent Bewley representation: the set of probabilities and the state dependent utilities cannot be separated and identified. This is in contrast to the state independent Bewley representation where the set of probabilities is uniquely identified.

Theorem 1.13 (Identification, Bewley) Let (u, P) be a state dependent Bewley representation for $\succsim$, $\widetilde{u} = (\widetilde{u}_s)_{s\in S}$ be such that for all $s \in S$, $\widetilde{u}_s : \Delta(X_s) \to \mathbb{R}$ is affine, and $\widetilde{P} \subseteq \Delta(S)$ be closed, convex and nonempty. Then $(\widetilde{u}, \widetilde{P})$ is a state dependent Bewley representation for $\succsim$ if and only if there exist $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that

$$\widetilde{P} = \left\{ \left(\frac{a_s^{-1} p_s}{\sum_{s'\in S} a_{s'}^{-1} p_{s'}} \right)_{s\in S} \middle| p \in P \right\} \text{ and for all } s \in S^{\mathcal{P}},\ \widetilde{u}_s = a_s u_s + b_s.$$

This result highlights the difference in the identifications between state dependent and state independent representations. With state dependent utilities, the largest Independence satisfying core of the preferences does not allow us to separate the probabilities and the utilities. Instead, the intensities of preferences and the probabilities are interchangeable. This follows from the observation that the incomplete preferences do not reveal which acts are unambiguous. Instead, any act could be unambiguous.

In contrast, in the state independent case, the Independence satisfying core allows us to separate beliefs and tastes. This follows from the underlying assumption that constant acts are unambiguous. This highlights how the state dependent identification is more general since we do not make any assumptions on which acts are unambiguous. Instead, we infer it from the decision maker's behavior using the violations of the Independence axiom. These violations are not revealed in the Bewley preferences that satisfy Independence axiom and linearity.

In Appendix Appendix A.6, we also axiomatically characterize the existence of state dependent Bewley expected utility. Especially, with state independent prizes,[28] we show that the only axiom we relax is monotonicity with between states comparisons as stated in Remark 1.2 in Section 1.2.4.

1.6 Related Literature

State independent dual-self and dual-self variational expected utilities were introduced in Chandrasekher et al. (2020) building on Ghirardato et al.'s (2004) approach of using Clarke derivatives to capture the beliefs of the decision maker. These models are non-convex generalizations of multiple prior preferences (Gilboa and Schmeidler, 1989) and variational preferences (Maccheroni et al., 2006).

The previous literature on state dependent utility has mainly focused on axiomatizing the state dependent expected utility using some additional information: Karni (2007) assumes preferences on conditional acts which are acts conditional on a given event happening. Karni (2011a; 2011b) assumes preferences conditional on signals and uses updating of probabilities for identification. Karni et al. (1983) and Karni and Schmeidler (2016) assume preferences on hypothetical lotteries that are lotteries on state-consequence pairs. Lu (2019) establishes uniqueness for the utilities up to a common positive multiplication and addition of any

[28] Here, for all $s \in S$, $X_s = X$.

state specific constants by using two different random choice data based on updated random beliefs.

Chew and Wang (2020) exemplifies that state and rank dependent expected utility can be identified under two states of the world. State and rank dependent expected utility is a special case of state dependent dual-self expected utility. Karni (2020) shows that in state and rank dependent expected utility with rank-dependent probabilities the utilities and probabilities are not identified.

Drèze (1958; 1961; 1987; 2004) studies state dependent maxmax expected utility when the acts are lotteries of Anscombe-Aumann acts in the context of moral hazard. He characterizes the existence and the uniqueness of the state dependent maxmax expected utility when the intersection of utilities has a non-empty interior.[29] In contrast, we characterize the uniqueness of any state dependent dual-self expected utility in the standard state dependent Anscombe-Aumann setting and characterize behaviorally when the utilities and the set of probabilities are fully or partially identified. Baccelli (2019) discusses Drèze's contribution extensively.

Hill (2019) studies state dependent maxmin expected utility such that the best and the worst acts have a constant utility in the state independent Anscombe-Aumann setting without Risk Independence. He assumes that the best and the worst acts are crisp or unambiguous acts. In contrast under Risk Independence, we show the full generality of the dual-self and the maxmin expected utilities as assuming that there exist some unambiguous act and two acts that share the same ambiguity. Hill (2019) shows the identification of the representation when the best and the worst acts have a constant utility and the representation is linear between these best and worst acts. Under these restrictions, we have the standard state independent identification. In contrast under Risk Independence, we show the full identification of the representation among any state dependent dual-self representations which were ruled out by the above restriction.

[29]This additional property is used in the proof but not stated in the theorem.

Our application to social choice theory with symmetric state dependent maxmin contributes to the literature on preference aggregation. Harsanyi's (1953; 1955; 1977) seminal contribution axiomatized utilitarian aggregation of preferences. We generalize this aggregation to generalized Rawlsian social welfare functions. Grant et al. (2010) considers this axiomatization and generalizes it to generalized utilitarianism where the society has different risk attitudes over individual's utilities. Second, we contribute to the literature discussing the utility comparability requirements for Harsanyi's axiomatization. This is discussed in Sen (1986) and Weymark (1991). Third, our application contributes to the literature on generalizing Rawlsian social welfare function for a common class of social welfare functions with the utilitarian social welfare function. Symmetric state dependent maxmin generalizes mixed utilitarian–maximin welfare function that is a convex combination of Rawlsian and utilitarian social welfare functions as suggested in Roberts (1980). This was axiomatized in Bossert and Kamaga (2020) assuming that utilities are observable. We additionally show the identification of this representation and axiomatize its generalization when utilities are not observable.

1.7 Conclusion

The assumption of state independent utilities has been the simplifying, but non-ideal, assumption to separate the subjective probabilities from the utilities. This paper provided a novel foundation for state dependent utility by studying models of non-expected utilities. We showed that with a state dependent version of general dual-self expected utility (non-convex multiple priors preferences, Chandrasekher et al., 2020) state dependent utilities and probabilities can be identified. Additionally, with a state dependent version of more general dual-self variational expected utility (non-convex variational preferences, Chandrasekher et al., 2020) the intensities of preferences and probabilities can be separated; however, the levels of utilities cannot be identified. These identifications are characterized by full-dimensional set of probabilities.

These identifications encompass those for state dependent versions of all the special cases and alternative representations of dual-self and dual-self variational expected utilities: maxmin expected utility (multiple prior preferences, Gilboa and Schmeidler, 1989), Choquet expected utility (Schmeidler, 1989), α-maxmin expected utility (Ghirardato et al., 2004), invariant biseparable preferences (Ghirardato et al., 2004; Amarante, 2009), monotone mean-variance preferences (Maccheroni et al., 2009), multiplier preferences (Hansen and Sargent, 2001; Strzalecki, 2011), variational preferences (Maccheroni et al., 2006), monotonic mean-dispersion preferences (Grant and Polak, 2013), and vector expected utility (Siniscalchi, 2009).

We provided new characterizations for state dependent dual-self and dual-self variational expected utilities. First, we generalized standard preferences under ambiguity to state dependent setting. Under these preferences, state dependent dual-self and dual-self variational expected utilities are characterized by assumptions on the structure of ambiguity: dual-self variational was characterized by an axiom stating that each act does not contain a different source of ambiguity but instead there exist two acts that are statewise ranked and share exactly the same sources of ambiguity. This was operationalized as the two acts hedging ambiguity similarly. Dual-self expected utility was characterized by an additional assumption stating that there exists a least ambiguous act that shares the sources of ambiguity with all the other acts. This was operationalized as the least ambiguous act not hedging ambiguity.

Finally, we showed that these identifications have applications beyond ambiguity. We applied it to social choice theory. There we provided the identification of fairness of the society and for interpersonal utility comparisons and showed that a utilitarian social welfare function is a non-identified knife-edge case of Rawlsian type social welfare functions.

Chapter 2

On Preference for Simplicity, Probability Weighting, and Expected Utility

2.1 Introduction

In choice under risk, a preference for simple lotteries with simple distributions of possible prizes is well-documented (e.g. Huck and Weizsäcker, 1999; Kovářík et al., 2016) and has been suggested as an explanation for some observed paradoxical behavior (Stodder, 1997; Bernheim and Sprenger, 2020). Yet there are very few axiomatic models capturing a preference for simplicity. This paper provides an axiomatic behavioral foundation for preference for simplicity and highlights the close connection to probability weighting. Second, we provide a novel axiomatization for expected utility with a weak cancellation axiom and either first order stochastic dominance or topological continuity. This highlights the necessity of relaxing continuity and stochastic dominance in simplicity type preferences.

For our preference for simplicity, we consider a decision maker who cancels out in lotteries common consequences with common probabilities. After cancellations, the lotteries are evaluated without the canceled common consequences using conditional lotteries. Under

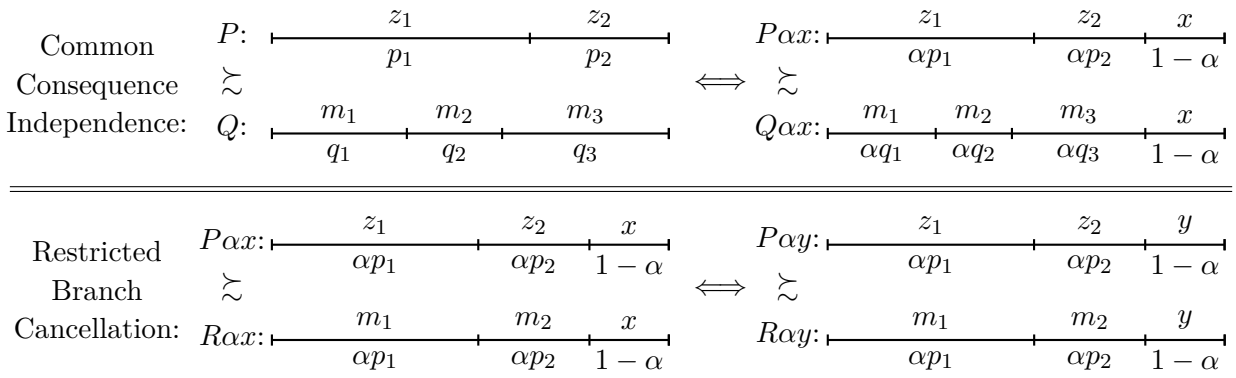

Figure 2.1. Examples of Common Consequence Independence and Restricted Branch Cancellation. Lottery P consists of different prizes z_i with probabilities p_i that sum to 1. Similarly, lottery Q consists of different prizes m_i with probabilities q_i. Lottery R shares the same probabilities as P but with different prizes m_i. Prizes x, y are different from z_i, m_i. $\alpha \in (0, 1)$. $P\alpha x$ denotes the lottery $\alpha P + (1 - \alpha)\delta_x$.

continuity assumptions, we provide a complete characterization for this decision maker. We show that either the decision maker is an expected utility maximizer with an additive cost of complexity where the complexity of the lottery is measured by Shannon's entropy (Shannon, 1948) or the decision maker is a power weighted expected utility maximizer that weights probabilities by a power function. This provides behavioral foundations for the cost of complexity that generalizes the expected utility by a single parameter that is the scale of entropy cost or the degree of probability weighting. Additionally, this highlights the close connection between the cost of complexity and power probability transformations. Especially, this result suggests that subcertainty resulting from subadditivity of probability weighting is capturing a proportional cost of complexity.

More specifically, we consider choice under risk on monetary lotteries. We study a decision maker who satisfies a Common Consequence Independence axiom exemplified in the first part of Figure 2.1. In the figure, we have two lotteries P that gives the prizes z_i with the probabilities p_i and Q that gives the prizes m_i with the probabilities q_i. We consider a probabilistic mixture of the lotteries P and Q with a common prize x that is different from z_i and m_i as on the right-hand side on the figure. The axiom states that the preference between P and Q is independent of this mixing with a common consequence x.

The standard Independence axiom assumes the above independence of mixing for any prize x that might be a part of P or Q. Our assumption relaxes Independence axiom by assuming that the decision maker might not be able to split probabilities to cancel out any common consequences. Instead, only common consequences with common probabilities are canceled. This enables a preference for lotteries with simple prize distributions without making any assumptions on what constitutes a simple prize distribution by relaxing Independence axiom's invariance to splitting probabilities and invariance to the resulting more complex lotteries.

Under continuity assumptions, we provide a full characterization for decision makers that satisfy this assumption. We show that the decision maker is either 1) an expected utility maximizer with an entropic cost of complexity or 2) a power weighted expected utility maximizer:

1) There exist a utility u and $\gamma \in \mathbb{R}$ such that for all lotteries P and Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)u(x) - \gamma H(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x) - \gamma H(Q),$$

where $H(P) = -\sum_{x \in \operatorname{supp} P} P(x) \log P(x)$ is Shannon's entropy and $\operatorname{supp} P$ denotes the set of prizes that the lottery P gives a positive probability.

2) There exist a utility u and $\lambda > 0$ such that for all lotteries P and Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^{\lambda} u(x) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\lambda} u(x).$$

Common Consequence Independence maintains Independence axiom's invariance to scaling the lottery comparison by mixing the comparison with a common consequence x as in Figure 2.1. The above result shows how this scale invariance characterizes either entropic cost of distribution or power weighted probabilities. This highlights their behavioral foundations as an ability to focus on conditional lotteries after canceling common consequences with common probabilities.

The entropy of a lottery captures the information content of the lottery. This result suggests a close relation between probability weighting and the information content of the

lottery. In the rational inattention literature (Sims, 2003), the cost of attention is measured by the information content of the lottery. This result suggests a connection between the cost of attention and probability weighting. In this rational inattention interpretation, the lotteries are interpreted as the posterior distributions of prizes that the decision maker is making a choice on.

Second, we reconsider expected utility and provide novel simple characterizations for it. We consider a decision maker who evaluates each possible prize independently of the other possible prizes, but the evaluation might depend on the probabilities of the other prizes. We show that under continuity in prizes and stochastic dominance or joint continuity in prizes and probabilities, this decision maker is an expected utility maximizer.

More specifically, we consider a decision maker who satisfies a Restricted Branch Cancellation axiom exemplified in the second part of Figure 2.1. In the figure, we have two lotteries $P\alpha x$ and $R\alpha x$ that have 2 different prizes, z_i and m_i respectively, with common probabilities αp_i and a common prize x with a probability $1-\alpha$. The axiom states that the preference between $P\alpha x$ and $R\alpha x$ is independent of the common prize and changing the common prize to y that is different from z_i and m_i does not affect the preference.

This condition relaxes Independence axiom in three ways. First, by assuming that $P\alpha x$ and $R\alpha x$ have the same probabilities but different prizes. Second, by assuming that x, y are different from the other prizes and not a part of P or Q. Third, the decision maker does not take the conditional lottery without the canceled prizes. This relaxation allows for a limited understanding of probabilities and only assumes cancellation of common prizes with common probabilities if the probabilities of (different) prizes are the same.

However, we show that Restricted Branch Cancellation is equivalent to Independence axiom if the preferences are either continuous in prizes and satisfy first order stochastic dominance or jointly continuous in prizes and probabilities[1]. In this case, the decision

[1]Formally, topologically continuous in the weak convergence.

maker is an expected utility maximizer. This axiomatization shows the necessity of relaxing stochastic dominance and joint continuity in our generalizations of expected utility.

Finally, we consider extensions of our preference for simplicity. First, we characterize more general costs of complexity. Second, we characterize power weighted expected utility where the weights are normalized to sum to 1 and show how this is characterized by a lack of intrinsic cost of complexity.

Our characterizations contribute to the literature on non-expected utility. First, we provide behavioral foundations for entropy modified expected utility and power weighted expected utility. This foundation suggests that these models might be relevant to some applications where the decision maker has a preference for simple distributions or uses the distribution of prizes as a frame to evaluate the lottery but has difficulties in evaluating lotteries across different frames. Luce et al. (2008b) show how these models can accommodate many common choice paradoxes.

Second, we provide a novel characterization for expected utility. This characterization highlights the crucial role of Restricted Branch Cancellation and first order stochastic dominance in understanding violations of expected utility. Under first order stochastic dominance, non-expected utility models must violate Restricted Branch Cancellation. That is the decision maker's evaluation of possible prizes must depend on the other possible prizes as for example with rank dependent expected utility. On the other hand, if the decision maker evaluates each possible prize as a different possibility without interactions, then we must violate FOSD for a non-expected utility model. This contributes to the recent literature comparing and testing traditional prospect theory and cumulative prospect theory (Bernheim and Sprenger, 2020; Abdellaoui et al., 2020).

Additionally, our models of complexity contribute to the literature on the utility of gambling. In this interpretation, the cost of complexity becomes an intrinsic utility or disutility for the process of gambling that is not related to the consequences of the gambling as suggested in Conlisk (1993) and Le Menestrel (2001). Here, the process of gambling is captured

by the distribution of prizes of the lottery. Additionally, entropy captures the amount of information that the gamble produces so expected utility with entropic utility for gambling captures directly a preference or dispreference for the information contained in the gamble. The utility of gambling has been discussed in Diecidue et al. (2004) and Luce et al. (2008a; 2008b).

The idea of using power weighted utility with entropy cost of information of complexity was introduced by Meginniss (1976). Luce et al. (2008b) studied and axiomatized this type of preferences in a Savagian context when the decision maker has an additive utility over joint receipt of multiple lotteries and has a status quo prize with zero utility. This different context leads to a very different axiomatic approach with multiple behavioral assumptions for the representations and to a utility function that is only unique up to a positive scaling. In contrast, our axiomatization provides simple behavioral foundations for this type of preferences with only Common Consequence Independence and continuity in prizes and probabilities. Additionally, we show that power weighted utility with entropy cost of complexity is behaviorally equivalent to power weighted expected utility when there is no status quo prize with zero utility.

Recently, Puri (2020) studied a preference for simplicity using an expected utility with an additive cost for the number of different possible prizes. In contrast, our additive entropic cost depends on the probabilities of the prizes and on the information content of the distribution.

The remainder of the paper proceeds as follows: We begin, in Section 2.2, by introducing the setting and studying preference for simplicity. Section 2.2.2 shows our first main results and axiomatizes expected utility with entropy cost of complexity and power weighted expected utility. Section 2.3 shows our second main result and reconsiders axioms for expected utility. Sections 2.4.1 and 2.4.2 extends on our preference for simplicity by axiomatizing normalized power weighted expected utility and more general costs of complexity. Section 2.5 shows the uniqueness properties for our representations. Section 2.6 sketches the proofs



Figure 2.2. An example of lotteries $P = \sum_{i=1}^{3} p_i x_i$, $Q = \sum_{i=1}^{3} p_i y_i$ with the same co-probability (p_1, p_2, p_3) where for all $i \neq j$, $x_i \neq x_j$ and $y_i \neq y_j$.

for expected utility and entropy cost of complexity. Finally, Section 2.7 discusses related literature and Section 2.8 concludes. Proofs for all the results are in the Appendix.

2.2 Preliminaries and Preference for Simplicity

2.2.1 Preliminaries

We consider a standard setting in choice under risk with monetary prizes without compound lotteries. The prize space is an open interval of monetary prizes $X = (m_*, m^*)$ where $m_*, m^* \in \mathbb{R} \cup \{\infty, -\infty\}$ and $m_* < m^*$. Denote the simple lotteries on X by $\Delta(X)$. We consider preferences $\succsim$ on the lotteries $\Delta(X)$. $\operatorname{supp} P$ denotes the support of a lottery P.[2] We endow the set of lotteries $\Delta(X)$ with the topology of weak convergence.[3]

We define mixtures of lotteries prizewise and assume that compound lotteries are reduced to single stage lotteries: Define for all $\alpha \in [0,1]$, $P, Q \in \Delta(X)$, and $x \in X$,

$$\Big(\alpha P + (1-\alpha)Q\Big)(x) = \alpha P(x) + (1-\alpha)Q(x).$$

We discuss this definition at the end of this section.

Our approach is to consider the probabilities and the prizes of a lottery separately. This approach is illustrated in Figure 2.2. Here the line represents the probability line from 0 to 1 and each probability is a section of the line and associated with a prize. Here, we have a lottery P with probabilities (p_1, p_2, p_3) and different prizes (x_1, x_2, x_3). We call the probabilities

[2]For simple lotteries, $\operatorname{supp} P = \{x \in X | P(x) > 0\}$.

[3]For simple lotteries, weak convergence simplifies to convergence of cumulative distribution function outside the support of the limit lottery: A sequence of lotteries $(P_n)_{n=1}^{\infty} \subseteq \Delta(X)$ converges weakly to $P \in \Delta(X)$ if for all $a \in \mathbb{R}$ and $a \notin \operatorname{supp} P$,

$$\sum_{\substack{x \in \operatorname{supp} P_n \\ x \leq a}} P_n(x) \to \sum_{\substack{x \in \operatorname{supp} P \\ x \leq a}} P(x) \text{ as } n \to \infty.$$

This is equivalent to convergence for the expectation of any continuous and bounded function.

(p_1, p_2, p_3), the co-probability of a lottery P and this is a probability distribution on indexes, $\Delta(\mathbb{N})$. This co-probability captures the probabilities for different prizes independently of the prizes. It corresponds to the probabilities when P is written in a standard form as a convex combination of different degenerate lotteries $P = \sum_{i=1}^{3} p_i \delta_{x_i}$. With an abuse of notation, we omit δ for degenerate lotteries and denote $P = \sum_{i=1}^{3} p_i x_i$.

In Figure 2.2, we have another lottery, Q with the same co-probability as P but different prizes so it has a similar lottery representation as P. Formally:

Definition Lotteries P and Q have the same co-probability if there exist $p \in \Delta(\mathbb{N})$ and $(x_i)_{i \in \operatorname{supp} p}, (y_i)_{i \in \operatorname{supp} p} \in X^{\operatorname{supp} p}$ such that for all $i \neq j$, $x_i \neq x_j$, and $y_i \neq y_j$ and $P = \sum_{i \in \operatorname{supp} p} p_i x_i$ and $Q = \sum_{i \in \operatorname{supp} p} p_i y_i$.

In Figure 2.2, if $y_1 = y_2$, then the lottery Q simplifies since the probabilities p_1 and p_2 with the same prize can be joined together simplifying the graphical representation and changing the co-probability. Formally, this change follows from our definition of mixtures since the following two lotteries are identically the same

$$p_1 y_1 + p_2 y_1 + p_3 y_3 = (p_1 + p_2) y_1 + p_3 y_3.$$

Our definition of mixtures and lottery simplifications incorporate the idea of Prospect Theory's editing phase (Kahneman and Tversky, 1979). In this editing phase, the decision maker combines probabilities with the same prizes and reduces multistage lotteries to a single stage. In the literature, the first property is known as Coalescing Axiom (Birnbaum, 1999; Humphrey, 1995; Starmer and Sugden, 1993) or combination operation of the editing phase (Kahneman and Tversky, 1979) and the second property as Reduction of Compound Lotteries (Segal, 1988; 1990).

2.2.2 Entropy Modified and Power Weighted Expected Utility

We start off with the main result of the paper. We provide behavioral foundations for preference for simplicity with expected utility that is adjusted by the entropic cost of complexity and with power weighted expected utility. We show that under continuity assumptions these representations are characterized by canceling common prizes with common probabilities when comparing lotteries and viewing the lotteries as conditional lotteries without the canceled prizes.

Before defining our two representations, we denote Shannon's entropy for a lottery P by

$$H(P) = -\sum_{x \in \text{supp}\, P} P(x) \log P(x).$$

Definition $\succsim$ is *Entropy Modified Expected Utility* if there exist a utility $u : X \to \mathbb{R}$ and $\gamma \in \mathbb{R}$ such that for all $P, Q \in \Delta(X)$,

$$P \succsim Q \iff \sum_{x \in \text{supp}\, P} P(x)u(x) - \gamma H(P) \geq \sum_{x \in \text{supp}\, Q} Q(x)u(x) - \gamma H(Q).$$

This representation was introduced by Meginniss (1976), Yang and Qiu (2005), and Luce et al. (2008b).

Remark The choice of logarithm base for the entropy does not affect the representation. The base of the logarithm can be changed by a rescaling that only affects γ.

Definition $\succsim$ is *Power Weighted Expected Utility* if there exist a utility $u : X \to \mathbb{R}$ and $\lambda > 0$ such that for all $P, Q \in \Delta(X)$,

$$P \succsim Q \iff \sum_{x \in \text{supp}\, P} P(x)^{\lambda} u(x) \geq \sum_{x \in \text{supp}\, Q} Q(x)^{\lambda} u(x).$$

We characterize these two representation by the following four axioms. Our first three assumptions are standard or technical continuity assumptions that we maintain throughout the paper. The first standard assumption is that preferences are weak order and monotonic in certain amounts money.

Axiom 2.1 (Weak Order and Monotonicity) $\succsim$ is complete and transitive and for all $x > y$, $\delta_x \succ \delta_y$.

The next two axioms are technical continuity axioms in prizes and in probabilities. First, we assume continuity in prizes when the probabilities of the prizes are unchanged.

Axiom 2.2 (Continuity in Prizes) For all lotteries Q, $(P_n)_{n=1}^{\infty}$, and P such that for all $n \in \mathbb{N}$, P_n and P have the same co-probability and P_n converges weakly to P as $n \to \infty$,

$$\text{if } P_n \succsim Q \text{ for all } n \in \mathbb{N}, \text{ then } P \succsim Q$$

and

$$\text{if } Q \succsim P_n \text{ for all } n \in \mathbb{N}, \text{ then } Q \succsim P.$$

Remark The following observation makes the connection between convergence in prizes and weak convergence conditional on a co-probability more clear. If lotteries $(P_n)_{n=1}^{\infty}$, and P have the same co-probability, then P_n converges weakly to P as $n \to \infty$ iff for each $n \in \mathbb{N}$, there exist $(x_i^n)_{i \in \operatorname{supp} p}, (x_i)_{i \in \operatorname{supp} p} \in X^{\operatorname{supp} p}$ such that $P^n = \sum_{i \in \operatorname{supp} p} p_i x_i^n$, $P = \sum_{i \in \operatorname{supp} p} p_i x_i$, and for all $i \in \operatorname{supp} p$, $x_i^n \to x_i$ as $n \to \infty$ in the Euclidean topology.

The next axiom is continuity in the probabilities when the prizes are unchanged. We assume this with the standard mixture continuity.

Axiom 2.3 (Mixture Continuity) For all lotteries P, Q, R, the sets

$$\left\{\alpha \in [0,1] \middle| \alpha P + (1-\alpha)Q \succsim R\right\} \text{ and } \left\{\alpha \in [0,1] \middle| R \succsim \alpha P + (1-\alpha)Q\right\} \text{ are closed.}$$

The last axiom is the main axiom of the representation. This axiom states that when two different lotteries have a common consequence x with a common probability $1 - \alpha$, this can be canceled without affecting the comparison. After canceling it out, the lotteries are viewed as conditional lotteries without the canceled consequence. This axiom relaxes the standard Independence axiom and does not assume cancellation of common consequences with different probabilities that are less obvious to cancel. This axiom is illustrated in Figure 2.3.

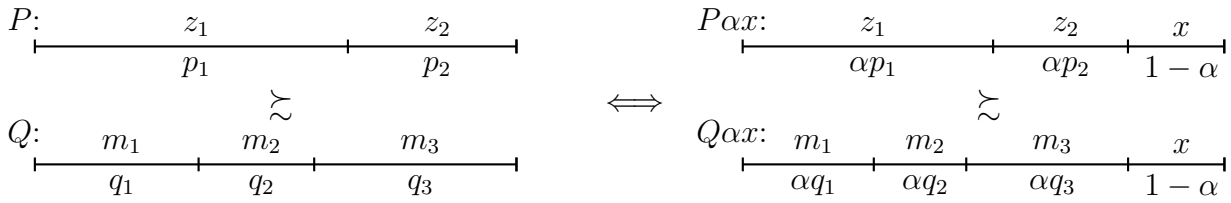

Figure 2.3. An example of Common Consequence Independence, Axiom 2.4. The lotteries are $P = \sum_{i=1}^{2} p_i z_i$, $Q = \sum_{i=1}^{3} q_i m_i$, $P\alpha x = \alpha P + (1-\alpha)\delta_x$ and $Q\alpha x = \alpha Q + (1-\alpha)\delta_x$.

Axiom 2.4 (Common Consequence Independence) For two lotteries P, Q, if $x \notin \operatorname{supp} P \cup \operatorname{supp} Q$ and $\alpha \in (0,1)$, then

$$P \succsim Q \iff \alpha P + (1-\alpha)\delta_x \succsim \alpha Q + (1-\alpha)\delta_x.$$

An alternative interpretation for this axiom is that the decision maker understands that only the relative likelihoods of the prizes matter: When canceling or adding a common consequence, the relative likelihoods of the non-common prizes are not affected. That is the comparison between P and Q does not depend on the probability scale of the comparison since rescaling the comparison by mixing with a common prize x does not affect the comparison.

This axiom has a close formal connection to C-Independence axiom in the ambiguity literature as introduced in Gilboa and Schmeidler (1989). However, in choice under risk, it captures a different type of behavior.[4]

These four axioms characterize Entropy Modified or Power Weighted Expected Utility:

Theorem 2.1 (Entropy Modified and Power Weighted EU) $\succsim$ satisfies Axioms 2.1-2.4 iff one of the following two conditions holds:

1. There exist a continuous, strictly increasing, $u : X \to \mathbb{R}$ and $\gamma \in \mathbb{R}$ such that for all lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)u(x) - \gamma H(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x) - \gamma H(Q).$$

[4]We follow the naming convention from Birnbaum (1999). This is a different axiom than the similarly named Common Consequence Independence in the context of stochastic choice (Blavatskyy, 2008) that is a stochastic version of the axiom we call Branch Cancellation.

2. There exist a continuous, strictly increasing, $u : X \to \mathbb{R}$ and $\lambda > 0$ such that for all lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^{\lambda} u(x) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\lambda} u(x).$$

Adding the following restricted first order stochastic dominance characterizes Entropy Modified Expected Utility in the above theorem.

Axiom 2.5 (Restricted FOSD) If two lotteries P, Q have the same co-probability and P first order stochastically dominates Q, then $P \succsim Q$.

This theorem provides a full characterization for preferences under Common Consequence Independence and separate continuities in probabilities and prizes. This characterization shows a simple behavioral interpretation for entropy modified and power weighted expected utility maximizers with Axiom 2.4. They have difficulties in evaluating the same prizes with different probabilities while being able to cancel out common consequences with common probabilities and being able to take the conditional lotteries without the canceled consequences. That is these decision makers understand that only the relative likelihoods of different prizes matter.

The characterization shows that this understanding of relative likelihoods gives power weighted probabilities. Here, the preferences are invariant to the choice of base for relative likelihood comparisons. This gives log-log-linearity[5] for the probability weighting that characterizes power probability weighting.

Second, this characterization without probability weighting gives an insight to entropy as an additive cost of distribution. Axiom 2.4 assumes that the additive cost of probability $(1-\alpha)$ is independent of the other probabilities in αP or αQ. That is the cost of distribution is additive across different probabilities which characterizes entropy as shown in Section 2.6. With probability weighting, these additive costs of distributions are a part of utility function as shown in the uniqueness result in Section 2.5.

[5]Formally, the function w is log-log-linear if for all $\alpha > 0$, $w(p^{\alpha}) = w(p)^{\alpha}$.

This result highlights the closeness of probability weighting and cost of complexity. Especially, it shows how the subcertainty resulting from probability weighting, that is the weighted probabilities sum to less than 1, is a form of complexity aversion and in the limit as $\lambda \to 1$ this subcertainty gives the entropic cost of complexity. Our characterization for Normalized Power Weighted Expected Utility in Section 2.4.1 illustrates this connection more.

In the context of complexity, the characterizing Axiom 2.4 can be interpreted as stating that the obvious cancellations of common consequences with common probabilities are simple and do not change the complexity of the evaluation. Additionally, the axiom is assuming that the additional complexity of the common consequence with common probability is independent of the complexity of P or Q. Thus additional complexity has a constant marginal cost. This characterizes entropy under Restricted FOSD.

Finally, this result highlights the important role of Restricted FOSD. This axiom rules out probability weighting in the above result and differentiates between entropic cost function and power weighted probabilities. However, it is an open question if this axiom is satisfied in practice.

Remark The connection between Entropy Modified and Power Weighted Expected Utility comes from the non-uniqueness of the limiting preferences as $\lambda \to 1$ since additive transformations of utilities do not cancel out. Intuitively, this comes from doing an additive transformation of utilities by $\frac{\gamma}{\lambda-1}$ and subtracting $\frac{\gamma}{\lambda-1}$ from the representation. Then we have for all lotteries P,

$$\sum_{x\in\operatorname{supp} P} P(x)^\lambda\left(u(x)+\frac{\gamma}{\lambda-1}\right)-\frac{\gamma}{\lambda-1}=\sum_{x\in\operatorname{supp} P} P(x)^\lambda u(x)+\frac{\gamma}{\lambda-1}\left(\sum_{x\in\operatorname{supp} P} P(x)^\lambda-1\right).$$

By taking $\lambda \to 1$, we have by L'Hôpital's rule

$$\lim_{\lambda\to 1}\sum_{x\in\operatorname{supp} P} P(x)^\lambda u(x)+\frac{\gamma}{\lambda-1}\left(\sum_{x\in\operatorname{supp} P} P(x)^\lambda-1\right)=\sum_{x\in\operatorname{supp} P} P(x)u(x)+\gamma\sum_{x\in\operatorname{supp} P} P(x)\log P(x).$$

Formally, as shown later in Corollary 2.8, doing an additive transformation of utilities creates generalized entropy cost of lottery.

Remark 2.1 The separate continuity assumptions on prizes and probabilities, Axioms 2.2 and 2.3, are equivalent under Axiom 2.4 to the following continuity in weighted weak convergence. Say that lotteries $(P_n)_{n=1}^{\infty}$ converge to P in weighted weak convergence if for all continuous and bounded functions $f : X \to \mathbb{R}$ and continuous and increasing functions $w : [0,1] \to [0,1]$ with $w(0) = 0$ and $w(1) = 1$,

$$\sum_{x \in \operatorname{supp} P_n} w\Big(P_n(x)\Big) f(x) \to \sum_{x \in \operatorname{supp} P} w\Big(P(x)\Big) f(x) \text{ as } n \to \infty.$$

Then $\succsim$ is continuous in weighted weak convergence if for all lotteries $(P_n)_{n=1}^{\infty}, P, Q$ such that P_n converges to P in weighted weak convergence,

$$\text{if } P_n \succsim Q \text{ for all } n \in \mathbb{N}, \text{ then } P \succsim Q$$

and

$$\text{if } Q \succsim P_n \text{ for all } n \in \mathbb{N}, \text{ then } Q \succsim P.$$

This is weaker than continuity in weak convergence as shown in the next section. But under stochastic dominance equivalent to weak convergence as shown in Remark 2.3. Without Axiom 2.4, this is stronger than Axioms 2.2 and 2.3.

2.3 Expected Utility

We move on to our second main result. We provide a novel axiomatization for expected utility using a weak form of cancellation assumption stating that the decision maker evaluates each prize independently of the other possible prizes, but the evaluation might depend on the probabilities of the other prizes. We show that under continuity in prizes and first order stochastic dominance or topological continuity on lotteries, the decision maker is an expected utility maximizer. This shows the necessity of relaxing continuity and stochastic dominance in our representations with preference for simplicity.

The cancellation axiom used in this section is Restricted Branch Cancellation. This is weaker than our other cancellation axioms, Axioms 2.4, 2.10, and 2.12. It states that for two lotteries with the same co-probability that share a common prize with a common probability,

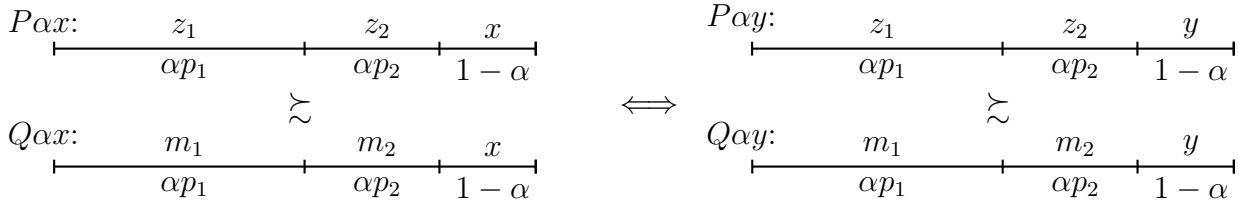

Figure 2.4. An example of Restricted Branch Cancellation, Axiom 2.6. The lotteries are $P = \sum_{i=1}^{2} p_i z_i$, $Q = \sum_{i=1}^{2} p_i m_i$, $P\alpha x = \alpha P + (1-\alpha)\delta_x$ and $Q\alpha x = \alpha Q + (1-\alpha)\delta_x$.

the common prize does not affect the choice and can be canceled. This assumes that the evaluation of prizes is independent of the other possible prizes. However, their evaluation can depend on the probability of the prize or on the probabilities of the other prizes. This axiom is illustrated in Figure 2.4.

Axiom 2.6 (Restricted Branch Cancellation) If two lotteries P, Q have the same co-probability and if $x, y \notin \operatorname{supp} P \cup \operatorname{supp} Q$ and $\alpha \in (0,1)$, then

$$\alpha P + (1-\alpha)\delta_x \succsim \alpha Q + (1-\alpha)\delta_x \iff \alpha P + (1-\alpha)\delta_y \succsim \alpha Q + (1-\alpha)\delta_y.$$

This axiom does not make any restrictions on how different probabilities are compared and requires only a limited understanding of probabilities as different possibilities for prizes. This axiom relaxes the Independence axiom in three ways. First, after canceling the common prizes with common probabilities, the lotteries might not be viewed as conditional lotteries without the canceled prizes. This relaxes the scaling property of Axiom 2.4. Second, the cancellations are done only when the lotteries P and Q share the same co-probabilities. Third, only common prizes with common probabilities are canceled and the decision maker might not be able to split probabilities to cancel all common prizes. However, we show that this axiom is equivalent to the Independence axiom under topological continuity on lotteries or continuity in prizes and first order stochastic dominance.

A stronger non-restricted version of this axiom was suggested by Kahneman and Tversky (1979) as a cancellation operation in the editing phase of lotteries. This axiom is formally related to Weak C-Independence in choice under uncertainty (Maccheroni et al., 2006).[6]

[6]This axiom is also called Restricted Branch Independence in Birnbaum (2008) and Restricted Common Branch Substitution in Luce et al. (2008b). Luce and Marley (2005) call an equivalent axiom Restricted Branch Cancellation.

2.3.1 Expected Utility and First Order Stochastic Dominance

Our first characterization for expected utility shows that if the decision maker satisfies Restricted Branch Cancellation, first order stochastic dominance, and a weak form of continuity in shifting all the prizes, then the decision maker is an expected utility maximizer. Our first axiom is continuity in shifting all the prizes of the lottery. First, we define a prize shift.

Definition For a lottery P and $\varepsilon \in \mathbb{R}$, define a lottery $P + \varepsilon$ by for all $x \in X$

$$P + \varepsilon(x + \varepsilon) = P(x).$$

Here, we shift the prize distribution of P by increasing every prize by ε. That is the probability of prize $x + \varepsilon$ is $P(x)$. Since X is an open interval, for a small enough ε, we have $P + \varepsilon \in \Delta(X)$. The next axiom is continuity in shifting all the prizes of a lottery.

Axiom 2.7 (Prize Shift Continuity) For all lotteries $P, Q \in \Delta(X)$, if $P \succ Q$, then there exist $\varepsilon, \gamma > 0$ such that $P - \varepsilon, Q + \gamma \in \Delta(X)$ and

$$P - \varepsilon \succ Q + \gamma.$$

This axiom weakens our previous prize continuity, Axiom 2.2, by only considering uniform shifts of prizes. This continuity allows for example rank dependent discontinuities.

The last axiom for expected utility is first order stochastic dominance.

Axiom 2.8 (FOSD) For all lotteries P, Q, if P first order stochastically dominates Q, then $P \succsim Q$.

The next result shows that under continuity in prizes and stochastic dominance, Restricted Branch Cancellation is equivalent to standard Independence axiom.

Theorem 2.2 (EU and FOSD) The following two are equivalent:

1. $\succsim$ satisfies Axioms 2.1 and 2.6 to 2.8.

2. There exists a strictly increasing and continuous $u : X \to \mathbb{R}$ such that for all lotteries $P, Q \in \Delta(X)$,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)u(x) \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x).$$

This result highlights the strength of the first order stochastic dominance. It gives a novel characterization for expected utility with only consistency and continuity in prizes and monotonicity with stochastic dominance. This also highlights the prevalence of violations of stochastic dominance in models of non-expected utility that satisfy Restricted Branch Cancellation such as subjectively weighted expected utility (Hong, 1983; Fishburn, 1983). The intuition for this result is discussed after our next result that characterizes expected utility with weak convergence continuity.

Remark In the proof, Appendix B.1, we consider general prize intervals. In this case, for the shift continuity, we use censored shifts that are censored to remain in the closure of the prize interval. We show that for a compact prize interval X, Theorem 2.2 generalizes to general Borel lotteries. Additionally, we show that with simple lotteries, Theorem 2.2 generalizes to any prize interval X with a possibly unbounded utility function.

Remark 2.2 Under FOSD, Prize Shift Continuity and Prize Continuity are equivalent without any Independence assumption: Assume $\succsim$ satisfies Axioms 2.1 and 2.8. Then $\succsim$ satisfies Axiom 2.2 iff $\succsim$ satisfies Axiom 2.7.

Remark 2.3 Additionally, under FOSD, Weak Convergence Continuity and Weighted Weak Convergence as defined in Remark 2.1 are equivalent without any Independence assumptions for any prize interval X. Assume $\succsim$ satisfies Axioms 2.1 and 2.8 and X is an interval. Then $\succsim$ is continuous in weak convergence, Axiom 2.9, iff $\succsim$ is continuous in weighted weak convergence as defined in Remark 2.1.

2.3.2 Expected Utility and Continuity

Second, we show how Restricted Branch Cancellation with topological continuity on lotteries gives expected utility. The next axiom is a joint continuity in prizes and probabilities using continuity in weak convergence.[7]

Axiom 2.9 (Weak Convergence Continuity) For all lotteries Q, $(P_n)_{n=1}^{\infty}$, and P such that P_n converges weakly to P as $n \to \infty$,

$$\text{if } P_n \succsim Q \text{ for all } n \in \mathbb{N}, \text{ then } P \succsim Q$$

and

$$\text{if } Q \succsim P_n \text{ for all } n \in \mathbb{N}, \text{ then } Q \succsim P.$$

Our next result shows that under monotonicity in certain money and continuity in weak convergence, Restricted Branch Cancellation is equivalent to Independence axiom.

Theorem 2.3 (EU and Continuity) The following two are equivalent:

1. $\succsim$ satisfies Axioms 2.1, 2.6, and 2.9.
2. There exists a strictly increasing, bounded, and continuous $u : X \to \mathbb{R}$ such that for all lotteries $P, Q \in \Delta(X)$,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)u(x) \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x).$$

This result highlights the strength of the joint continuity in prizes and probabilities. It gives a novel characterization for expected utility with only consistency in prizes and continuity in prizes and probabilities. This also highlights the prevalence of violations of continuity models of non-expected utility that satisfy Restricted Branch Cancellation.

As discussed later in Section 2.6.2, the joint continuity in prizes and probabilities rules out preference for simplicity by making the preferences continuous when joining different probabilities with the same prizes. This gives expected utility representation without assuming

[7]Continuity in weak convergence is a standard assumption in choice under risk and used e.g. in Segal (1990), Border and Segal (1994), Castagnoli and Calzi (1996), Machina (2001), Cerreia-Vioglio et al. (2015; 2020), and Hara et al. (2019).

linearity by instead using uniform lotteries and symmetry since all the uniform lotteries are connected by joining probabilities. Additionally, these uniform lotteries can approximate any lottery.

Remark In the proof, Appendix B.1, we show that Theorem 2.3 generalizes to any interval X and to general Borel lotteries over any interval.

2.4 Extensions on Preference for Simplicity

Next, we consider extensions of Theorem 2.1 by a preference for simplicity when an intrinsic cost of complexity is ruled out and with more general forms of costs functions.

2.4.1 Normalized Power Weighted Expected Utility

We start off by ruling out any intrinsic cost of complexity. We show that this corresponds to Power Weighted Expected Utility when the weights are normalized to sum to one. This representation does not have subcertainty in the probability weighting. This highlights the connection between cost of complexity and subcertainty. This probability weighting was suggested by Karmarkar (1978).

First, we relax Common Consequence Independence to the following Restricted Common Consequence Independence. This axiom assumes that the decision maker cancels out common consequences with common probabilities and takes the conditional lotteries without the canceled consequence if the co-probabilities of the lotteries are the same. That is the decision maker might not be able to scale different distributions similarly. In the context of complexity, this is assuming that the cost of complexity in scaling up the distribution is the same regardless of the prizes. However, scaling more complex distributions might be more difficult and costly than scaling simple distributions. This axiom is illustrated in Figure 2.5.

Axiom 2.10 (Restricted Common Consequence Independence) If two lotteries P, Q have the same co-probability and if $x \notin \operatorname{supp} P \cup \operatorname{supp} Q$ and $\alpha \in (0,1)$, then

$$P \succsim Q \iff \alpha P + (1-\alpha)\delta_x \succsim \alpha Q + (1-\alpha)\delta_x.$$

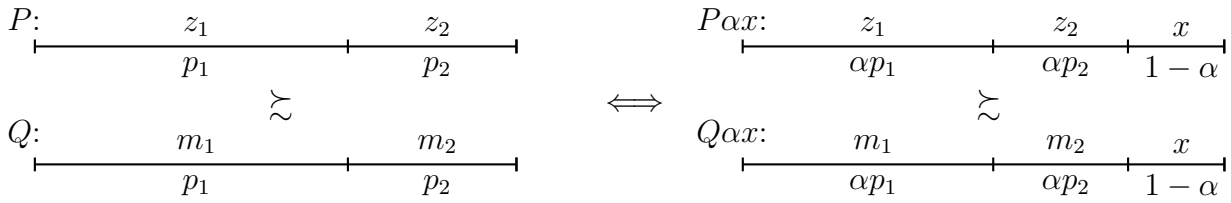

Figure 2.5. An example of Restricted Common Consequence Independence, Axiom 2.10. The lotteries are $P = \sum_{i=1}^{2} p_i z_i$, $Q = \sum_{i=1}^{2} p_i m_i$, $P\alpha x = \alpha P + (1-\alpha)\delta_x$ and $Q\alpha x = \alpha Q + (1-\alpha)\delta_x$.

The next assumption is the main feature of normalized probability weights: We assume that a certainty equivalent of a lottery is between its smallest and the largest prize. This is a weak form of stochastic dominance.

Axiom 2.11 (Certainty Equivalent Bounds) For all lotteries P and $x \in X$,

$$\delta_x \sim P \Rightarrow \max\{y \in \operatorname{supp} P\} \geq x \geq \min\{z \in \operatorname{supp} P\}.$$

By relaxing Common Consequence Independence with the two axioms, we characterize an expected utility with normalized power probability weighting.

Theorem 2.4 (Normalized Power Weighted EU) $\succsim$ satisfies Axioms 2.1 to 2.3, 2.10, and 2.11 iff there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that for all lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) \geq \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x).$$

Comparing this result to previous Theorem 2.1, shows that the normalized probability weights and the lack of subcertainty comes from ruling out ordinal costs of different distributions. This representation only reflects the misperceptions or misprocessing of probabilities.

In the context of complexity, Certainty Equivalent Bounds rule out any intrinsic cost of complexity. This representation without subcertainty and the axiomatization by ruling out any intrinsic cost of complexity highlights how we can interpret the subcertainty of probability weighting as reflecting a proportional cost of complexity. In this context, the normalized probability weighting reflects a cardinal cost of complexity that depends on the prizes. This is absent in the entropic cost of complexity that has only an ordinal cost of complexity depending on the distribution of prizes or co-probability but not on the prizes.

Remark This theorem shows a method to axiomatize expected utility models with weighted probabilities that use normalized power weighted probabilities where the normalization depends only on the co-probability of the lottery. This includes for example, the probability weighting used in Tversky and Kahneman (1992), where there exists $\lambda > 0$ such that each probability p_i in the co-probability p is weighted as

$$\frac{p_i^\lambda}{(\sum_{j\in\operatorname{supp} p} p_j^\lambda)^{\frac{1}{\lambda}}}.$$

These models satisfy Axioms 2.1 to 2.3 and 2.10 and different variations of Certainty Equivalent Bounds characterize these models.

2.4.2 Variational Expected Utility

Finally, we study preferences with general cost functions for different distributions. First, we define our cost function as an information function that depends only on the distribution of prizes of the lottery.

Definition A function $c : \Delta(X) \to \mathbb{R}$ is an information function if for all $P, Q \in \Delta(X)$ with the same co-probability, $c(P) = c(Q)$ and for all $x \in X$, $c(\delta_x) = 0$.

Additionally, an information function $c : \Delta(X) \to \mathbb{R}$ is linearly continuous if for all $P, Q \in \Delta(X)$,

$$c(P) = \lim_{\alpha\to 1} c\Big(\alpha P + (1-\alpha)Q\Big).$$

Our representation combines power weighted probabilities and additive costs of distributions.

Definition $\succsim$ is *Power Weighted Variational Expected Utility* if there exist a utility $u : X \to \mathbb{R}$, $\lambda > 0$, and an information function $c : \Delta(X) \to \mathbb{R}$ such that for all $P, Q \in \Delta(X)$,

$$P \succsim Q \iff \sum_{x\in\operatorname{supp} P} P(x)^\lambda u(x) - c(P) \geq \sum_{x\in\operatorname{supp} Q} Q(x)^\lambda u(x) - c(Q).$$

We relax Common Consequence Independence with the previous Restricted Common Consequence Independence, Axiom 2.10, and a new Branch Cancellation axiom. This new

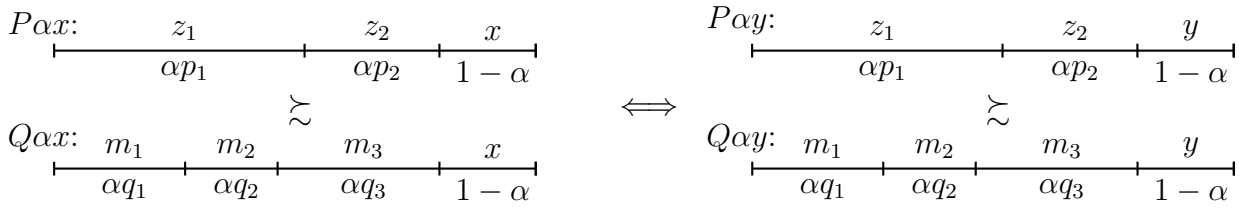

Figure 2.6. An example of Branch Cancellation, Axiom 2.12. The lotteries are $P = \sum_{i=1}^{2} p_i z_i$, $Q = \sum_{i=1}^{3} q_i m_i$, $P\alpha x = \alpha P + (1-\alpha)\delta_x$ and $Q\alpha x = \alpha Q + (1-\alpha)\delta_x$.

axiom states that the decision maker cancels common consequences with common probabilities but does not take the conditional distributions without the canceled consequences as in Common Consequence Independence. That is the decision maker evaluates each possible prize independently of the other possible prizes and the probabilities of the other prizes. This is illustrated in Figure 2.6.

Axiom 2.12 (Branch Cancellation) For two lotteries P, Q, if $x, y \notin \operatorname{supp} P \cup \operatorname{supp} Q$ and $\alpha \in (0,1)$, then

$$\alpha P + (1-\alpha)\delta_x \succsim \alpha Q + (1-\alpha)\delta_x \iff \alpha P + (1-\alpha)\delta_y \succsim \alpha Q + (1-\alpha)\delta_y.$$

As mentioned in discussing Restricted Branch Cancellation, this axiom has a close formal connection to Weak C-Independence axiom in the ambiguity literature and this axiom was suggested by Kahneman and Tversky (1979) as a cancellation operation in the editing phase of lotteries.

Our continuity assumptions with Restricted Common Consequence Independence and Branch Cancellation characterize Power Weighted Variational Expected Utility with a linearly continuous information cost function.

Theorem 2.5 (Power Weighted Variational EU) $\succsim$ satisfies Axioms 2.1 to 2.3, 2.10, and 2.12 iff there exist a continuous, strictly increasing, $u : X \to \mathbb{R}$, $\lambda > 0$, and a linearly continuous information function $c : \Delta(X) \to \mathbb{R}$ such that for all lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^{\lambda} u(x) + c(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\lambda} u(x) + c(Q).$$

Additionally, $\succsim$ satisfies Axiom 2.5 iff $\lambda = 1$.

This representation allows for general costs of distributions combined with probability weighting. This gives both an ordinal cost of distribution from the information function and a cardinal cost of distribution from the subcertainty of probability weighting. This result shows that the formally symmetrical axioms of Branch Cancellation and Weak C-Independence in choice under uncertainty (Maccheroni et al., 2006) give similar variational representations in the two different contexts.

Comparing this characterization to Theorem 2.1, shows how entropy cost comes from assuming that the comparison of lotteries does not depend on the scale of the comparison as discussed in Axiom 2.4. Here, we relax this assumption by only assuming scale invariance within the same co-probability.

Under complexity interpretation, Branch Cancellation assumes that the evaluation of consequences is independent of the complexity of the lottery. That is the complexity of a lottery is ordinal and does not dependent on the consequences. Compared to Axiom 2.4, this axiom relaxes the assumption that the cost of additional complexity is independent of the complexity of the lottery that characterized the entropy cost of complexity.

The representations and the axioms studied are summarized in Figure 2.7 in the concluding section.

2.5 Uniqueness of the Representations

Finally, we show the uniqueness of all our representations. We will only consider the uniqueness of the variational representations since it induces the uniqueness for all the other representations as different special cases. First, without probability weighting, the utility is unique up to a positive affine transformation and the information cost function is unique up to the scaling of utilities.

Theorem 2.6 (Variational EU Uniqueness) Let preferences $\succsim$, strictly increasing and continuous $u, \tilde{u} : X \to \mathbb{R}$, and linearly continuous information functions $c, \tilde{c} : \Delta(X) \to \mathbb{R}$ be such that for all lotteries P, Q,

$$\begin{aligned} P \succsim Q &\iff \sum_{x \in \operatorname{supp} P} P(x)u(x) + c(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x) + c(Q) \\ &\iff \sum_{x \in \operatorname{supp} P} P(x)\tilde{u}(x) + \tilde{c}(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)\tilde{u}(x) + \tilde{c}(Q). \end{aligned}$$

Then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that for all $x \in X, P \in \Delta(X)$,

$$\tilde{u}(x) = \alpha u(x) + \beta, \text{ and } \tilde{c}(P) = \alpha c(P).$$

Next, we consider the uniqueness of power weighted variational representation. In this case, the utilities are unique up to a positive affine transformation, but additive transformations induce changes to the cost function coming from the subcertainty of probability weighting.

Theorem 2.7 (Power Weighted Variational EU Uniqueness) Let preferences $\succsim$, strictly increasing and continuous $u, \tilde{u} : X \to \mathbb{R}, \lambda, \tilde{\lambda} > 0$ and linearly continuous information functions $c, \tilde{c} : \Delta(X) \to \mathbb{R}$ be such that for all lotteries P, Q,

$$\begin{aligned} P \succsim Q &\iff \sum_{x \in \operatorname{supp} P} P(x)^{\lambda} u(x) + c(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\lambda} u(x) + c(Q) \\ &\iff \sum_{x \in \operatorname{supp} P} P(x)^{\tilde{\lambda}} \tilde{u}(x) + \tilde{c}(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\tilde{\lambda}} \tilde{u}(x) + \tilde{c}(Q). \end{aligned}$$

Then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that for all $x \in X, P \in \Delta(X)$,

$$\tilde{\lambda} = \lambda, \tilde{u}(x) = \alpha u(x) + \beta, \text{ and } \tilde{c}(P) = \alpha c(P) + \Big(1 - \sum_{x \in \operatorname{supp} P} P(x)^{\lambda}\Big)\beta.$$

As a corollary, this gives the uniqueness for power probability weighted expected utility. To state it, we define a generalized entropy of order $\lambda > 0, \lambda \neq 1$[8] for lottery P is

$$H^{\lambda}(P) = \frac{1}{\lambda - 1}\Big(\sum_{x \in \operatorname{supp} P} P(x)^{\lambda} - 1\Big).$$

This converges to Shannon's entropy as $\lambda \to 1$.

[8]This is known also as Tsallis entropy.

Corollary 2.8 Let preferences $\succsim$, strictly increasing and continuous $u, \tilde{u} : X \to \mathbb{R}$, $\lambda, \tilde{\lambda} > 0$, $\lambda, \tilde{\lambda} \neq 1$, and $\tilde{\gamma} \in \mathbb{R}$ be such that for all lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^{\lambda} u(x) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\lambda} u(x)$$

$$\iff \sum_{x \in \operatorname{supp} P} P(x)^{\tilde{\lambda}} \tilde{u}(x) + \tilde{\gamma} H^{\tilde{\lambda}}(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\tilde{\lambda}} \tilde{u}(x) + \tilde{\gamma} H^{\tilde{\lambda}}(Q).$$

Then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that for all $x \in X$,

$$\tilde{\lambda} = \lambda, \tilde{u}(x) = \alpha u(x) + \beta, \text{ and } \tilde{\gamma} = (1 - \lambda)\beta.$$

A representation with $u : X \to \mathbb{R}, \lambda > 0, \gamma \in \mathbb{R}$ such that for all lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^{\lambda} u(x) + \gamma H^{\lambda}(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\lambda} u(x) + \gamma H^{\lambda}(Q)$$

was introduced in Luce et al. (2008b) as entropy-modified linear weighted utility. Corollary 2.8 shows that any entropy-modified linear weighted utility with $\lambda \neq 1$ has an alternative equivalent representation as a Power Weighted Expected Utility. In here, the entropy arises from additive transformations of utility since they do not cancel out across different lotteries with weighted probabilities.

2.6 Proof Sketches

2.6.1 Proof Sketch for Entropy Modified Expected Utility

First, we have additive representation for preferences conditional on a co-probability $p = (p_1, \ldots, p_n)$. We keep the co-probability fixed and consider different prizes. Here we can think of each probability as a state or a dimension. First, by Common Consequence Independence (CCI), these conditional preferences are separable[9]. With continuity, this is

[9]Preferences on X^n is separable, if for all $1 \leq i \leq n$, $x, y \in X^n$, $z_i, z_i' \in X$

$$(z_i, x_{-i}) \succsim (z_i, y_{-i}) \iff (z_i', x_{-i}) \succsim (z_i', y_{-i})$$

where e.g. (z_i, x_{-i}) denotes an element of X^n that has in coordinate i, z_i and in all other coordinates j, x_j.

Debreu's characterization (Debreu, 1960; Wakker, 1988) for an additive representation with different utilities in each state, $u_1^p, \ldots, u_n^p$. Now we have,

$$\sum_{i \in \operatorname{supp} p} p_i x_i \succsim \sum_{i \in \operatorname{supp} p} p_i y_i \iff \sum_{i \in \operatorname{supp} p} u_i^p(x_i) \geq \sum_{i \in \operatorname{supp} p} u_i^p(y_i).$$

Second, preferences conditional on different co-probabilities are connected. If in the preferences conditional on a co-probability $(p_1, \ldots, p_n)$, we keep the prize in the first state p_1 as fixed, then this is the same as preferences conditional on the co-probability $\left(\frac{p_2}{1-p_1}, \ldots, \frac{p_n}{1-p_1}\right)$ by CCI. This gives that all representations conditional on different co-probabilities must use the same utilities and so there exist $w : [0,1] \to [0,1]$ and $u : X \to \mathbb{R}$ such that

$$\sum_{i \in \operatorname{supp} p} p_i x_i \succsim \sum_{i \in \operatorname{supp} p} p_i y_i \iff \sum_{i \in \operatorname{supp} p} w(p_i) u(x_i) \geq \sum_{i \in \operatorname{supp} p} w(p_i) u(y_i).$$

Next, the weighting function is a power function. As discussed after Axiom 2.4, the decision maker understands that only the relative likelihoods of the prizes matter. That is only the relative likelihoods affect the weighting function w. However, the choice of base for relative likelihoods does not matter for w and so for all $p, q \in (0,1), p < q$, $w(\frac{p}{q}) = \frac{w(p)}{w(q)}$. This log-log-linearity characterizes power functions and there exists $\lambda > 0$ such that $w(p) = p^\lambda$. Restricted FOSD rules out probability weighting for $\lambda = 1$. Now conditional on each co-probability, the preferences have an expected utility representation.

Next, we connect the preferences conditional on different co-probabilities into a single representation. By applying CCI twice, we have for all lotteries $P, Q, \alpha \in (0,1), x, y \in\notin \operatorname{supp} P \cup \operatorname{supp} Q$,

$$\alpha P + (1-\alpha)x \succsim \alpha Q + (1-\alpha)x \iff \alpha P + (1-\alpha)y \succsim \alpha Q + (1-\alpha)y.$$

This states that the intensities of utilities conditional on different co-probabilities are the same but their levels might differ. So there exists a function on co-probabilities $c : \Delta(\mathbb{N}) \to \mathbb{R}$ such that

$$\sum_{i \in \operatorname{supp} p} p_i x_i \succsim \sum_{i \in \operatorname{supp} q} q_i y_i \iff \sum_{i \in \operatorname{supp} p} p_i u(x_i) + c(p) \geq \sum_{i \in \operatorname{supp} q} q_i u(y_i) + c(q).$$

We omit the last index of c and normalize c with $c(\delta_1) = 0$.

Finally, c is Shannon's entropy. Denote by $\mathrm{CE}(P)$ the certainty equivalent of a lottery P and assume they always exist. By CCI, we have for any $p_1, p_2 \geq 0$ such that $p_1 + p_2 \leq 1$ and different prizes $x, y, z \in X$

$$p_1 z + (1-p_1)\Big(\frac{p_2}{1-p_1}x + \frac{1-p_1-p_2}{1-p_1}y\Big) \sim p_1 z + (1-p_1)\,\mathrm{CE}\left(\frac{p_2}{1-p_1}x + \frac{1-p_1-p_2}{1-p_1}y\right),$$

By the above representation, this is equivalent to

$$c(p_1, p_2) - (1-p_1)c\Big(\frac{p_2}{1-p_1}\Big) = c(p_1)$$

and symmetrically for p_2,

$$c(p_2) + (1-p_2)c\Big(\frac{p_1}{1-p_2}\Big) = c(p_1, p_2) = c(p_1) + (1-p_1)c\Big(\frac{p_2}{1-p_1}\Big).$$

This recursivity is a known functional equation characterization for Shannon's entropy (Tverberg, 1958; Aczél and Daróczy, 1975).

2.6.2 Proof Sketch for Expected Utility with Continuity

The proof idea for the expected utility characterization is simple.

First, following the first step of Section 2.6.1 conditional on a co-probability, the preferences are separable by Axiom 2.6 and have an additive representation.

Second, dyadic co-probabilities[10] use the same utilities: By symmetry, each dimension uses the same utility. By continuity and joining probabilities continuously, the utilities across dyadic co-probabilities must be the same but only normalized differently to sum to 1. This gives an expected utility representation for dyadic lotteries.

Finally, we can use dyadic lotteries to approximate any lottery.

2.7 Related Literature

The idea of using expected utility with entropy cost of information or complexity was introduced in Meginniss (1976). Meginniss characterized power weighted expected utility with

[10]Co-probability p is dyadic if there exists $n \in \mathbb{N}$ such that p is uniform lottery with 2^n outcomes, $p = (2^{-n})_{i=1}^{2^n}$.

a generalized entropy as an additive cost by a recursivity of the functional form but did not provide a behavioral interpretation for the assumption. Yang and Qiu (2005) studied entropy-modified expected utility as a measure of risk.

Luce et al. (2008a; 2008b; 2009) study expected utility with entropy cost function and with and without power probability weighting as in Theorem 2.1 and Corollary 2.8. They study this representation in a Savagian setting with induced lotteries. Additionally, they assume that the decision maker can receive multiple gambles simultaneously and assume additivity of utilities over joint receipt of lotteries and that there exists a neutral status quo prize. Luce et al. axiomatize expected utility with entropy cost in this context. Due to the different settings, there is a limited connection between this axiomatization and the axiomatization provided here. In their setting utilities are identified up to a positive scaling due to neutral status quo prize and our Uniqueness results do not hold in that setting. Especially, Corollary 2.8 does not hold and in their setting, power weighted expected utility with generalized entropy cost is not behaviorally equivalent to power weighted expected utility. Additionally, they apply the entropy representation to different paradoxes and show how it explains them.

Puri (2020) studies preference for simplicity on lotteries using an expected utility with an additive cost of complexity when the cost of complexity of lottery depends only on the number of possible outcomes of the lottery or on the support of the lottery that gives a discontinuous model of complexity. Here, we focus on the cost of complexity as captured by the information content of the distribution that is continuous in the distribution.

Ortoleva (2013) studies thinking aversion in menu choices. Here the complexity of a menu is captured by the number of possible outcomes.

Valenzuela-Stookey (2020) and Saponara (2020) study complexity in choice under uncertainty by modelling a decision maker who has limitations in evaluating complex alternatives that are captured by the number of outcomes in Valenzuela-Stookey (2020) or by a given well understood collection of events in Saponara (2020).

Prelec (1998) and Luce (2001) axiomatizes expected utility with two parameter compound-invariance family probability weighting. This includes non-normalized power weighting that are combined with exponential transformations.

Our expected utility characterization is related to the literature characterizing expected utility by substitution methods. Arrow (1951) and Luce and Raiffa (1957) use Common Consequence Independence, Axiom 2.4, and continuity to characterize expected utility. Wakker (2010) use Common Consequence Independence and first order stochastic dominance for the characterization.

Birnbaum (2018) documents the experimental evidence on violations of Restricted Branch Cancellation and first order stochastic dominance. There is additionally Allais-type evidence on violations of Common Consequence Independence documented in Birnbaum (1999).

Complexity aversion has been documented in Huck and Weizsäcker (1999), Mador et al. (2000), Sonsino et al. (2002), Moffatt et al. (2015), and Kovářík et al. (2016). Carvalho and Silverman (2019) show how complexity has larger effects on people with less sophisticated choices.

2.8 Conclusion

This paper has studied choice under risk when the decision maker has a preference for simplicity. The axioms and the models studied are summarized in Figure 2.7.

First, we studied preference for simplicity. Here, we considered a decision maker who cancels out in lotteries common consequences with common probabilities and evaluates them using conditional lotteries without the canceled consequences. Under a continuity assumption, we showed that either this decision maker is an expected utility maximizer with an additive cost of complexity where the complexity of the lottery is measured by Shannon's entropy or the decision maker is a power weighted expected utility maximizer. Especially, this result suggested a close connection between probability weighting and cost of complexity: subcertainty resulting from subadditivity of probability weighting is capturing a proportional

Axiom	Representation	Utility of Lottery P
Common Consequence Independence	Entropy Modified EU or Power Weighted EU	$\sum P(x)u(x) - \gamma H(P)$ or $\sum P(x)^\lambda u(x)$
Branch Cancellation; Restricted Common Consequence Independence	Power Weighted Variational EU	$\sum P(x)^\lambda u(x) - c(P)$
Restricted Branch Cancellation	Under Weak Convergence Continuity or FOSD, EU	$\sum P(x)u(x)$
Certainty Equivalent Bounds; Restricted Common Consequence Independence	Normalized Power Weighted EU	$\frac{\sum P(x)^\lambda u(x)}{\sum P(y)^\lambda}$

Figure 2.7. Summary of the independence type axioms studied and the representations they characterize under continuities in prizes and probabilities, Axioms 2.1 to 2.3. $H(P)$ denotes the entropy of P. $\gamma \in \mathbb{R}, \lambda > 0$, and c is a function that only depends on the probabilities of P but not on the prizes. The arrows on the axioms signify implication relations.

cost of complexity. Finally, we generalized this result to more general forms of cost of complexity and considered the case of normalized probability weights without any intrinsic cost of complexity.

Second, we reconsidered expected utility. We considered a decision maker whose evaluation of prizes does not depend on the other possible prizes but might depend on their distribution. We showed that if the preferences are either continuous in prizes and satisfy first order stochastic dominance or jointly continuous in prices and probabilities, then the decision maker is an expected utility maximizer. This highlighted the necessity of relaxing topological continuity and stochastic dominance for our non-expected utility representations.

Chapter 3

The Foundations of Measuring (Cardinal) Rationality

3.1 Introduction

One of the most fundamental assumptions in economics is the assumption of rationality. Yet, a substantial body of literature (e.g. Echenique et al. (2011), Choi et al. (2014), Dean and Martin (2016)) has questioned its validity. This lack of rationality is unsettling since virtually every study of individuals' consumption behavior assumes rationality in the form of utility functions. Hence it is fundamental to understand how close people's behavior is to the assumption of rationality. Is the assumption of rationality a better approximation for some groups than others? Or are some markets more suitable for economic analysis than others?

Over the past decades, several measures of rationality have been proposed to quantify the extent of rationality violations. However, almost all the measures are ad hoc measures without any theoretical foundations. This makes it difficult to compare them to each other in order to choose the appropriate measure for evaluating the suitability of economic analysis.

This paper provides axiomatic foundations for the most used measures of rationality and compares them based on simulated choices. This offers both theoretical and practical perspectives into the choice of an appropriate measure of rationality.

First, this paper provides novel axiomatic foundations for the most used cardinal measures of rationality capturing the intensity of the rationality violations. We axiomatize them in the standard budget-choice setting with finite choices. Exploiting this theoretical framework, we highlight the commonalities and the differences of the measures; especially, they differ only on a few critical properties.

Second, this paper provides novel simulations for the measures of rationality. Previously computational constraints have restricted the computations of measures, but we provide a new efficient method to compute them that enables simulations. We do simulations based on a calibrated model of choice mistakes. We compare the performance of the measures at capturing the size of these choice mistakes. We show that the best measure at capturing the choice mistakes crucially depends on the fat-tailedness of the mistake distribution.

Finally, exploiting the simulations we test new variations of the measures of rationality. We first test if giving more weight to larger mistakes improves the performance of the measures. Surprisingly, overweighting larger mistakes makes the measures perform worse; a single large mistake is not more indicative of lower rationality than multiple small mistakes if the total size of mistakes is the same.

We then test if combining the cardinal intensity measures of rationality with ordinal counting measures and overweighting small mistakes would improve their performance. We find that, based on the simulation comparisons, adding some weight to small mistakes is always beneficial. That is multiple small mistakes is more indicative of lower rationality than a single large mistake of the same size. Especially if the distribution of choice mistakes is fat-tailed, then the best measures use lexicographically ordinal and cardinal measures prioritizing the ordinal measures. This shows a novel perspective to the connection between ordinal and cardinal measures of rationality. The value of cardinal information of rationality

violations depends on how fat-tailed or dispersed the choice mistakes are. Especially the choice of an appropriate measure depends on the application at hand.

This paper contributes to the literature studying the axiomatic foundations of rationality measures following Apesteguia and Ballester (2015). Second, we contribute to the computational study of rationality measures.

The remainder of the paper proceeds as follows: We begin, in Section 3.1.1, by highlighting the importance of choosing the appropriate measure with an empirical application. Next, Section 3.2 introduces the cardinal measures of rationality and axiomatizes them. Next, we move on to computational results in Section 3.3. First, we introduce our model of choice mistakes, how to compare measures independently of their units, and how the sizes of mistakes are calibrated in Sections 3.3.1, 3.3.2, and 3.3.3 respectively. Next, in Section 3.3.4, we compare the measures of rationality to each other. Section 3.3.5 considers different variations of the measures by overweighting either small or large mistakes. Section 3.3.6 discusses the computational aspects of the simulations. Finally, Section 3.4 discusses related literature and Section 3.5 concludes. Additional simulations and proofs are in the Appendix.

3.1.1 Example: The Choice of Measure Matters

We begin with an empirical application highlighting that the choice of measure of rationality matters in practice.

We use Choi et al.'s (2014) experiment on risky choice. In the experiment subjects allocated a budget between two risky accounts where either one of them was paid with 50% probability. This choice was repeated 25 times with different budgets and prices for the two accounts. We study if the assumption of rationality is a better approximations for some groups than others.[1]

In Table 3.1, we have calculated rationality scores for the subjects' choices using different measures of rationality. Here higher score means less rational choices. We study the

[1]For this application, we assume that rationality includes an assumption that there is no account preference between the two similar accounts.

Measure:	#Removed Choices		Afriat's Efficiency Ind.		Money Pump Ind.		Minimum Cost Ind.	
Focus:	#Mistaken Choices		Largest Violation		Average Violation		Total Violations	
Female	.018**	(.007)	.012	(.427)	.00012	(.982)	-.027*	(.042)
Age 35-49	.020*	(.039)	.0067	(.732)	.0057	(.431)	-.0038	(.817)
Age 50-65	.067***	(.000)	.077***	(.000)	.032***	(.000)	.042*	(.014)
Age 65+	.072***	(.000)	.081*	(.012)	.030**	(.006)	.042	(.154)
Sec. Voc. Edu	-.011	(.175)	-.021	(.211)	-.0075	(.207)	-.011	(.424)
Voc. or Uni. Edu	-.022**	(.006)	-.059***	(.001)	-.016*	(.013)	-.0064	(.730)
Income Effects	yes		yes		yes		yes	
N	1182		1182		1182		1182	

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.1. Measure is the rationality of choices in an experimental choice under risk. Higher measure signifies less rational choices. The table shows a regression of rationality of choices on demographics.

correlations between the demographics and the rationality measures using a regression and especially focus on the gender coefficient.

Here, we draw a significantly different picture depending on which rationality measure we use. The first measure on #Mistaken Choices counts how many choices were mistakes. Based on this measure, women's choices are statistically significantly less rational than men's choices. The second measure Afriat's Efficiency Index calculates the size of the largest rationality violation and the third measure Money Pump Index calculates the size of the average violation. Based on these two measures, there are no gender differences in the rationality of choices. Finally, Minimum Cost Index focuses on total rationality violations. Based on this measure, women's choices are statistically significantly more rational than men's choices.

What should we conclude about the gender differences in the rationality of choices? Different measures give contradictory pictures. This paper provides novel perspectives to choosing the appropriate measure using both theoretical and computational approaches.

3.2 Measures of Rationality and Axiomatizations

We start off with a theoretical analysis of the measures of rationality. We first define abstract rationality measure and then characterize three rationality measures: Afriat's Efficiency Index, Normalized Money Pump Index, and Normalized Minimum Cost Index.

We model choices in the standard finite budget-choice setting. At each time period t, the consumer makes non-negative purchases x_t at positive prices p_t for g goods over T time periods or observations. For simplicity, we assume that at each time period the purchased bundle is different.[2] These prices p_t and purchases x_t form data D that is a sequence of pairs for T periods

$$D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big).$$

We consider the collection of all the feasible data that could be observed. This is every finite length purchase data over some number of goods where all the purchases are different as above.

$$\mathcal{D} \coloneqq \Big\{\Big((p_1, x_1), \ldots, (p_T, x_T)\Big)\Big| \exists\, g \in \mathbb{N}, T \in \mathbb{N}, \forall\, t, p_t \in \mathbb{R}^g_{++}, x_t \in \mathbb{R}^g_{+}, \forall\, t' \neq t, x_t \neq x_{t'}\Big\}.$$

Next, we define that data is rationalizable with a utility function if the purchases can be explained by some strictly increasing utility function. That is the choices are done as if maximizing a utility function.[3]

Definition Data $\Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ with g goods is *rationalizable* if there exists a strictly increasing utility function $u : \mathbb{R}^g_+ \to \mathbb{R}$ such that for all $1 \leq t \leq T$,

$$x_t \in \arg\max \Big\{u(x)\Big| p_t \cdot x \leq p_t \cdot x_t\Big\}$$

and if $x, y \in \mathbb{R}^g_+$ are such that for all $1 \leq i \leq g$, $x_i > y_i$, then $u(x) > u(y)$.

[2]The model could be generalized to include duplicate quantities by generalizing the below definition and using revealed preference on time indices instead of quantities at the cost of using a non-standard notation.

[3]Here assuming that the utility function is strictly increasing is without loss of generality: This is equivalent to assuming that the choices are rationalizable with a locally non-satiated utility function u stating that for all $x \in \mathbb{R}^g_{++}$ and $\varepsilon > 0$, there exists $y \in \mathbb{R}^g_{++}$ such that $||x - y|| < \varepsilon$ and $u(x) < u(y)$ (Chambers and Echenique, 2016).

As is well known, the rationalizability of the choice data is characterized by the Generalized Axiom of Revealed Preferences (GARP) stating that there are no choice cycles. That is the revealed preference is acyclical as defined below. We formalize this next with two standard definitions. First, the revealed preference is a pair of weak and strict revealed preference.

Definition Let $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big) \in \mathcal{D}$. (R, P) is the *revealed preference* for D if for all $1 \leq t, t' \leq T$,

$$x_t \,\mathrm{R}\, x_{t'} \iff p_t \cdot x_t \geq p_t \cdot x_{t'}$$

and

$$x_t \,\mathrm{P}\, x_{t'} \iff p_t \cdot x_t > p_t \cdot x_{t'}.$$

Second, a cycle is a violation of transitivity of the revealed preference.

Definition Let $\Big((p_1, x_1), \ldots, (p_T, x_T)\Big) \in \mathcal{D}$. $(x_{t_1}, \ldots, x_{t_n})$ is a *cycle* if for all $1 \leq i \leq n$, when $x_{t_{n+1}} \coloneqq x_{t_1}$,

$$x_{t_i} \,\mathrm{R}\, x_{t_{i+1}} \text{ and for some } 1 \leq i' \leq n,\ x_{t_{i'}} \,\mathrm{P}\, x_{t_{i'+1}}.$$

The collection of cycles for data D up to different rotations is denoted $\mathcal{C}^D$. (R, P) is *acyclical* if $\mathcal{C}^D = \varnothing$.

Here each cycle gives a money pump cycle as a sequence of trades where the consumer always is willing to pay money to do the trade but at the end of the cycle is left with the original bundle. Especially, the choices are rationalizable if the decision maker cannot be used as a money pump.

Finally, we define a measure of rationality. This measures how far the choices are from being rationalizable or how irrational the choices are. Formally, A measure of rationality is a function from feasible data to non-negative numbers such that if the data is rationalizable, then it has a measure of zero.

Definition (Measure of Rationality) A function $I : \mathcal{D} \to \mathbb{R}_+$ is *a measure of rationality* if for all $D \in \mathcal{D}$ such that $\mathcal{C}^D = \varnothing$, $I(D) = 0$.

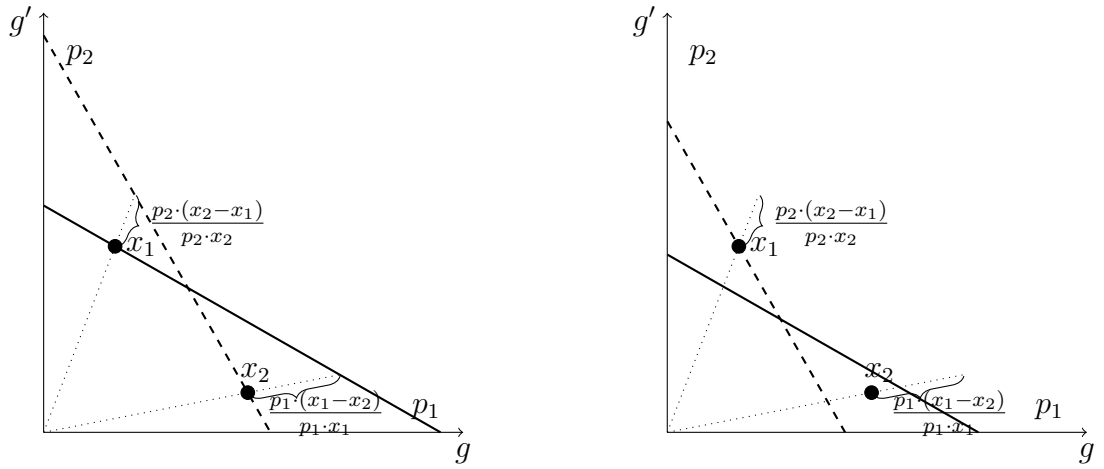

Figure 3.1. An example of calculating Afriat's Efficiency Index. We have two goods g and g'. Prices are represented by lines defining the budget sets. The data is $D = ((p_1, x_1), (p_2, x_2))$ with a cycle between x_1 and x_2. The budgets are adjusted downwards until there are no cycles left. The size of budget adjustment to break the cycles is the value of Afriat's Index.

Remark Under this definition, there can be data $D \in \mathcal{D}$ such that $\mathcal{C}^D \neq \varnothing$, but $I(D) = 0$. Additionally, if I is continuous as a function of the quantities, then almost any data with an acyclical strict revealed preference P has a measure zero, except for some technical special cases.

3.2.1 Afriat's Efficiency Index

We start off with Afriat's Efficiency Index introduced by Afriat (1967; 1972). This is the standard measure of rationality used in the literature and provides a benchmark for the other measures. We first define it and then axiomatize it. The axiomatization provides a novel insight into the measure and shows how it is essentially the only way to take an average over the rationality violations continuously.

Afriat's Efficiency Index is defined using the following relaxed revealed preference pair. For data $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ and adjustment factor $e \in [0, 1]$, define

$$x_t \, \mathrm{R}^e \, x_{t'} \iff (1-e)p_t \cdot x_t \geq p_t \cdot x_{t'}$$

and

$$x_t \, \mathrm{P}^e \, x_{t'} \iff (1-e)p_t \cdot x_t > p_t \cdot x_{t'}.$$

Afriat's Efficiency Index is

$$I_{\text{Afr}}(D) = \inf_{e\in[0,1]} e \text{ such that } (\mathrm{R}^e, \mathrm{P}^e) \text{ is acyclical.}$$

The intuition for this index is that the budgets are adjusted downwards by fraction e until there are no cycles left. This is illustrated in Figure 3.1. In the figure, we have two goods g and g' and our data consists of two observations: x_1 at the price p_1 and x_2 at the price p_2. At the first observation, everything under the budget line defined by p_1 is available and similarly at the second observation. Especially, there is a cycle between x_1 and x_2. Afriat's Efficiency Index is calculated by pulling the budget lines towards the origin until there are no cycles left. The fraction of distance the budget lines are moved towards the origin is the value of Afriat's Efficiency Index.

Remark Formally, Afriat's Efficiency Index is breaking the largest cycle at the weakest link:

$$I_{\text{Afr}}(D) = \max_{(x_{t_1},\dots,x_{t_n})\in\mathcal{C}^D} \min_{1\le i\le n} \frac{p_{t_i}\cdot(x_{t_i} - x_{t_{i+1}})}{p_{t_i}\cdot x_{t_i}}. \tag{3.1}$$

Here all the costs are calculated as fractions of income and $\max \varnothing = 0$. With this definition, this measure can be interpreted as capturing the largest fraction of income that could be extracted from the consumer from a single mistaken purchase. Here the idea is that in each choice cycle there must be at least one mistaken purchase or trade. We take a conservative estimate on which one of the trades is a mistake and use the smallest trade.

Axioms for Afriat's Efficiency Index

Next, we characterize Afriat's Index with the following five axioms. Our first axiom assumes that the measure of rationality satisfies price homogeneity of degree zero in each period's prices.

Axiom 3.1 (Strong Homogeneity) For all $\alpha \in (0,\infty), 1 \le t \le T$,

$$I\Big((p_1,x_1),\dots,(p_t,x_t),\dots,(p_T,x_T)\Big) = I\Big((p_1,x_1),\dots,(\alpha p_t,x_t),\dots,(p_T,x_T)\Big).$$

This axiom is illustrated in Figure 3.2. The figure illustrates that by doubling the prices at observation x_2, the budget set does not change. Instead, the choice and the set of possible

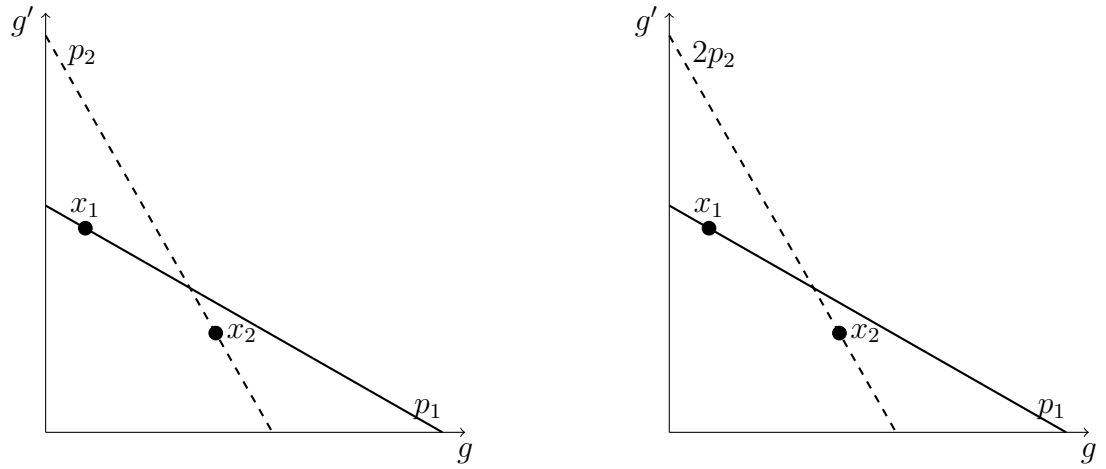

Figure 3.2. An example of Strong Homogeneity. The data is $D = \big((p_1, x_1), (p_2, x_2)\big)$ and $D' = \big((p_1, x_1), (2p_2, x_2)\big)$. By Strong Homogeneity $I(D) = I(D')$.

alternatives are exactly the same. This shows that multiplying the prices should not affect the measure of rationality: The measure of rationality measures the distance of the observed choices to rationality. That is to the collection of choices that could have resulted from utility maximization over the budget sets. Especially, these choices can only depend on the budget sets or the sets of possible alternatives. When the prices for a single period are doubled, then any rationalizable choice cannot change since the budget sets remain the same. Thus doubling the price level for a single period does not change the chosen bundle or the collection of rationalizable choices; especially the distance between the chosen bundle and the collection of rationalizable choices cannot change. Hence also the measure of rationality cannot change.

The second axiom states that violations of rationality are continuous in quantities.

Axiom 3.2 (Continuity) Let $(p_1, \dots, p_T)$ be prices and for each $n \in \mathbb{N}$, let $\Big((p_1, x_1^n), \dots, (p_T, x_T^n)\Big) \in \mathcal{D}$ and $\Big((p_1, x_1), \dots, (p_T, x_T)\Big) \in \mathcal{D}$ be such that for all $1 \leq t \leq T$, $\lim_{n\to\infty} x_t^n = x_t$.[4] Then

$$\lim_{n\to\infty} I\Big((p_1, x_1^n), \dots, (p_T, x_T^n)\Big) = I\Big((p_1, x_1), \dots, (p_T, x_T)\Big).$$

[4]Here, the convergence is in the Euclidean topology. Since the prices are fixed, all the quantities use the same number of goods.

This axiom continues on the idea that the measure of rationality is measuring the distance to rationalizable choices. The idea is that the distance is measured in terms of chosen quantities. This makes the measure continuous in quantities for given prices and given the number of observations.

However, this continuity axiom is strong: if a cycle contains a single trade with equality, then it must have measure 0 since we can make the equal revealed preference vanish by slightly perturbing the quantities. Thus, especially some consumers have measure 0 but their choices cannot be represented with a utility function. In this case, the revealed preference (R, P) is not rationalizable, but the strict revealed preference P is acyclical.[5] This axiom is stating that the distance to acyclical revealed preference is the same as the distance to acyclical strict revealed preference.[6]

Next, using Strong Homogeneity, we can focus on unit income normalized data. This allows us to compare monetary amounts across periods since in each period the income is the same.

Definition $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ is *unit income normalized* if for all $1 \leq t \leq T$,

$$p_t \cdot x_t = 1.$$

The next assumption on monotonicity continues on the idea of measuring the distance to satisfying GARP, that is to having an acyclical revealed preference. The axiom states that if two data sets have the same income and the same cycles but the cycles are always bigger for one of the data sets, then the data set with the bigger cycles is less rational.

Axiom 3.3 (Monotonicity) If $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$, $D' = \Big((p'_1, x'_1), \ldots, (p'_T, x'_T)\Big)$ are unit income normalized data such that

$$(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D \iff (x'_{t_1}, \ldots, x'_{t_n}) \in \mathcal{C}^{D'}$$

and for all $(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D$, $1 \leq i \leq n$,

$$p'_{t_i} \cdot (x'_{t_i} - x'_{t_{i+1}}) \geq p_{t_i} \cdot (x_{t_i} - x_{t_{i+1}}),$$

[5] Formally, (P, P) is acyclical.

[6] Except in some technical special cases.

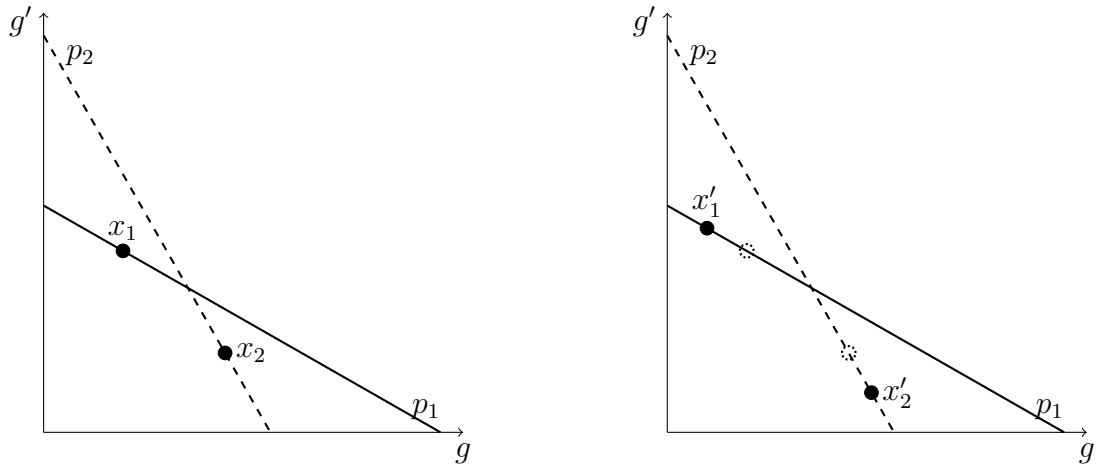

Figure 3.3. An example of Monotonicity. The data is $D = ((p_1, x_1), (p_2, x_2))$ and $D' = ((p_1, x'_1), (p_2, x'_2))$. By Monotonicity $I(D') \geq I(D)$.

then

$$I(D') \geq I(D).$$

In the axiom, the first condition states that the two unit income normalized data sets have the same cycles as a function of the indexes. The second condition states that the cycles are always bigger for the data set D'. The sizes of cycles are comparable since the income has been normalized to 1. In this situation, the axiom assumes that the choices in D' are less rational.

This axiom assumes two things. First, for a fixed data length, only the violations of GARP affect the measure of rationality. Second, within the same GARP violations that is the same ordinal violations of rationality, the size of these violations is measured as the money that an arbitrager could extract from the consumer. This axiom states that conditional on the same ordinal violations of rationality, uniformly larger cardinal violations of rationality are less rational.

The idea for the axiom is, as before, that the measure of rationality is a distance measure to satisfying the assumption of rationality. First, if a data is rationalizable depends only on if the choices satisfy GARP. Thus any distance measure to rationalizability can only depend on how close the data is to satisfying GARP, that is on the violations of GARP. Second,

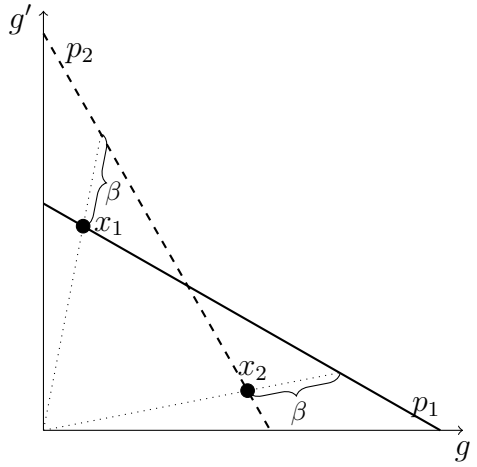

Figure 3.4. An example of Normalization. The unit income normalized data is $D = ((p_1, x_1), (p_2, x_2))$ where $p_1 \cdot (x_1 - x_2) = p_2 \cdot (x_2 - x_1) = \beta$. By Normalization $I(D) = \beta$.

conditional on the same GARP violations, increasing uniformly the size of all the trades involved in these violations increases the distance to rationalizability.

The next axiom defines the unit of the rationality measure as the average fraction of income extracted from constant cycles.

Axiom 3.4 (Normalization) If $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ is unit income normalized data such that

$$\{(x_{t_1}, \ldots, x_{t_n})\} = \mathcal{C}^D$$

and for some $\beta > 0$, for all $1 \leq i \leq n$,

$$p_{t_i} \cdot (x_{t_i} - x_{t_{i+1}}) = \beta,$$

then

$$I(D) = \beta.$$

This axiom defines the unit for the measure of rationality as constant cycles. Additionally, this axiom assumes that different length constant cycles of the same size have the same irrationality. This is a restrictive assumption but standard among measures of rationality. Especially, the unit of measurement is the average fraction extracted from constant cycles.

The last axiom assumes that the measure of rationality is focusing on the average violation across all the cycles. The axiom is stating that if we have two data sets where the second

one consist of a single constant cycle and both data are equally rational. Then increasing the amount extracted from each trade in each cycle in both data sets does not make the first data more irrational than the second data with a single cycle. This states that the measure is a sub-average across the cycles since uniformly increasing all the trades increases the measure less than increasing the equivalent average cycle by the same uniform amount.

Axiom 3.5 (Averaging) If $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ and $D' = \Big((p'_1, x'_1), \ldots, (p'_T, x'_T)\Big)$ are unit income normalized data such that for some $\beta > 0$, for all $1 \leq t, t' \leq T$, $t \neq t'$,

$$p'_t \cdot (x'_t - x'_{t'}) = p_t \cdot (x_t - x_{t'}) + \beta,$$

then

$$I(D') \leq I(D) + \beta.$$

These five axioms characterize Afriat's Efficiency Index.

Theorem 3.1 Let I be a measure of rationality on $\mathcal{D}$. I satisfies Axioms 3.1-3.5 iff

$$I = I_{\text{Afr}}.$$

This result provides a novel insight into the Afriat's Efficiency Index. It shows that under price invariance and monotonicity, Afriat's Efficiency Index is the only continuous average measure that is measured in the constant cycle units. This shows that the main restriction on Afriat's Index is that it is focusing only on the average violation. Especially, the index's extremeness of focusing only on the largest cycle arises simply from Continuity and Averaging. This result formalizes the intuition of the measure.

3.2.2 Normalized Money Pump Index

Next, we move onto another averaging measure that is a normalized version of Money Pump Index from Echenique et al.'s (2011). This measure looks at the average fraction of income that could be extracted from the consumer on average from all the violations of rationality. Formally, Normalized Money Pump Index is taking the average fraction of income that the

consumer would be willing to pay in each trade of a cycle and then taking the average of these costs of cycles:

$$I_{\text{NMPI}}(D) = \frac{1}{\mathcal{C}^D} \sum_{(x_{t_1},\ldots,x_{t_n})\in\mathcal{C}^D} \frac{1}{n}\sum_{i=1}^{n} \frac{p_{t_i}\cdot(x_{t_i}-x_{t_{i+1}})}{p_{t_i}\cdot x_{t_i}}.$$

This measure has a close connection to Afriat's Index in (3.1). Afriat's Index focused on the minimum trade of the maximum cycle. Here, we are focusing on the average trade of the average cycle.

We will show that Normalized Money Pump Index is characterized by Strong Homogeneity, Axiom 3.1, and four new axioms. The two main axioms are Strong Monotonicity that allows us to evaluate each cycle separately, the first part of below Axiom 3.3^{NMPI}, and Linearity, below Axiom 3.4^{NMPI}, which give the arithmetic averaging.

Axioms for Normalized Money Pump Index

Our first new axiom for NMPI is a weak continuity for not purchasing any goods at some time period. These zero purchases cannot be a part of any cycle.

Axiom 3.1^{NMPI} (Continuity at 0) Let $(p_1,\ldots,p_T)$ be prices, g be the number of goods, $1\leq t\leq T$, and $\Big((p_1,x_1),\ldots,(p_t,0),\ldots,(p_T,x_T)\Big)\in\mathcal{D}$. If for each $n\in\mathbb{N}$, $x_t^n\in\mathbb{R}^g_+$ is such that $\Big((p_1,x_1),\ldots,(p_t,x_t^n),\ldots,(p_T,x_T)\Big)\in\mathcal{D}$ and $\lim_{n\to\infty}x_t^n=0$, then

$$\lim_{n\to\infty} I\Big((p_1,x_1),\ldots,(p_t,x_t^n),\ldots,(p_T,x_T)\Big) = I\Big((p_1,x_1),\ldots,(p_t,0),\ldots,(p_T,x_T)\Big).$$

The second axiom combines the ideas behind Axioms 3.4 and 3.5. It states that the measure of rationality is measured in constant cycle units and takes an average across cycles. That is if every cycle is a constant cycle of the same size, then it is also the measure of rationality.

Axiom 3.2^{NMPI} (Normalization and Averaging) If $D=\Big((p_1,x_1),\ldots,(p_T,x_T)\Big)$ is unit income normalized data such that $\mathcal{C}^D\neq\varnothing$, for some $\beta>0$, for all $(x_{t_1},\ldots,x_{t_n})\in\mathcal{C}^D$ and $1\leq i\leq n$,

$$p_{t_i}\cdot(x_{t_i}-x_{t_{i+1}})=\beta,$$

then

$$I(D) = \beta.$$

As mentioned previously this axiom is assuming that the unit of measurement for the rationality measure is the average fraction extracted from constant cycles and it is taking the average across cycles. This combines Axioms 3.4 and 3.5, but this is not formally comparable to Axiom 3.5 without Continuity.

For the next monotonicity assumption, we first define two ways to measure dominance across cycles. The first way is elementwise dominance. That is if two cycles are of the same length and each trade in one of the cycles is bigger than in the other one, then the bigger cycle dominates the other one.

Definition (Cycle Dominance) Let D, D' be data and $c \coloneqq (x_{i_1}, \ldots, x_{i_n}) \in \mathcal{C}^D$ and $c' \coloneqq (x'_{i'_1}, \ldots, x'_{i'_n}) \in \mathcal{C}^{D'}$. Define the ordering

$$c \geq c' \iff \forall\, 1 \leq j \leq n, p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}}) \geq p'_{i'_j} \cdot (x'_{i'_j} - x'_{i'_{j+1}})$$

The second way is uniform dominance. This uniform dominance compares two cycles that might be of different length but if every trade in one of the cycles is bigger than any of the trades in the other cycle, then the bigger cycle uniformly dominates the other one.

Definition (Uniform Cycle Dominance) Let D, D' be data and $c \coloneqq (x_{i_1}, \ldots, x_{i_n}) \in \mathcal{C}^D$ and $c' \coloneqq (x'_{i'_1}, \ldots, x'_{i'_{n'}}) \in \mathcal{C}^{D'}$. Define the ordering

$$c \gg c' \iff \forall\, 1 \leq j \leq n, 1 \leq k \leq n', p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}}) \geq p'_{i'_k} \cdot (x'_{i'_k} - x'_{i'_{k+1}})$$

Our next axiom is that the measure of rationality respects elementwise cycle dominance for the same length cycles and uniform cycle dominance for different length cycles. That is, first, if we have two data sets with the same number of cycles and all the cycles across the data sets are of the same length but every cycle is bigger in one of the data sets than in the other one, then the data with bigger cycles is less rational. And second, if we have two data sets with the same number of cycles and all the cycles in one of the data sets are uniformly bigger than the cycles in the other data set, then the data with bigger cycles is less rational.

In contrast to the other axioms, here we allow the two data sets compared to be of different lengths.

Axiom 3.3^{NMPI} (Strong Monotonicity) If $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ and $D' = \Big((p'_1, x'_1), \ldots, (p'_{T'}, x'_{T'})\Big)$ are unit income normalized data and $\mathcal{C}^D = \{c_i\}_{i=1}^n$, $\mathcal{C}^{D'} = \{c'_i\}_{i=1}^n$, and either for all $1 \leq i \leq n$,

$$|c_i| = |c'_i| \text{ and } c_i \geq c'_i,$$

or for all $1 \leq i \leq n$,

$$c_i \gg c'_i,$$

then

$$I(D) \geq I(D').$$

This makes two assumptions. The first elementwise monotonicity is monotonicity in cycles ignoring the structure of cycles. That is it is assuming that only the cycles in the data matter for the rationality. Especially the intersections of cycles do not matter for the rationality measure. That is, doing one mistake at one observation that creates multiple cycles is as bad as doing multiple mistakes at different observations if the total number and sizes of cycles created are the same. This is the first of the two main assumptions for the normalized money pump index.

The second assumption on uniform monotonicity is monotonicity across different cycle lengths. Here, elementwise comparisons of different cycles are not possible, so instead we only assume weaker uniform monotonicity of dominance for every element. This allows us to compare different length cycles consistently. This axiom continues on the idea that the measure is focusing on the average fraction of income extracted through the cycles.

The last axiom for NMPI is a linearity axiom for equally rational data. It states that if we have two equally rational data sets using the same prices and with the same cycles, then mixing up the data sets does not make the mixed data seem more or less rational than the individual data sets if there are no new cycles created when the data sets are mixed. Before

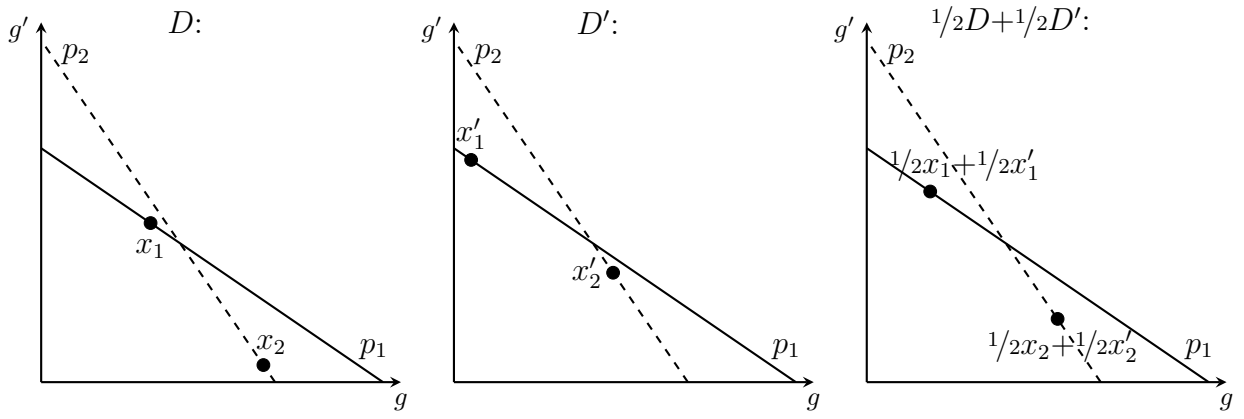

Figure 3.5. Linearity: If $I(D) = I(D')$, then $I(\frac{1}{2}D + \frac{1}{2}D') = \frac{1}{2}I(D) + \frac{1}{2}I(D') = I(D)$.

introducing the axiom, we define how to mix two data sets with the same prices as convex combinations of the quantities.

Definition If $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ and $D' = \Big((p_1, x'_1), \ldots, (p_T, x'_T)\Big)$ are two unit income normalized data with the same prices, then for all $\alpha \in [0, 1]$,

$$\alpha D + (1-\alpha)D' = \Big(\Big(p_1, \alpha x_1 + (1-\alpha)x'_1\Big), \ldots, \Big(p_T, \alpha x_T + (1-\alpha)x'_T\Big)\Big).$$

Axiom 3.4^{NMPI} (Linearity) Let $D=\Big((p_1, x_1), ..., (p_T, x_T)\Big)$ and $D'=\Big((p_1, x'_1), ..., (p_T, x'_T)\Big)$ be unit income normalized data such that

$$(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D \iff (x'_{t_1}, \ldots, x'_{t_n}) \in \mathcal{C}^{D'},$$

and $I(D) = I(D')$, then for all $\alpha \in [0, 1]$, such that $\left|\mathcal{C}^D\right| = \left|\mathcal{C}^{\alpha D+(1-\alpha)D'}\right|$,

$$I(\alpha D + (1-\alpha)D') = \alpha I(D) + (1-\alpha)I(D').$$

This axiom is the second main axiom of NMPI. In it, we have two data sets with the same prices which makes it possible to mix the data sets with each other. Then if the two data sets are equally rational, have the same sets of cycles, and there are no new cycles in the mixed data set, then the mixed data set is as rational as the original data sets. The underlying idea for this axiom is that the cycles contain the sources of irrationality. Now, if we have two data sets with the same cycles, they contain the same sources of irrationality and especially mixing the data does not affect the sources and does not make the data appear any more or less rational.

This assumption is illustrated in Figure 3.5. The figure illustrates the strength of the linearity. The choices in the data set D can be explained by one small mistake at 1 and the choices in the data set D' can be explained by one small mistake at 2. The axiom assumes that these data sets are as rational as the data set $0.5D + 0.5D'$ with two average-sized mistakes. The point of this example is that if the mistakes are done at the observation level at 1 and 2 respectively, then the data sets D and D' would contain different sources of irrationality. As mentioned, this axiom is, on the contrary, assuming that the mistakes are done on the level of cycles and hence the two cycles with the same elements would have the same sources of irrationality.

These previous 4 new axioms with Strong Homogeneity characterize Normalized Money Pump Index.

Theorem 3.2 Let I be a measure of rationality on $\mathcal{D}$. I satisfies Axiom 3.1 and Axioms 3.1^{NMPI}-3.4^{NMPI} iff

$$I = I_{\text{NMPI}}.$$

This result gives an insight into Normalized Money Pump Index. First, it satisfies a price homogeneity. Second, it is an averaging measure that is measured in constant cycle units. Third, only cycles matter for the measure but not the intersections of the cycles by strong monotonicity. Additionally, the lengths of cycles do not matter if one data has uniformly bigger cycles. Fourth, it satisfies linearity stating that mixing two equally rational data sets with the same prices and cycles does not make the data more or less rational.

The usefulness of this characterization comes from comparing this result to Afriat's characterization, Theorem 3.1. This gives us the following corollary.

Corollary 3.3 (Comparison: Afriat's Efficiency Index and Normalized Money Pump Index) The only difference between Afriat's Efficiency Index and Normalized Money Pump Index is that Afriat's Efficiency Index does not satisfy Linearity, Axiom 3.4^{NMPI} and Normalized Money Pump Index does not satisfy Continuity, Axiom 3.2.

This result shows a novel insight into the closeness of Afriat's Efficiency Index and Normalized Money Pump Index: The only difference between them is that Normalized Money Pump Index trades-off Continuity for Linearity. First, this highlights the similarity between them. Both are monotonic averaging measures. Second, it highlights that both measures are only different methods of taking an average: Afriat's Efficiency Index takes an average continuously whereas Normalized Money Pump Index takes the average linearly. This axiomatization and comparison provide an insight into the intuition of the Afriat's Efficiency Index and shows the strength of the Continuity axiom.

3.2.3 Normalized Minimum Cost Index

Next, we move on to non-averaging measures. We start off with introducing and axiomatizing a normalized version of Minimum Cost Index introduced by Dean and Martin (2016).

We define Normalized Minimum Cost Index (NMCI) similarly to the Afriat's Efficiency Index. However, this time we relax preferences by adjustments for each comparison. Let $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ be data and $(e_{t,t'}) \in [0,1]^{T \times T}$ be adjustments for each comparison. Relaxed revealed preference $\Big(\mathrm{R}^{(e_{t,t'})}, \mathrm{P}^{(e_{t,t'})}\Big)$ is defined by

$$x_t \; \mathrm{R}^{(e_{t,t'})} \; x_{t'} \iff (1 - e_{t,t'}) p_t \cdot x_t \geq p_t \cdot x_{t'}$$

and

$$x_t \; \mathrm{P}^{(e_{t,t'})} \; x_{t'} \iff (1 - e_{t,t'}) p_t \cdot x_t > p_t \cdot x_{t'}. \tag{3.2}$$

Normalized Minimum Cost Index is the smallest amount of adjustment required to rationalize the data.

$$I_{\mathrm{NMCI}}(D) = \frac{1}{T} \inf_{(e_{t,t'}) \in [0,1]^{T \times T}} \sum_{t=1}^{T} \sum_{t'=1}^{T} e_{t,t'} \text{ such that } \Big(\mathrm{R}^{(e_{t,t'})}, \mathrm{P}^{(e_{t,t'})}\Big) \text{ is acyclical.}$$

Here, $e_{t,t'}$ is the fraction of income that could be extracted from the agent from trading $x_{t'}$ for x_t at the price p_t. Now, NMCI is the most conservative estimate for the per observation average of the fraction of income that could be extracted from all the possible trades. This allows the same object to be traded possibly multiple times.

We can simplify this definition since only the acyclicality of the strict part of the revealed preference affects the NMCI. Removing indifferent parts of the revealed preference is at zero cost. We can simplify the measure to

$$I_{\text{NMCI}}(D) = \frac{1}{T} \inf_{(e_{t,t'}) \in [0,1]^{T \times T}} \sum_{t=1}^{T} \sum_{t'=1}^{T} e_{t,t'} \text{ such that } \mathrm{P}^{(e_{t,t'})} \text{ is acyclical,} \tag{3.3}$$

where, formally, the strict preference is acyclical if the pair $(\mathrm{P}^{(e_{t,t'})}, \mathrm{P}^{(e_{t,t'})})$ is acyclical.

Next, we will define a random version of Normalized Minimum Cost Index (RNMCI). Our axiomatization gives an insight into the NMCI by showing when a measure of rationality is between NMCI and RNMCI. As we will argue later on, this bound is tight and allows us to avoid technical axioms in order to gain more useful insight into NMCI. First, we define random strict preferences and when they are acyclical.

A random preference $\succ$ on a set X gives a probability for every pairwise comparison, that is for all $x, y \in X$, $x \succ y \in [0, 1]$. This includes standard preferences when the probabilities are always 0 or 1. Define the complement preference $\not\succ$ as $x \not\succ y = 1 - x \succ y$.

Next, we will define when a random preference is acyclical by relating it to deterministic acyclicality. First, a deterministic preference $\succ$ is acyclical if there does not exist a sequence $(x_1, \ldots, x_n) \subseteq X$ such that for all $1 \leq i \leq n$, $x_i \succ x_{i+1}$ when $x_{n+1} = x_1$. When we interpret $\succ$ as a deterministic random preference, this can be written as for all $1 \leq i \leq n$, $x_i \not\succ x_{i+1} = 0$. Finally, this is equivalent to

$$\sum_{i=1}^{n} x_i \not\succ x_{i+1} < 1$$

since all the deterministic preferences are either 0 or 1. Our random acyclicality condition applies this condition to random preferences.

Definition A random preference $\succ$ on X is acyclical if there does not exist a sequence $(x_1, \ldots, x_n) \subseteq X$ such that

$$\sum_{i=1}^{n} x_i \not\succ x_{i+1} < 1.$$

This condition states that $\succ$ is acyclical if one cannot guarantee that there is a cycle with a positive probability regardless of the correlation of pairwise comparisons. That is if in the

best-case scenario when each one of the events $x_i \not\succ x_{i+1}$ is disjoint from others, they do not fill the state space. Then especially, there is some positive probability that each one of the events $x_i \succ x_{i+1}$ is simultaneously true.

Now we can define Random Minimum Cost Index. Here instead of deterministic adjustments, we allow for a distribution of adjustments. RNMCI is the smallest expected adjustments that make the relaxed revealed preference acyclical in expectations. Let $(E_{t,t'}) \in \Delta([0,1]^{T\times T})$. For each realization $(e_{t,t'})$ of $(E_{t,t'})$, we define $\mathrm{P}^{(e_{t,t'})}$ as in (3.2). RNMCI is

$$I_{\mathrm{RNMCI}}(D) = \frac{1}{T} \inf_{(E_{t,t'})\in\Delta([0,1]^{T\times T})} \mathbb{E}\sum_{t=1}^{T}\sum_{t'=1}^{T} E_{t,t'} \text{ such that } \mathbb{E}\mathrm{P}^{(E_{t,t'})} \text{ is acyclical}$$

where the expectation of $\mathrm{P}^{(E_{t,t'})}$ is taken coordinatewise.

First, if the optimal $(E_{t,t'})$ is deterministic, then $I_{\mathrm{NMCI}}(D) = I_{\mathrm{RNMCI}}(D)$. So RNMCI relaxes NMCI by allowing for probabilistic removals of strict revealed preferences. Especially, we always have $I_{\mathrm{NMCI}}(D) \geq I_{\mathrm{RNMCI}}(D)$. Additionally, for the data sets used by Dean and Martin (2016), Echenique et al. (2011), and Choi et al. (2014) that includes over 3000 observations, we always have $I_{\mathrm{NMCI}}(D) = I_{\mathrm{RNMCI}}(D)$. This suggests that the bounds between NMCI and RNMCI are tight.

Remark The minimization problem in the definition of NMCI can be simplified. First, as noted in (3.3), we can focus only on the strict part of revealed preference. Secondly, the optimal $e_{t,t'}$s are either 0 or $\frac{p_t\cdot(x_t - x_{t'})}{p_t\cdot x_t}$ since each $e_{t,t'}$ affects only a single pair of revealed preference. This allows us to simplify the measure to

$$I_{\mathrm{NMCI}}(D) = \frac{1}{|D|}\min_{B\subseteq \mathrm{P}} \sum_{(x_t,x_{t'})\in B} \frac{p_t\cdot(x_t - x_{t'})}{p_t\cdot x_t} \text{ such that } \mathrm{P}\setminus B \text{ is acyclical.}$$

That is NMCI is removing the revealed preference for acyclicality where the cost of removing a relation is the fraction of income that could be extracted at that trade.

A similar alternative definition holds for RNMCI. First, denote by P^{Δ}, the strict revealed preference when interpreted as a random preference. Second, let $\beta : \mathrm{P} \to [0,1]$ be the probability of removing each strict revealed preference. Then define $\mathrm{P}^{\Delta} - \beta$ by doing the

subtraction elementwise: for all $x, y \in X$, $x(\mathrm{P}^{\Delta} - \beta)y = x\mathrm{P}^{\Delta}y - \beta(x, y)$. This corresponds to standard set subtraction if β is 0 or 1. Now RNMCI is

$$I_{\mathrm{RNMCI}}(D) = \frac{1}{|D|} \min_{\beta:\mathrm{P}\to[0,1]} \sum_{(x_t,x_{t'})\in \mathrm{P}} \beta(x_t, x_{t'}) \frac{p_t \cdot (x_t - x_{t'})}{p_t \cdot x_t} \text{ such that } \mathrm{P}^{\Delta} - \beta \text{ is acyclical.}$$

Remark If in the definition of RNMCI, instead of the condition $\mathbb{E}\mathrm{P}^{(E_{t,t'})}$ is acyclical, we would assume that $\mathbb{E}\mathrm{P}^{(E_{t,t'})}$ is in the convex hull of deterministic acyclical preferences, this measure would be NMCI: Using the above definition this is equivalent to assuming that $\mathrm{P}^{\Delta} - \beta$ is in the convex hull of deterministic acyclical preferences. Now we have a linear minimization problem over a convex set, so the optimal point is an extreme point of the convex set i.e. a deterministic acyclical preference. This highlights the closeness of RNMCI and NMCI.

Axioms for Normalized Minimum Cost Index

Next, we characterize when a measure of rationality is between NMCI and RNMCI. We will use Strong Homogeneity, Continuity, and Monotonicity, Axioms 3.1-3.3, from Afriat's Efficiency Index. NMCI is normalized differently from Afriat's Index and so we will use a slightly different normalization axiom. Finally, instead of Afriat's Index's Averaging axiom, we have two new axioms: First, we assume that there is no double counting of trades. The second new axiom is the main axiom of NMCI and assuming that the measure of rationality is sensitive to every violation of rationality.

We start with a normalization. The previous measures worked with unit income normalized data. NMCI instead is measured with total income normalized data.

Definition $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ is *total income normalized* if for all $1 \leq t, t' \leq T$,

$$p_t \cdot x_t = p_{t'} \cdot x_{t'} \text{ and } \sum_{t=1}^{T} p_t \cdot x_t = 1.$$

Now the normalization axiom is applied to total income normalized data sets.

Axiom 3.1$^{\mathbf{NMCI}}$ (Normalization) If $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ is total income normalized data such that

$$\{(x_{t_1}, \ldots, x_{t_n})\} = \mathcal{C}^D$$

and for some $\beta > 0$, for all $1 \leq i \leq n$,

$$p_{t_i} \cdot (x_{t_i} - x_{t_{i+1}}) = \beta,$$

then

$$I(D) = \beta.$$

This axiom assumes that the unit of measurement for the measure of rationality is the average fraction extracted from constant cycles for total income normalized data. With Strong Homogeneity, these total income normalized data sets give that the measure is looking at the per observation average violation of rationality and gives the normalization $\frac{1}{|D|}$ for the NMCI.

The next axiom is assuming that there is no double counting of trades.

Axiom 3.2$^{\mathbf{NMCI}}$ (Limited Effect of a Trade) If $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ and $D' = \Big((p'_1, x'_1), \ldots, (p'_T, x'_T)\Big)$ are total income normalized data such that for some $\beta > 0$ and t_0, t'_0,

$$p'_{t_0} \cdot (x'_{t_0} - x'_{t'_0}) = p_{t_0} \cdot (x_{t_0} - x_{t'_0}) + \beta$$

and for all other $1 \leq t, t' \leq T$,

$$p'_t \cdot (x'_t - x'_{t'}) = p_t \cdot (x_t - x_{t'}),$$

then

$$I(D') \leq I(D) + \beta.$$

This axiom states that if we can extract β more from a single trade, then the measures of rationality can increase at most by β. That is there is no double counting of a single trade.

On the other hand, we can use the previous axiom to provide an insight into this axiom. The previous axiom defines the unit of measurement as constant cycles and gives meaning to the cardinal numbers. So this axiom states that a single trade is not more violating than a constant cycle. The interpretation of this axiom is that a single trade can affect at most a single mistake that is counted by the measure of rationality.

The last axiom rules out negative interactions between cycles. This is the main axiom for NMCI type measures.

Axiom 3.3^{NMCI} (All Cycles Matter) If $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$, $D' = \Big((p'_1, x'_1), \ldots, (p'_T, x'_T)\Big)$ are total income normalized data such that for some $(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D$ and $\beta > 0$, for all $1 \leq i \leq n$,

$$p'_{t_i} \cdot (x'_{t_i} - x'_{t_{i+1}}) = p_{t_i} \cdot (x_{t_i} - x_{t_{i+1}}) + \beta$$

and for all other $1 \leq t, t' \leq T$,

$$p'_t \cdot (x'_t - x'_{t'}) = p_t \cdot (x_t - x_{t'}),$$

then

$$I(D') \geq I(D) + \beta.$$

This axiom assumes that if the cost of a single cycle c is increased by β at every trade, then the measure must increase at least by β. The reason for the increase is that we have increased the cost of at least a single mistake that created the cycle and is measured by the measure of rationality. That is the sources of irrationality are included in the trades of the cycles.

The idea for the axiom is that if c were the only cycle, then increasing the cost of its every trade would increase the measure by β by the previous two axioms. Now, this axiom is assuming that in the presence of other cycles, this single violation is not less significant but it must be at least as significant as without any other cycles. So this axiom rules out negative interactions between cycles and assumes that every violation of rationality or cycle affects the measure of rationality.

Remark Normalization, Axiom 3.1^{NMCI}, is implied by the last two axioms. However, it useful at interpreting them by giving the unit of measurement.

These three new assumptions with Strong Homogeneity, Continuity, and Monotonicity guarantee that the measure of rationality is between NMCI and RNMCI. Additionally, these assumptions are necessary for NMCI and RNMCI.

Theorem 3.4 If a measure of rationality I satisfies Axioms 3.1-3.3 and Axioms 3.1^{NMCI}-3.3^{NMCI}, then for all $D \in \mathcal{D}$,

$$I_{\text{NMCI}}(D) \geq I(D) \geq I_{\text{RNMCI}}(D).$$

Additionally, all the above axioms are necessary for I_{NMCI} and I_{RNMCI}.

Since the bounds between NMCI and RNMCI are tight, this result gives an insight into NMCI. This result shows that under the price invariance of Strong Homogeneity, NMCI type measures are the only measures that are continuous and monotonic in rationality violations, measured in the constant cycle units, and sensitive to every violation of rationality so that there are no negative interactions between violations of rationality.

Second, using this result and Theorem 3.1, we can axiomatically compare NMCI to Afriat's Efficiency Index. This comparison shows the close connection between the two measures. The only difference between NMCI and Afriat's Index is the way they aggregate violations of rationality: Afriat's Index takes the average violation, whereas NMCI is sensitive to every violation of rationality. That is NMCI is a more refined version of Afriat's Index.

Third, using Theorem 3.2 we can compare NMCI to Normalized Money Pump Index. We find out that NMCI does not satisfy any of Axioms 3.2^{NMPI}-3.4^{NMPI}. But on the other hand, ignoring the different normalization, NMPI satisfies all of NMCI's axioms except Continuity, Axiom 3.2, and Axiom 3.3^{NMCI}. That is NMPI trades-off continuity and sensitivity to every cycle for taking an average, strong monotonicity ignoring intersections of cycles, and linearity.

3.2.4 Varian's Goodness of Fit

Next, we introduce Varian's Goodness of Fit (Varian, 1990) that is a measure between Afriat's Efficiency Index and Normalized Minimum Cost Index. This measure can be axiomatized using the same method as Normalized Minimum Cost Index using bounds with a random version of the measure. However, these bounds are not as tight for Varian's measure. We will omit the axiomatization of this measure and only introduce the measure.

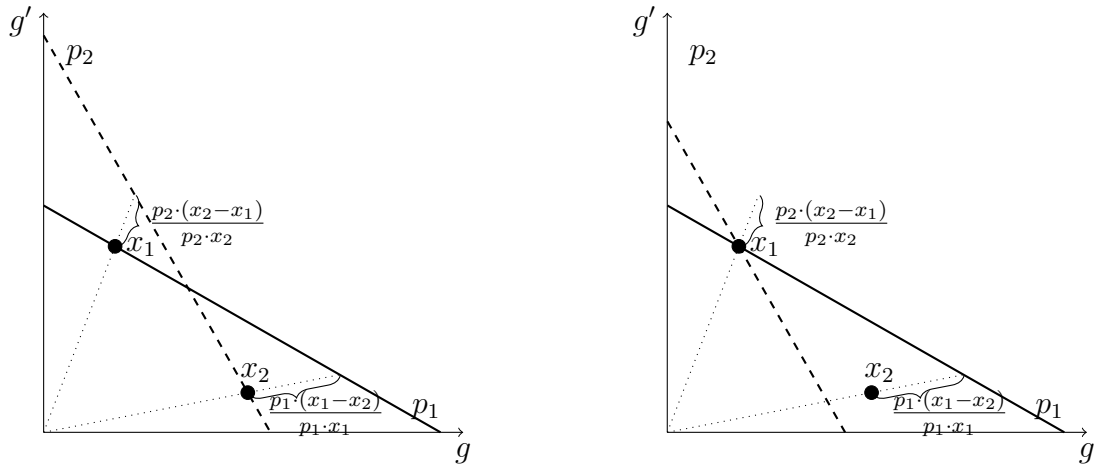

Figure 3.6. An example of calculating Varian's Index. The data is $D = ((p_1, x_1), (p_2, x_2))$ with a cycle between x_1 and x_2. Individuals budgets are adjusted downwards until there are no cycles left. The smallest size of budget adjustments to break the cycles is the value of Varian's Index. In this example we have $\frac{p_2\cdot(x_2-x_1)}{p_2\cdot x_2} < \frac{p_1\cdot(x_1-x_2)}{p_1\cdot x_1}$ and so Varian's Index adjusts budget line p_2 and the index is $\frac{p_2\cdot(x_2-x_1)}{p_2\cdot x_2}$.

We define Varian's Index similarly to NMCI and Afriat's Efficiency Index. However, this time we relax preferences by adjustments for each observation. Let $D=\Big((p_1, x_1), ..., (p_T, x_T)\Big)$ be data and $(e_t) \in [0, 1]^T$ be adjustments for each observation. Relaxed revealed preference $\Big(\mathrm{R}^{(e_t)}, \mathrm{P}^{(e_t)}\Big)$ is defined by

$$x_t \, \mathrm{R}^{(e_t)} \, x_{t'} \iff (1 - e_t)p_t \cdot x_t \geq p_t \cdot x_{t'}$$

and

$$x_t \, \mathrm{P}^{(e_t)} \, x_{t'} \iff (1 - e_t)p_t \cdot x_t > p_t \cdot x_{t'}.$$

Varian's Index is the least amount of adjustment required to rationalize the data.

$$I_{\mathrm{Var}}(D) = \frac{1}{T} \inf_{(e_t)\in[0,1]^T} \sum_{t=1}^{T} e_t \text{ such that } \Big(\mathrm{R}^{(e_t)}, \mathrm{P}^{(e_t)}\Big) \text{ is acyclical.}$$

The idea is that the Varian's Index is the most conservative estimate for the per observation average of income that could be extracted from all the possible trades when each bundle can be traded at most once. This is the difference to NMCI that allows the same bundles to be traded multiple times.

The idea of Varian's Index is similar to Afriat's Index and illustrated in Figure 3.6: we are pulling the budget lines towards the origin to break the revealed preference cycles. However,

this time we are pulling a single budget line at a time. Varian's Index is the least amount of adjustment of budget lines required to break all the cycles.

3.2.5 Inverse Varian's Goodness of Fit

Our last measure is Inverse Varian's Goodness of Fit that is a novel alternative method of taking a measure between Afriat's Efficiency Index and Normalized Minimum Cost Index. This measure can be axiomatized using the same method as Normalized Minimum Cost Index using bounds with a random version of the measure, but we will omit the axiomatization.

We define Inverse Varian's Index similarly to Varian's Index. Again, we relax preferences by adjustments for each observation but this time the adjustments depend on the compared bundle or the bundle on the right-hand side. Let $D = \Big((p_1, x_1), \ldots, (p_T, x_T)\Big)$ be data and $(e_{t'}) \in [0,1]^T$ be adjustments for each observation. Relaxed revealed preference $\Big(\mathrm{R}^{(e_{t'})}, \mathrm{P}^{(e_{t'})}\Big)$ is defined by

$$x_t \, \mathrm{R}^{(e_{t'})} \, x_{t'} \iff (1 - e_{t'}) p_t \cdot x_t \geq p_t \cdot x_{t'}$$

and

$$x_t \, \mathrm{P}^{(e_{t'})} \, x_{t'} \iff (1 - e_{t'}) p_t \cdot x_t > p_t \cdot x_{t'}.$$

Inverse Varian's Index is the least amount of adjustment required to rationalize the data.

$$I_{\mathrm{InvVar}}(D) = \frac{1}{T} \inf_{(e_{t'}) \in [0,1]^T} \sum_{t=1}^{T} e_{t'} \text{ such that } \Big(\mathrm{R}^{(e_{t'})}, \mathrm{P}^{(e_{t'})}\Big) \text{ is acyclical.}$$

The idea of this measure is that we can write the relaxed revealed preference as

$$x_t \, \mathrm{R}^{(e_{t'})} \, x_{t'} \iff p_t \cdot x_t \geq p_t \cdot (1 - e_{t'})^{-1} x_{t'}$$

and

$$x_t \, \mathrm{P}^{(e_{t'})} \, x_{t'} \iff p_t \cdot x_t > p_t \cdot (1 - e_{t'})^{-1} x_{t'}.$$

In this definition instead of decreasing the income and pulling the budget lines inwards to break the cycles, we are increasing the costs of bundles and pushing the bundles outwards away from the origin to break the revealed preference cycles. This is illustrated in Figure 3.7.

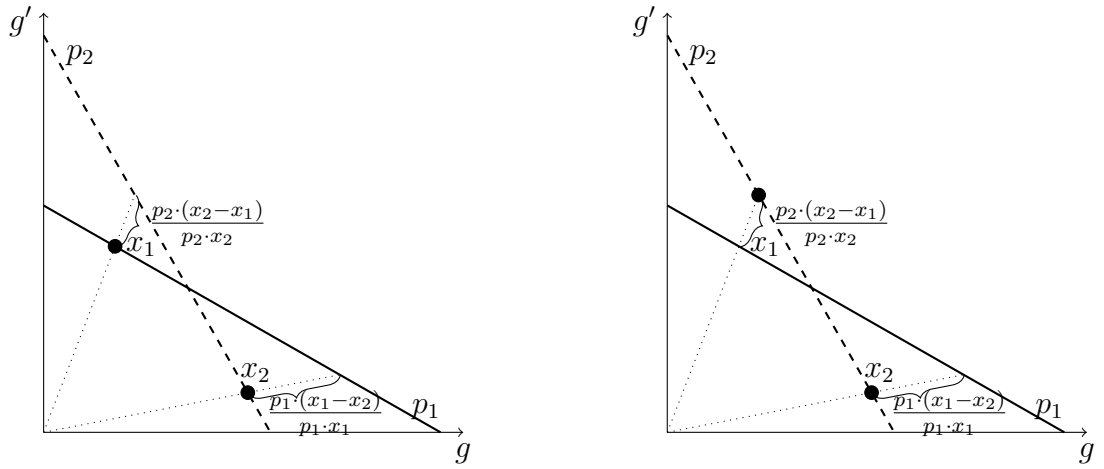

Figure 3.7. An example of calculating Inverse Varian's Index. The data is $D = ((p_1, x_1), (p_2, x_2))$ with a cycle between x_1 and x_2. The individual bundles are pushed outwards until there are no cycles left. The smallest size of bundle adjustments to break the cycles is the value of Inverse Varian's Index. In this example we have $\frac{p_2\cdot(x_2-x_1)}{p_2\cdot x_2} < \frac{p_1\cdot(x_1-x_2)}{p_1\cdot x_1}$ and so Inverse Varian's Index adjusts bundle x_1 and the index is $\frac{p_2\cdot(x_2-x_1)}{p_2\cdot x_2}$.

As with Varian's Index, the Inverse Varian's Index is the least amount of adjustment of the bundles required to break all the cycles. However, the cost of adjusting bundles is not additive this time but is measured in the reciprocal units $(1-e)^{-1}$. Instead of capturing how much the decision maker would have needed to misperceive the income to explain the irrational choices like Varian's Index, this index focuses on how much the decision maker would have needed to misperceive the costs of bundles to explain the irrational choices. This measure is focusing on the most conservative estimate on increasing the cost of bundles that would rationalize the data.

3.2.6 Axiomatic Summary

Table 3.2 summarizes our axiomatic comparisons of measures. The table does not take into account the axiomatic differences coming from different normalizations of NMCI, Var, and InvVar that are looking at per observation average measure. The necessary axioms characterize Afriat's Efficiency Index, Normalized Money Pump Index, and essentially Normalized Minimum Cost Index.

Axiom	3.1	3.2	3.3	3.4	3.5
Measure	Str.Homogeneity	Continuity	Monotonicity	Normalization	Averaging
Afr	✓	✓	✓	✓	✓
NMPI	✓	-	✓	✓	✓
NMCI	✓	✓	✓	✓	-
Var	✓	✓	✓	✓	-
InvVar	✓	✓	✓	✓	-

Axiom	3.2^{NMPI}	3.3^{NMPI}	3.4^{NMPI}	3.2^{NMCI}	3.3^{NMCI}
Measure	Norm. & Averaging	Str.Monot.	Linearity	Lim. Effect Trade	All Cycles Matter
Afr	✓	✓	-	✓	-
NMPI	✓	✓	✓	✓	-
NMCI	-	-	-	✓	✓
Var	-	-	-	✓	-
InvVar	-	-	-	✓	-

✓: Necessary, - : Does not satisfy.

Table 3.2. Summary of axioms satisfied by different measures of rationality. The table does not take into an account the axiomatic differences coming from different normalizations of the measures.

Remark Using these axiomatizations, we can axiomatize for example an additive Normalized Money Pump Index that sums all the mistakes or use different price invariance assumptions as in Mononen (2020).

3.3 Simulation Comparisons

We move on to practical comparisons of the measures of rationality using simulations. In this section, we first introduce a model of choice mistakes. Second, we calibrate the size of mistakes to match an empirically reasonable range. Third, we use the model for ordinal comparisons of measures based on pairwise comparisons of different size mistakes and calculating hit rates of ranking the mistakes correctly. Fourth, we study extensions of standard measures by under and overweighting small violations of rationality. Finally, we discuss the computational aspects of these simulations.

We focus on comparing continuous measures of rationality: Afriat's Efficiency Index, Minimum Cost Index, Varian's Index, and Inverse Varian's Index. Normalized Money Pump Index has additional computational constraints that restrict their effective simulations.

Our first main finding of this simulation study is that the best measure depends on how fat-tailed the mistake distribution is. If the distribution is fat-tailed, then a more aggregated measure such as Varian's Index is the appropriate choice. Whereas if the mistakes are concentrated, then a more refined measure such as Normalized Minimum Cost Index is the appropriate choice.

The second main finding of the simulation is that combining the measures with ordinal counting measures and overweighting small violations of rationality always improves the performance of the measures. Especially, with fat-tailed mistake distributions, the best variations of the measures use lexicographically ordinal (the number of rationality violations) and cardinal (size of the violations) measures prioritizing the ordinal measures. This gives a unifying perspective to the literatures on ordinal and cardinal measures of rationality. Additionally, it suggests that multiple small mistakes are more indicative of lower rationality than a single large mistake.

In this section, we focus on mistakes in choosing the optimal quantity with random mistakes in choosing the income shares. The results are robust to different types of mistakes: Randomizing uniformly on the budget plane. Random utility with an additive utility error for logit-choices on the budget plane. Making the utility function observation dependent with variance in the parameters of the utility function.

3.3.1 Choice Model

We model a decision maker (DM) who does purchases given the prices but does mistakes at purchasing the optimal quantities. Specifically, the DM does choices with a trembling hand around the optimal bundle.

We assume that the DM has a constant income I of 8.[7] We assume that the prices are independent draws from truncated Normal distribution with mean 4 and variance 1 that is truncated to be above 0.1. These results are robust to using a uniform distribution between 0.1 and 8. The prices are truncated away from zero to have more overlap in the budget sets to increase violations of rationality. The number of goods G is varied in different specifications.

The decision maker is assumed to have a symmetric additive log utility over the goods: The utility from a bundle $(x_g)_{g=1}^G \in \mathbb{R}_+^G$ is

$$u\Big((x_g)_{g=1}^G\Big) = \sum_{g=1}^{G} \log(x_g).$$

The decision maker has an optimal demand $(x_g^*)_{g=1}^G$ given the income 8 and prices $(p_g)_{g=1}^G$. The decision maker makes choice mistakes $(\varepsilon_g)_{g=1}^G$ in the budget shares of goods and scales uniformly the budget shares to consume the whole income. That is for each good g, the observed income shares with mistakes are

$$\frac{\frac{p_g}{I}x_g^* + \varepsilon_g}{\sum_{g'=1}^G \frac{p_{g'}}{I}x_{g'}^* + \varepsilon_{g'}} = \frac{p_g x_g^* + I\varepsilon_g}{I(1+\sum_{g'=1}^G \varepsilon_{g'})}$$

when I is the constant income of 8. Then the observed choices with mistakes are

$$\frac{x_g^* + \frac{I}{p_g}\varepsilon_g}{1+\sum_{g'=1}^G \varepsilon_{g'}}.$$

We test for different independent distributions of mistakes. The distributions are chosen to have a mode at 0 and truncated to be above $-\frac{p_g}{I}x_g^*$ for the income shares to be positive. Here we focus on two very different mistake distributions: Normal and Cauchy distributions. Normal distribution is chosen as a concentrated standard distribution. Cauchy distribution (Student's t-distribution of degree 1) is chosen as a standard example of fat-tailed distribution. These two distributions illustrate two extremities of concentration of distributions. The results are robust to using alternatively log-Normal, log-Cauchy, uniform, Weibull, Fréchet, Gumbel, or Gamma distributions.

[7]With a constant income normalized and non-normalized measures of rationality are linear scalings of each other and give the same values in our ordinal comparisons.

3.3.2 Ordinal Measure Comparisons

The previous mistake model gives an objective size of mistakes from the variance or scaling of the distributions. Next, we study how well the measures of rationality capture these objectively sized mistakes. The first difficulty in this is that these objective mistakes might be measured in different units from the measures of rationality and the units of different measures of rationality might not be comparable. To avoid these unit problems, we focus on ordinal comparisons and focus on finding out if different measures rank the irrationality of different data sets correctly according to the objective criterion. We explain this in more detail next.

We focus on ordinal non-trivial pairwise comparisons of data sets. We have two populations with different mistake levels e_1 and e_2 such that $e_1 > e_2$. Then we draw data sets $\tilde{D}_1$ and $\tilde{D}_2$ with different prices from these two populations using the previous section's choice model. Now $\tilde{D}_1$ is objectively more irrational than $\tilde{D}_2$ based on the mistake levels in the population or the data generating process. First, we redraw these data sets until they have violations of rationality. This gives us data sets D_1 and D_2 where D_1 is objectively more irrational than D_2. We focus only on comparing data sets with violations of rationality since otherwise by the definition of measure of rationality all measures of rationality would agree on the ranking of the comparisons and the comparison would be trivial.

We calculate measures of rationality for these two data sets and check if the measures of rationality agree that D_1 is more irrational and has a higher measure than D_2. We focus on the average number of correct pairwise comparison that gives us hit rates for each measure of rationality. For measure I, these hit rates estimate the probability

$$P\Big(I(D_1) > I(D_2)\Big|D_1, D_2 \text{ not rational}\Big)$$

where the probability is taken over the price distribution and the mistake distributions for D_1 and D_2. Comparing these probability estimates gives us ordinal comparisons of measures.

3.3.3 Calibration

Previously, we have studied how to do ordinal comparisons of measures of rationality given different sized mistakes. Next, we will calibrate the mistakes levels for appropriate comparisons. The idea for the calibration is to calibrate the mistake levels for the populations using the share of rational people in the population and focus on reasonable shares of rational people.

Formally, for each mistake level e, we estimate the share of rational people in the population. This is done by drawing data D from the population and by estimating the probability

$$P(D \text{ is rationalizable})$$

by sample average. We calibrate the mistake level e based on this estimate. Based on empirical field studies Echenique et al. (2011) and Dean and Martin (2016) around 20% to 30% of people are perfectly rational.

In the calibration, we keep the average level of rationality constant at 50%. Here, we start from two populations that consist of 50% rational individuals. Then we consider making the mistake level comparisons across these two populations easier by taking mean preserving spreads of rationality for these two populations. That is we increase the share of rational people for one population and decrease the share of rational people for the other population while keeping the average share of rational people at 50% across these two populations. We vary the level of mean preserving spreads from 5% to 45%. We have chosen the average share of rationality to be 50% in order to have as many mean preserving spreads for comparison as possible. This is slightly above the empirical estimates.

We consider an alternative calibration method in the appendix that keeps the difficulty of rationality comparisons as constant and changes the levels of rationality compared. We implement these calibrations using a simulated method of moments.

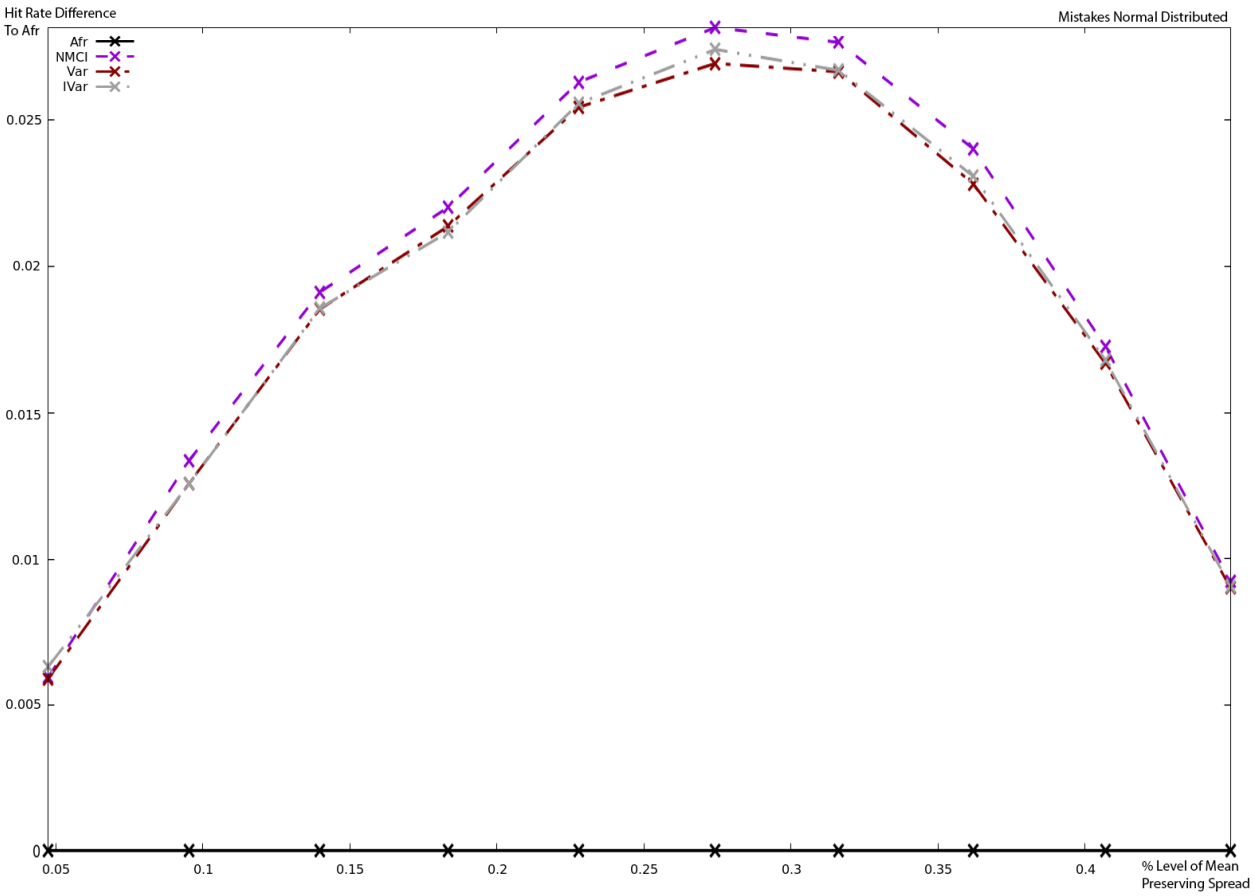

Figure 3.8. Measures of rationality's hit rate differences to Afriat's Index's hit rate when mistakes are drawn from Normal distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread. For example, at 5% the low mistake population's share of rational people is 55% and the high mistake population's share of rational people is 45%.

3.3.4 Comparisons of Measures

Next, we move on to comparing the measures of rationality in the simulations. We focus on 6 goods with 45 time periods/observations. The low number of goods follows the empirical studies on the measures of rationality that aggregate goods to a few categories. The results are robust to changing the number of goods or time periods. The simulation sample size for each point is 100,000 pairwise comparisons when both of the compared data sets have rationality violations.

Figures 3.8 and 3.9 compare the hit rates of different measures to Afriat's Index's hit rate when mistakes are Normal and Cauchy distributed respectively. The two populations

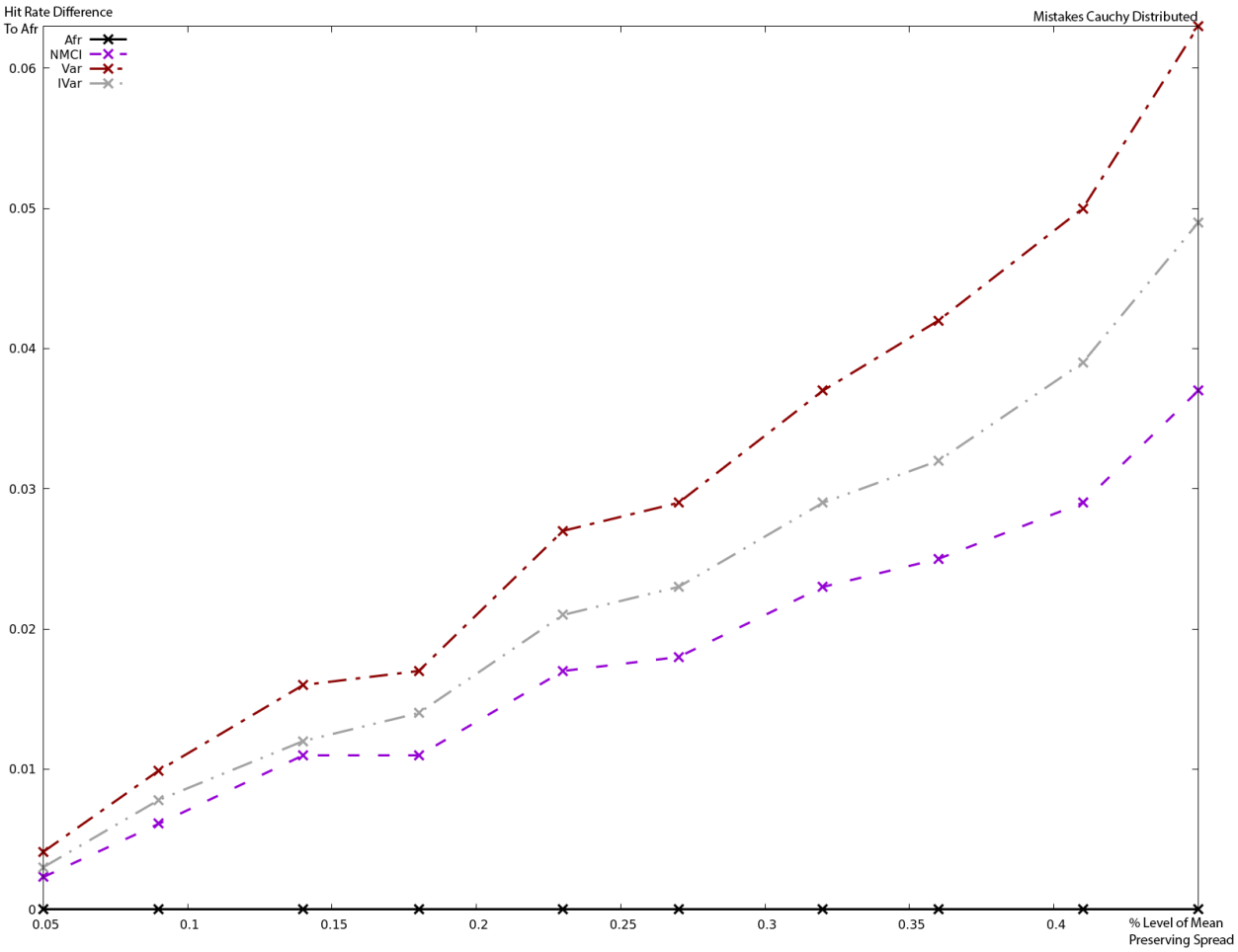

Figure 3.9. Measures of rationality's hit rate differences to Afriat's Index's hit rate when mistakes are drawn from Cauchy distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread.

compared are calibrated to have a level of rationality that is a mean preserving spread from 50%. The x-axis changes the level of mean preserving spread from 5% to 45%. That is x-axis changes from comparing 55% rational population with 45% rational population to comparing 95% rational population with 5% rational population. So the x-axis changes the difficulty of the rationality comparison while keeping the average level of rationality at a constant level.

We start with normally distributed mistakes in Figure 3.8. This graph shows that all measures are doing better than Afriat's Index. First, the hit rate differences to Afriat's Index's hit rate are first growing and then decreasing. Appendix Figure C.1 shows this graph when the y-axis is the absolute level of hit rates. It shows that the decrease in hit rate

differences comes from all measures getting almost all comparisons correctly and so there is not much room for improvement anymore.

Second, NMCI slightly outperforms the other measures. The performance improvement compared to Varian's or Inverse Varian's Index is up to 0.2 percentage points. Varian's and Inverse Varian's Indexes are doing approximately equally well.

This graph illustrates a situation when the mistake distribution is concentrated. Other mistake distributions and mistake models with concentrated mistake distributions create the same results: Logit-choice mistakes or Trembling hand mistakes or observation dependent utility function parameters with concentrated distributions of Log-Normal, Uniform, Gumbel, or Reversed Weibull.

The dominance of NMCI is robust to changing the number of time periods and to changing the number of goods. However, with a low number of goods (3) and a low number of time periods (25), NMCI, Var, and InvVar do approximately equally well. Increasing the number of goods or the number of periods slightly increases the dominance of NMCI.

This result suggests that more refined measures such as NMCI are doing better when the mistakes are concentrated.

We move on to Cauchy distributed mistakes in Figure 3.9. First, this graph shows that again all measures are doing better than Afriat's Index. However, this time the differences to Afriat's Index are always increasing and the magnitude of differences is much greater than previously. Appendix Figure C.2 shows this graph when the y-axis is the absolute level of hit rates. It shows that even when the comparison of measures should be easy in terms of the average rationality of the populations compared, the measures are only getting around 75% of the comparisons correctly. This shows that the measures do not get saturated at the easy comparisons, as they did previously with normally distributed mistakes. Comparison of Appendix Figures C.1 and C.2 shows that the measures are worse at capturing fat-tailed mistakes.

Second, Figure 3.9 shows that Varian's Index does better than the other measures. The difference between Varian's and NMCI's hit rates is up to 2 percentage points. Inverse Varian's Index is between Varian's Index and NMCI. It is doing up to 1 percentage point worse than Varian's Index and up to 1 percentage point better than Varian's Index.

This graph illustrates a situation when the mistake distribution is fat-tailed. Other mistake distributions and mistake models with fat-tailed mistake distributions create the same results: Uniform random choice on the budget plane or Trembling hand mistakes or observation dependent utility function parameters with fat-tailed distributions of Log-Cauchy or Fréchet.

This result is robust to changing the number of time periods or the number of goods.

This result suggests that more aggregated measures such as Varian's Index or Inverse Varian's Index are doing better when the mistakes are fat-tailed and dispersed. Additionally, the magnitude of improvement is significant, but all the measures have difficulties at capturing these dispersed mistakes.

These two results show that the choice of appropriate measure depends on the application at hand.

3.3.5 Variations of Measures

In this section, we study different convexifications and concavifications of the measures. First, the convexifications follow Varian's (1990) suggestion of using power transformations to weigh larger violations of rationality more. In this case, a single large violation of rationality is more irrational than many small mistakes of the same total size. Second, concavifications use root transformations to weigh small violations of rationality more. In this case, many small mistakes are more irrational than a single large mistake of the same total size. We start off by defining how to do power and root transformations of the measures.

We begin with Normalized Minimum Cost Index. Using the definition of NMCI from Section 3.2.3, we define power transformation of NMCI to the power α as

$$I_{\text{NMCI}\alpha}(D) = \frac{1}{T} \inf_{(e_{t,t'}) \in [0,1]^{T \times T}} \sum_{t=1}^{T} \sum_{t'=1}^{T} e_{t,t'}^{\alpha} \text{ such that } \left(\mathrm{R}^{(e_{t,t'})}, \mathrm{P}^{(e_{t,t'})}\right) \text{ is acyclical,}$$

where the pair is the relaxed revealed preference from Section 3.2.3.

If $\alpha = 1$, this is NMCI. If $\alpha = 2$, then mistakes are measured in the Euclidean norm with the sum of squared mistakes. The idea in here is that bigger mistakes are worse for rationality and should be weighted more relative to small mistakes.

When $0 < \alpha < 1$, this gives concavifications of NMCI. In this case, we say that it is a root transformation of NMCI to the degree $\frac{1}{\alpha}$. The idea for the concavifications is that we give more weight to small mistakes relative to larger mistakes. Especially we give more weight to ordinal information of the violation, that is the existence of the violation, relative to the intensity or the size of the violation.

Next, we extend on the connection to the ordinal information and root transformation. When α goes to 0, e^{α} converges to 1 pointwise for all $e > 0$. So when α goes to 0, $I_{\text{NMCI}\alpha}$ converges to an ordinal counting measure that counts the number of revealed preferences needed to remove to make the revealed preference acyclical. Now since we are focusing on ordinal pairwise comparisons, the limiting measure of $I_{\text{NMCI}\alpha}$ still uses the sizes of the violations in choosing which data set is more rational when the two data sets have the same number of violations. That is the measure uses the sizes of violations as a tie-breaking rule when the comparison cannot be done purely on the number of violations. So this limiting measure corresponds to a lexicographic ordering of rationality that first uses ordinal information of the revealed preference to compare the rationality and secondly uses cardinal information of the revealed preference when the comparison cannot be done based on the ordinal information.

Next, we define symmetrically power transformations for Varian's Index and Inverse Varian's Index. Following Varian's Index's definition from Section 3.2.4, we define the power transformation of Varian's Index to the power α as

$$I_{\text{Var}\alpha}(D) = \frac{1}{T} \inf_{(e_t)\in[0,1]^T} \sum_{t=1}^{T} e_t^{\alpha} \text{ such that } \left(\mathrm{R}^{(e_t)}, \mathrm{P}^{(e_t)}\right) \text{ is acyclical.}$$

And symmetrically for Inverse Varian's Index from Section 3.2.5

$$I_{\text{InvVar}\alpha}(D) = \frac{1}{T} \inf_{(e_{t'})\in[0,1]^T} \sum_{t=1}^{T} e_{t'}^{\alpha} \text{ such that } \left(\mathrm{R}^{(e_{t'})}, \mathrm{P}^{(e_{t'})}\right) \text{ is acyclical.}$$

The above discussion applies to these transformations almost symmetrically. The only difference is that the ordinal information of these measures and the limiting measure when α goes to 0, is a measure that counts the number of observations needed to remove to make the revealed preference acyclical. This is the Houthman-Maks Index (Houtman and Maks, 1985). Especially, the ordinal information part of Varian's and Inverse Varian's Indexes limiting measures is the same. However, these limiting measures differ on the cardinal information.

Finally, a power transformation of Afriat's Index would be an ordinal transformation of Afriat's Index and hence would not affect our ordinal comparisons. These transformations are omitted.

As before in our simulations, we focus on 6 goods with 45 time periods/observations. The results are robust to changing the number of goods or time periods. The simulation sample size for each point is 100,000 pairwise comparisons when both of the compared data sets have rationality violations.

Convexifications

Next, we test different power transformations of the above three measures up to 32 degrees. This is done in Figure 3.10 when mistakes are normally distributed and in Figure 3.11 when mistakes are Cauchy distributed. The number in the measures' labels is the degree of power transformation.

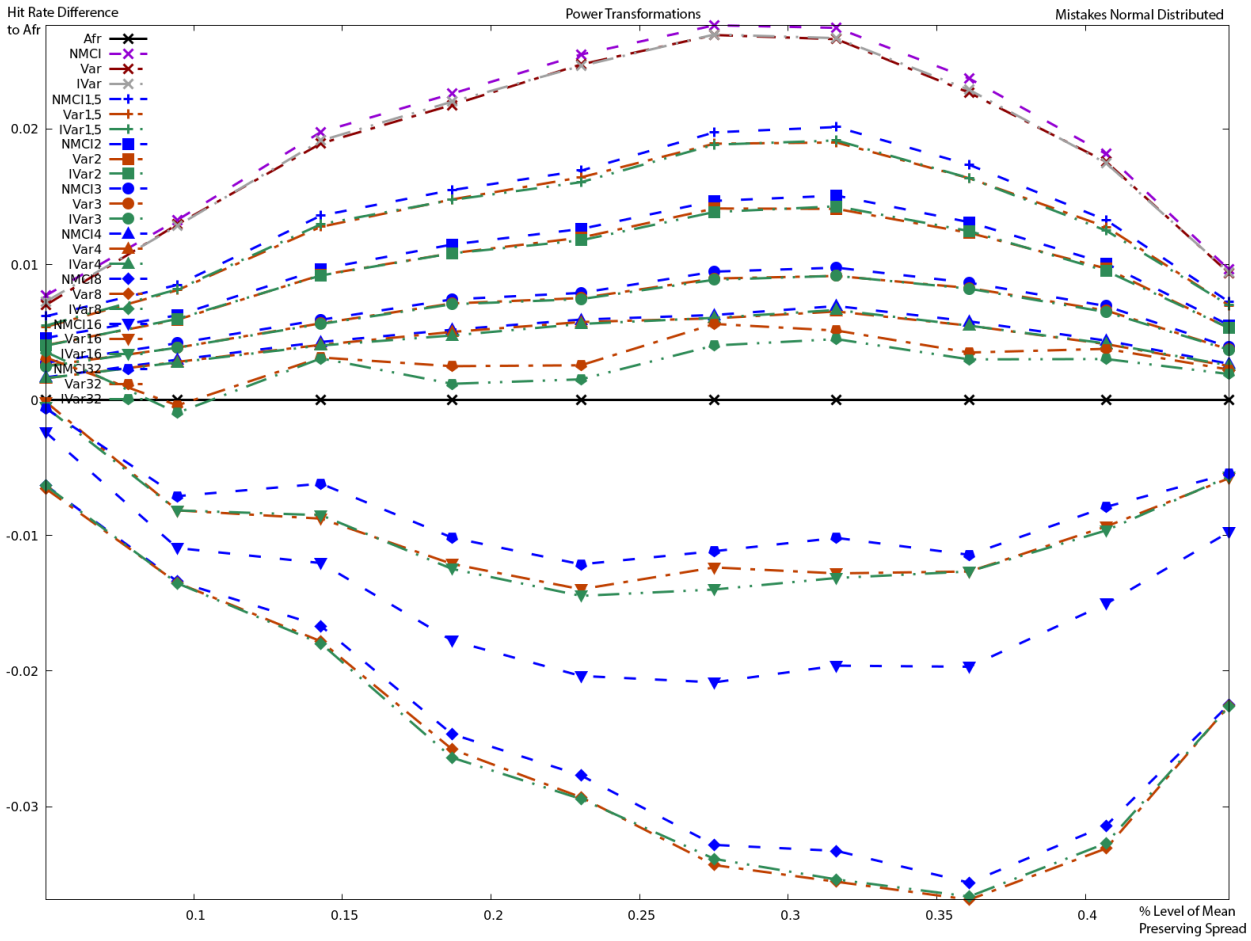

Figure 3.10. Convexifications of measures of rationality using power transformations for mistakes. Power transformations are varied between 1 and 32 and denoted in the labels. Measures' hit rate differences to Afriat's Index's hit rate when mistakes are drawn from Normal distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread.

Both figures show similar results. Any convexification makes the measures perform worse. Especially, with normally distributed mistakes at 8th degree power transformation, the measures perform significantly worse than Afriat's Efficiency Index.[8]

This decrease in performance of the measures of rationality is novel. These power transformations follow Varian's (1990) idea of using 2nd degree power transformation for the Varian's Index. Additionally, this convexification has symmetry to the use of the sum of

[8]The values seem to go above Afriat's Index with high degree power transformations. However, these values might have convergence problems from numerical limitations of very small numbers. The sizes of mistakes were multiplied by 10, which does not affect ordinal comparisons, to partially remedy this issue.

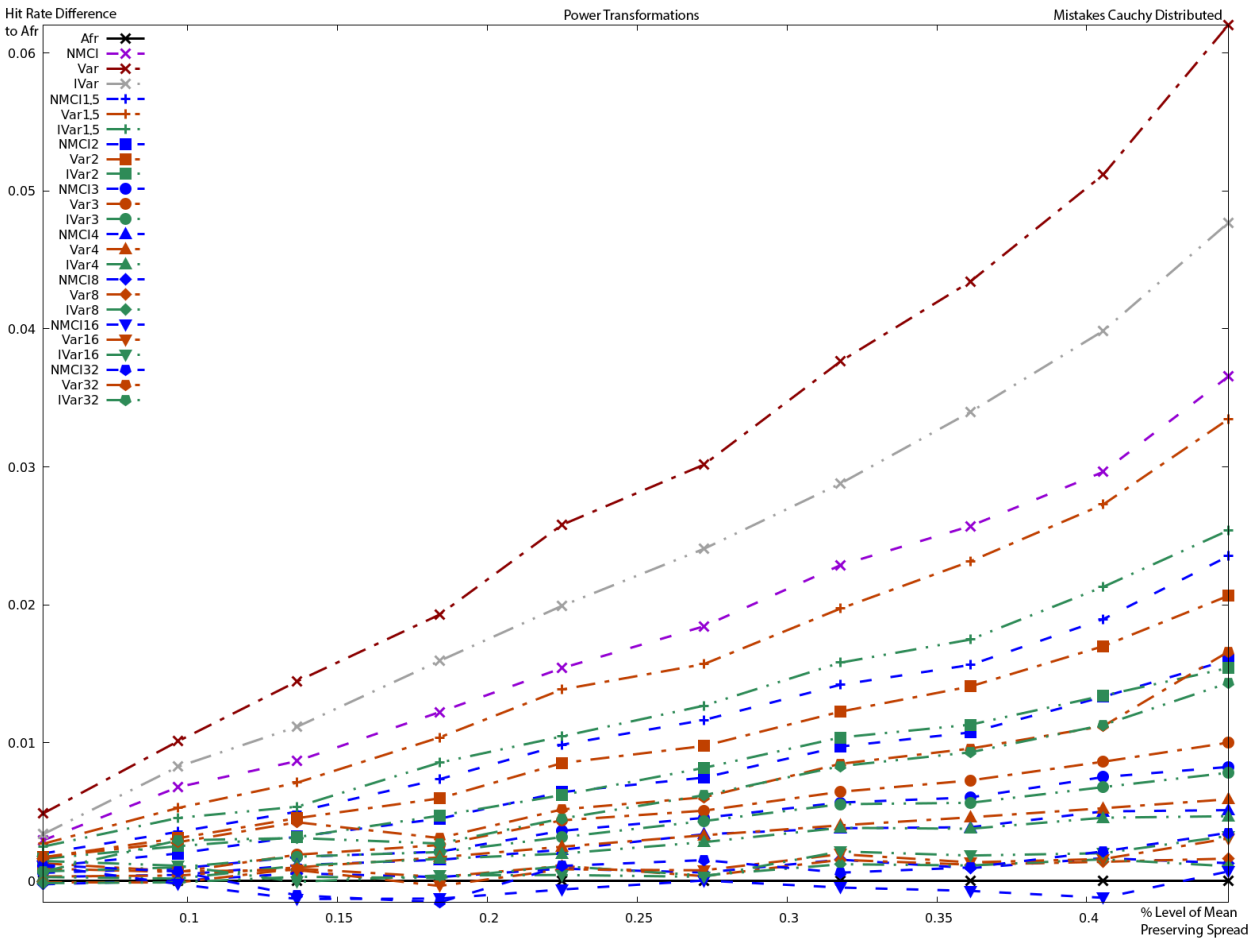

Figure 3.11. Convexifications of measures of rationality using power transformations for mistakes. Degrees of power transformations are varied between 1 and 32 and denoted in the labels. Measures' hit rate differences to Afriat's Index's hit rate when mistakes are drawn from Cauchy distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread.

squared errors in model fitting. Finally, this result suggests that a large mistake is not more indicative of lower rationality than multiple small mistakes.

This result is robust to all the variations of the simulations: Alternative distributions for the mistakes, alternative models of mistakes, CES utility function, uniform distribution of prices, and different numbers of goods and time periods.

Concavifications

Since the convexifications always made the measures perform worse, this raises the question if instead concavifications could improve the performance of the measures. We show that

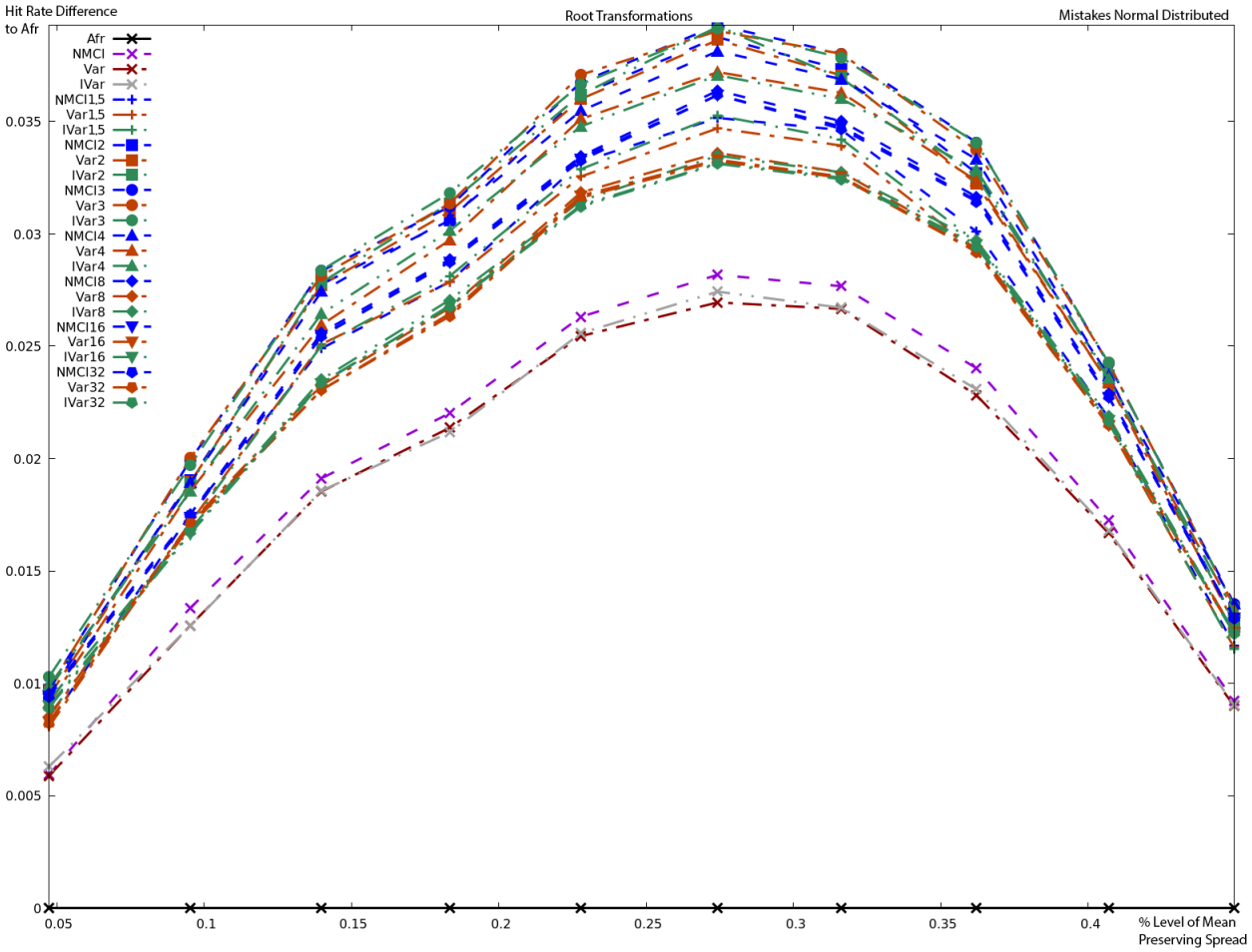

Figure 3.12. Concavifications of measures of rationality using root transformations for mistakes. Degrees of root transformations are varied between 1 and 32 and denoted in the labels. Measures' hit rate differences to Afriat's Index's hit rate when mistakes are drawn from Normal distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread.

this is the case but the optimal level of concavification depends on how fat-tailed the mistake distribution is.

Next, we test different root transformations of the above three measures up to 32 degrees. This is done in Figures 3.12 and 3.13 when mistakes are Normal and Cauchy distributed respectively. The number in the measures' labels is the degree of root transformation.

First, Figure 3.12 shows that with normally distributed mistakes some degree of root transformations increases the performance of the measures by up to 1.5 percentage points. The optimal degree of root transformations is around 3. At the optimal degree, the differences between the measures become very small.

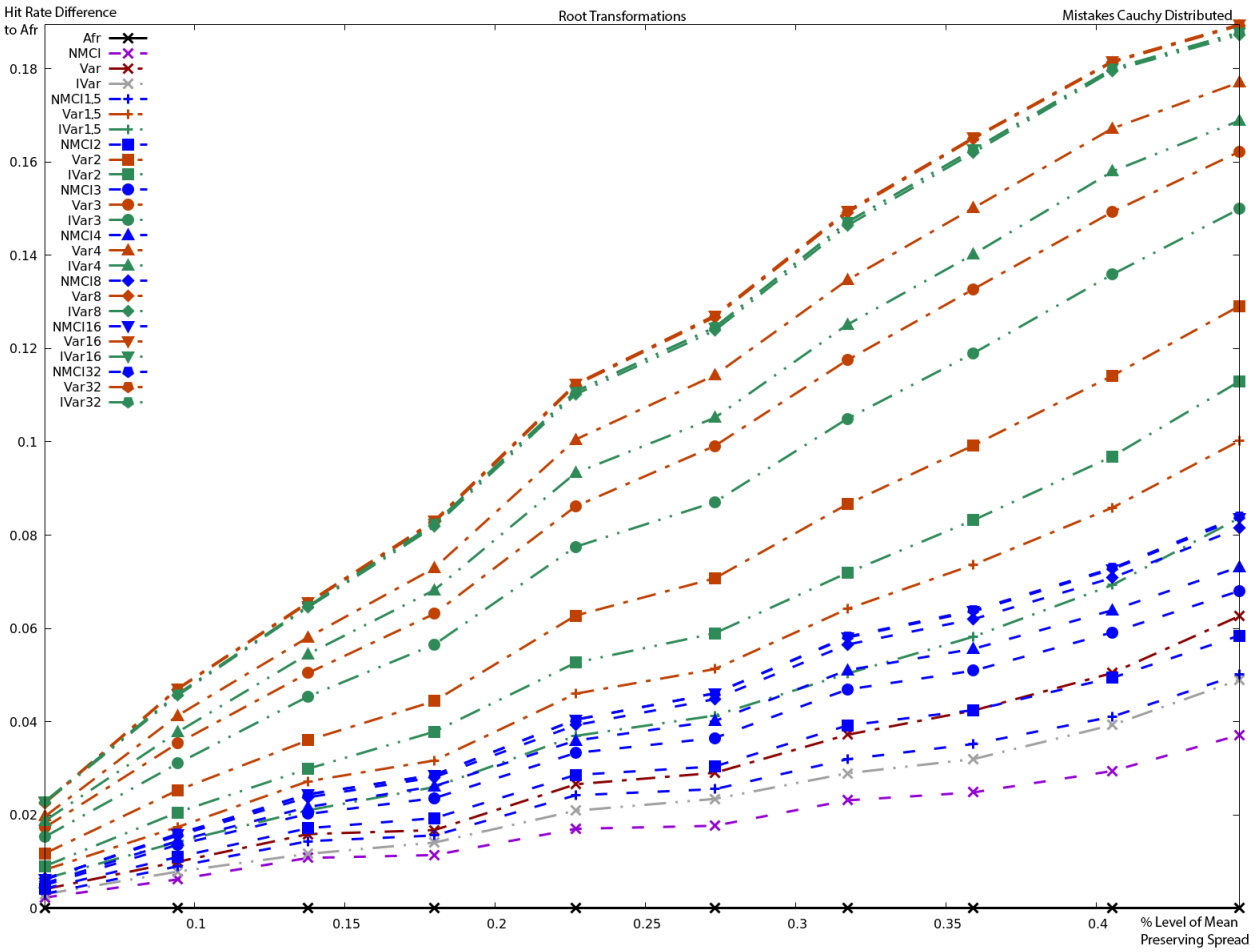

Figure 3.13. Concavifications of measures of rationality using root transformations for mistakes. Root transformations are varied between 1 and 32 and denoted in the labels. Measures' hit rate differences to Afriat's Index's hit rate when mistakes are drawn from Cauchy distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread.

After the optimal degree, additional degrees decrease the performance of the measures. However, even with high degrees of transformations of 32, which appears to be close to the limit measures, the measures significantly outperform the standard measures of rationality.

Second, Figure 3.13 shows that with Cauchy distributed mistakes higher degrees of root transformations always improve the performance of the measures by up to 14 percentage points. The performance improvements are the greatest for Varian's and Inverse Varian's Indexes that seem to converge to similar limiting measures. The improvements are much more limited for NMCI and up to 5 percentage points.

As discussed in Section 3.3.4 with fat-tailed mistakes the performance of the standard measures is worse than with concentrated mistakes and so there is much more room left for improvements. In Appendix Figures C.3 and C.4, we use absolute hit rates instead of differences to Afriat's Index's hit rate for Figures 3.12 and 3.13 respectively. Comparing those figures show that the highest degrees of root transformations with Cauchy distributed mistakes in Figure C.4 achieve almost the same level of performance as the measures with concentrated normally distributed mistakes in Figure C.3.

This result suggests that multiple small mistakes are more indicative of lower rationality than a single large mistake.

This result shows a connection between ordinal and cardinal measures of rationality. It shows how adding some weight to the ordinal information always improves the performance of the measures. However, the size of weight depends on how fat-tailed the distribution of mistakes is. If the mistakes are concentrated, cardinal information or intensities of violations are valuable at capturing the size of rationality violations and it is optimal to increase the weight of ordinal information only slightly. On the other hand, if the mistakes are fat-tailed, then cardinal information is less useful at capturing the size of rationality violations. But instead, it is optimal to prioritize the ordinal information and use the cardinal information only to break ties.

The intuition for this distinction is that with fat-tailed mistakes there are more big mistakes even at low mistake rates, so they are not useful at distinguishing between different mistake rates anymore. This highlights how the choice of measure of rationality depends on the application at hand and different measures are better at capturing different types of mistakes depending on how concentrated the mistakes are.

This result is robust to alternative mistake distributions and mistake models. Logit-choice mistakes or Trembling hand mistakes or observation dependent utility function parameters with concentrated distributions of Log-Normal, Uniform, Gumbel, or Reversed Weibull give the same results as normally distributed mistakes. On the other hand, fat-tailed mistakes

such as uniform choice on the budget plane or Trembling hand mistakes or observation dependent utility function parameters with fat-tailed distributions of Log-Cauchy or Fréchet give the same results as Cauchy distributed mistakes.

The results are robust to increasing or decreasing the number of time periods or observations in the data sets. The result that the limiting lexicographic measure outperforms standard measures with concentrated mistake distribution depends on the number of goods. With a low number of goods (3), the limiting measure does worse than the standard measures. Additionally, the optimal degree of root transformations is lower at around 2. With a higher number of goods (12), the optimal degree of root transformations is higher around 3 to 4 and the limiting measure outperforms the standard measures more.

Remark These results seem to suggest that using ordinal measures counting the number of revealed preferences or observations to remove while resolving ties by randomizing uniformly, could also perform well. This is, however, not the case and these ordinal measures with a coin perform worse than Afriat's Index. Cardinal information is crucial for tie-breaking.

3.3.6 Computation of Measures and Simulations

Computing most of the measures of rationality is computationally complex as discussed in Smeulders et al. (2014) except for Afriat's Efficiency Index. We develop a new program to calculate the measures of rationality in C++.[9] Our computational approach builds on the approach of Dean and Martin (2016) of representing the measures of rationality as an integer programming problem that can be efficiently solved. We extend this approach to Varian's Index and Inverse Varian's Index. This efficient solution approach with under 100 time periods gets around the complexity of the measures and allows us to do simulations on them. Additionally, we developed a novel method to compute Afriat's Efficiency Index based on a depth-first search. A pseudo algorithm for this is in Appendix Appendix C.2.

[9]The program is partly based on Dean and Martin's (2016) MCI calculator.

Dean and Martin's (2016) approach is to represent Minimum Cost Index as a cycle covering problem: The idea is to remove revealed preference relations so that all the cycles are covered and especially the remaining revealed preference graph is acyclical. This is solved using a relaxed problem with row generation iteratively: We solve a relaxed cycle covering problem with some cycles and check if solving this relaxed problem also solves the whole problem. If it does not solve it, then we find missing critical cycles and solve the new relaxed cycle covering problem.

We extend this approach to Varian's Index and Inverse Varian's Index. We reformulate Varian's Index and Inverse Varian's Index as a cycle covering problem. Instead of removing relations to cover cycles, we remove lower contour sets of relations to cover all the cycles. These lower contour sets of relations are all the cheaper relations at the same observation. This cycle covering problem can be solved symmetrically to the above one. However, this time some of the cycle constraints are redundant: One cycle's lower contour sets might contain some other cycles lower contour set always. So removing any one of the first cycle's lower contour sets removes also a lower contour set of the second cycle. We check for these redundancies and remove them to improve the performance.

We calibrate the levels of rationality for the populations compared using a simulated method of moments. Since choosing the mistake level is a unidimensional problem, we can implement the SMM as a root-finding problem. We solve the root-finding problem using a new combination of Brent's method (Brent, 1973) and Algorithm 748 (Alefeld et al., 1995) that combines inverse cubic and quadratic interpolation with double length secant and robust bisection methods.

The simulations for the SMM and for calculating the hit rates for the measures are done with parallel Monte-Carlo simulations. Each one of the Monte-Carlo simulations uses a (non-pseudo) random seed to make them independent.

3.4 Related Literature

We study the axiomatic foundations of measures of rationality. This literature was started by Apesteguia and Ballester (2015). They study the foundations of ordinal counting measures of rationality when we observe a continuum of data for random choices from general choice sets. Our approach complements this by providing the foundations for cardinal measures of rationality in the budget-choice setting using a finite number of observations.

Second, we use simulated choices to compare measures of rationality. Usually, simulated uniform random choices are used to provide a benchmark for measures of rationality using Selten's Score (Selten, 1991). We extend on this approach significantly to get objective sizes of mistakes using simulated choices with mistakes and compare how well different measures of rationality capture the sizes of these mistakes.

3.5 Conclusion

The assumption of rationality is a cornerstone of economic analysis. But how to measure how close reality is to satisfying the assumption of rationality is not clear. Instead, there is a multitude of ad hoc -measures of rationality. This paper provides a new perspective to different cardinal measures of rationality that capture the intensity of rationality violations. First, we provided theoretical foundations for Afriat's Efficiency Index, Normalized Money Pump Index, and Normalized Minimum Cost Index. Additionally, this approach can be extended to provide insight into the foundations for Varian's Index and Inverse Varian's Index.

This theoretical perspective highlighted the commonalities and differences among the measures. It showed a novel perspective to the similarity of Afriat's Index and Normalized Money Pump Index as two averaging measures where the first one is continuous and the second is linear that allows for discontinuities at new violations of rationality. Second, we showed that Varian's Index, Inverse Varian's Index, and Normalized Minimum Cost Index are more refined versions of Afriat's Index that do not take the average across violations.

Instead Normalized Minimum Cost Index is, essentially, characterized by a lack of negative interactions between different violations of rationality or sensitivity to every violation of rationality.

Second, we provided novel simulations for the continuous measures of rationality. We provided a model of choice mistakes and compared ordinally how well different measures capture the size of these objective mistakes. We calibrated the size of choice mistakes to an empirically reasonable range. The first major finding of this section was that the choice of an appropriate measure for capturing the mistakes depends on how dispersed the mistakes are. Normalized Minimum Cost Index slightly outperforms other measures on concentrated mistakes; whereas Varian's Index significantly outperforms other measures on fat-tailed dispersed mistakes. However, none of the measures does well on fat-tailed mistakes.

Next, we moved on to considering different variations of the measures of rationality. First, we considered convexifying the measures and giving more weight to bigger mistakes. This, contrary to suggestions in the literature, always made the measures perform worse in the simulations.

Second, we considered combining the measures with ordinal counting measures by concavifying them and giving more weight to smaller mistakes. Adding some more weight to smaller mistakes turned out to always improve the performance of the measures. Especially, under fat-tailed mistakes increasing the weight on smaller mistakes always is an improvement: The best measures used lexicographically first ordinal information on the number of rationality violations and second cardinal information on the intensities of rationality violations to rank the measures. This provided a novel perspective to the connection of ordinal and cardinal measures of rationality. The value of cardinal information depends on how fat-tailed or dispersed the mistakes are.

Appendix A

Appendix for Chapter 1

This appendix is organized as follows. Appendix A.1 shows the general identification result for the special cases and alternative representations of state dependent dual-self expected utility and state dependent dual-self variational expected utility. We do this by considering a representation consisting only of utility u and aggregator $I : u(H) \to \mathbb{R}$.

Appendix A.2 considers expected utility when we relax Independence to Weak Independence. We show that this gives expected utility with censoring for some utility intervals. That is there exist an affine utility u and a weakly increasing censoring function H such that for all $x, y \in \Delta(X)$,

$$x \succsim y \iff H(u(x)) \geq H(u(y)).$$

Appendix A.3 proves the existence of state dependent dual-self expected utility and state dependent dual-self variational expected utilities, Theorems 1.1 and 1.7. Appendix A.3.5 proves Remark 1.2.

Appendix A.4 proves the identification results for state dependent dual-self expected utility and state dependent dual-self variational expected utilities. First, Appendix A.4.1 shows the dual-self variational representation is additive between any equally crisp acts. Second, Appendix A.4.2 shows how we can capture beliefs or derivatives of the representation behaviorally. Third, Appendix A.4.3 shows the general partial identification results, Theorems 1.5 and 1.9. Next, Appendix A.4.4 shows the full-identification results under Axiom 1.7, Theorems 1.2 and 1.8. Next, Appendix A.4.5 shows how the identification relates to underlying preferences and crisp and equally crisp acts with Corollaries 1.6 and 1.11. Finally, Appendix A.4.6 shows the relative likelihood and probability charac-

terizations by showing how the general partial identification result simplifies in some cases with Propositions 1.3 and 1.4.

Appendix A.5 proves our application to social choice theory and characterizes symmetric state dependent maxmin expected utility with Theorem 1.12.

Finally, Appendix A.6 considers state dependent Bewley expected utility. We axiomatize its existence and characterize its uniqueness in Appendices A.6.1 and A.6.2 respectively. Appendix A.6.3 axiomatizes state independent Bewley expected utility and shows how the state dependent version only relaxes monotonicity with between states comparisons.

A.1 General Identification

This section shows the general identification results for any aggregator function $I : u(H) \to \mathbb{R}$. This shows the identification for all the special cases and alternative representations of state dependent dual-self expected utility and dual-self variational expected utility.

For this section, we assume that $S = S^{\mathcal{P}}$. The results extend to situations where not all states are proper and utilities are not identified on these states.

For $\psi, \varphi \in \mathbb{R}^S$, denote $\varphi \cdot \psi = \sum_{s \in S} \varphi_s \psi_s$. Let $A \subseteq \mathbb{R}^S$ be a convex set and $I : A \to \mathbb{R}$ be a function. We denote $\bar{1} \in \mathbb{R}^S$ such that for all $s \in S$, $\bar{1}_s = 1$. We say that

- For state dependent utility u, (u, I) is a state dependent representation for $\succsim$ if for all $f, g \in H$, $f \succsim g \iff I(u(f)) \geq I(u(g))$.
- I is C-additive if for all $\varphi \in A, \alpha \geq 0$ such that $\varphi + \alpha\bar{1} \in A$, $I(\varphi + \alpha\bar{1}) = I(\varphi) + \alpha$.
- I is positive homogeneous if for all $\varphi \in A, \alpha > 0$ such that $\alpha\varphi \in A$, $I(\alpha\varphi) = \alpha I(\varphi)$.
- I is monotonic if for all $\varphi, \psi \in A$ such that for all $s \in S$, $\varphi_s \geq \psi_s$, $I(\varphi) \geq I(\psi)$.

Dual-self variational aggregators correspond to C-additive and monotonic aggregators. Dual-self aggregators correspond to C-additive, positive homogeneous, and monotonic aggregators.

Let $A \subseteq \mathbb{R}^S$ be a convex set. For every $\varphi \in \operatorname{int} A, \xi \in \mathbb{R}^S$, the *Clarke upper derivative* of I at φ in the direction ξ is

$$I^\circ(\varphi; \xi) = \limsup_{\psi \to \varphi, t \searrow 0} \frac{I(\psi + t\xi) - I(\psi)}{t}.$$

The Clarke subdifferential of I at φ is the set

$$\partial I(\varphi) = \{\chi \in \mathbb{R}^S \,|\, \forall\, \xi \in \mathbb{R}^S, \chi \cdot \xi \leq I^\circ(\varphi; \xi)\}.$$

Ghirardato et al. (2004) has shown that the union of all Clarke subdifferentials, $\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)$, captures the decision maker's beliefs.

Our first result generalizes the partial identification for state dependent dual-self variational expected utility, Theorem 1.7. Here the identification is the same but for the general aggregator I.

Proposition A.1 Let (u, I) and $(\widetilde{u}, \widetilde{I})$ be state dependent monotonic and C-additive representations for $\succsim$. Then there exist $x \in \left(\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)\right)^{\perp}, \alpha > 0, B \in \mathbb{R}^S, \beta \in \mathbb{R}$ such that for all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$,

$$\partial\widetilde{I}\big(u(f)\big) = (1+x)\partial I\big(\widetilde{u}(f)\big),$$

especially

$$\bigcup_{\varphi\in\operatorname{int}\widetilde{u}(H)} \partial\widetilde{I}(\varphi) = (1+x) \bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi),$$

for all $s \in S$,

$$\widetilde{u}_s = \frac{\alpha}{1+x_s} u_s + B_s,$$

and for all $f \in H$,

$$\widetilde{I}\big(\widetilde{u}(f)\big) = \alpha I\big(u(f)\big) + \beta.$$

Our second result generalizes the partial identification for state dependent dual-self expected utility, Theorem 1.5.

Proposition A.2 Let (u, I) and $(\widetilde{u}, \widetilde{I})$ be state dependent monotonic, C-additive, and positive homogeneous representations for $\succsim$. Then there exist $x, y \in \left(\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)\right)^{\perp}, \alpha > 0, \beta \in \mathbb{R}$ such that for all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$,

$$\partial\widetilde{I}\big(u(f)\big) = (1+x)\partial I\big(\widetilde{u}(f)\big),$$

especially

$$\bigcup_{\varphi\in\operatorname{int}\widetilde{u}(H)} \partial\widetilde{I}(\varphi) = (1+x) \bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi),$$

for all $s \in S$,

$$\widetilde{u}_s = \frac{\alpha}{1+x_s}(u_s + y_s) + \beta,$$

and for all $f \in H$,

$$\widetilde{I}\big(\widetilde{u}(f)\big) = \alpha I\big(u(f)\big) + \beta.$$

These two partial identifications give all the following full identifications and identifications for probabilities for a single state and relative likelihoods between two states. The next result shows the full identification and gives Theorems 1.2 and 1.8.

Proposition A.3 Let (u, I) be a state dependent monotonic, C-additive representation for $\succsim$. The following conditions are equivalent:

1. $\succsim$ satisfies Axiom 1.7.
2. $\left(\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)\right)^{\perp} = \bar{0}$.
3. $\overline{\operatorname{co}}\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)$ has a non-empty interior in $\Delta(S)$.

The next result shows the identification for the probabilities for a single state. This generalizes Proposition 1.3 to variational representation.

Proposition A.4 Let (u, I) be a state dependent monotonic, C-additive representation for $\succsim$ and $s \in S^{\mathcal{P}}$. The following conditions are equivalent:

1. If $f, g \in H$ are such that $f \sim g$, $f_s \succ_s g_s$, and there exists $s' \in S$ such that $g_{s'} \succ_{s'} f_{s'}$, then there exist $h \in H$ and $\alpha \in (0, 1)$ such that
$$\alpha h + (1-\alpha)f \not\sim \alpha h + (1-\alpha)g.$$
2. $\operatorname{pr}_s \left(\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)\right)^{\perp} = \{0\}$.
3. $S^{\mathcal{P}} = \{s\}$ or there exist $p, q \in \overline{\operatorname{co}}\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)$ such that $p_s \neq q_s$ and for all $\tilde{s} \in S \setminus \{s\}$, $\frac{p_{\tilde{s}}}{1-p_s} = \frac{q_{\tilde{s}}}{1-q_s}$.

The last result shows the identification for the relative likelihoods for a single state. This generalizes Proposition 1.4 to variational representation.

Proposition A.5 Let (u, I) be a state dependent monotonic, C-additive representation for $\succsim$ and $s, s' \in S^{\mathcal{P}}, s \neq s'$. The following conditions are equivalent:

1. If $f, g \in H$ are such that and $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$, then there exist $h, h' \in H$ and $\alpha \in (0, 1)$ such that
$$\alpha h + (1-\alpha)f \succ \alpha h' + (1-\alpha)f \text{ and } \alpha h + (1-\alpha)g \prec \alpha h' + (1-\alpha)g.$$
2. Either
$$\operatorname{pr}_{s,s'} \Big(\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)\Big)^{\perp} = \{(0,0)\} \text{ or } \operatorname{pr}_{s,s'} \Big(\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)\Big)^{\perp} = \{(a,a) | a \in \mathbb{R}\}.$$
3. There exist $p, q \in \overline{\operatorname{co}}\bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi)$ such that $\frac{p_s}{p_{s'}} \neq \frac{q_s}{q_{s'}}$ and for all $\tilde{s} \in S \setminus \{s, s'\}$, $p_{\tilde{s}} = q_{\tilde{s}}$.

A.2 Latent Expected Utility Representation

For this section, we assume that X is a nonempty set, $\Delta(X)$ is the set of all (simple) lotteries on X, $H = \Delta(X)$, and $\succsim \subseteq H \times H$. This section studies the following weak independence assumption.

Axiom 4′ For all $f, g, h \in H, \alpha \in (0,1)$,

$$f \succ g \Longrightarrow \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h.$$

We will show that under continuity and weak order assumptions, Weak Independence characterizes censored expected utility where some utility differences are censored with the function H below. This gives a latent expected utility model.

Theorem A.6 (Latent Expected Utility) $\succsim$ satisfies Axioms 1.1, 1.2, 4′ iff there exist affine $u : \Delta(X) \to \mathbb{R}$ and continuous and weakly increasing $H : u(\Delta(X)) \to \mathbb{R}$ such that for all $P, Q \in \Delta(X)$,

$$P \succsim Q \iff H\Big(\sum_x P(x)u(x)\Big) \geq H\Big(\sum_x Q(x)u(x)\Big).$$

Additionally, if $(\widetilde{u}, \widetilde{H})$ is another censored expected utility for $\succsim$, then there exist $a > 0, b \in \mathbb{R}$ such that $\widetilde{u} = au + b$ and a strictly increasing $\varphi : H(u(\Delta(X))) \to \mathbb{R}$ such that for all $v \in u(\Delta(X))$,

$$\widetilde{H}(av + b) = \varphi(H(v)).$$

The uniqueness part of this theorem shows that the affine utility is unique up to a positive affine transformation and we can determine for every lottery $P \in \Delta(X)$ if it is censored. That is the intervals on which the utility is censored are unique given the utility. However, the values of the censoring function are not unique and any ordinal transformation are allowed.

Secondly, in this representation, u gives a weak affine representation for $\succsim$.

Definition u is a *weak affine representation for* $\succsim$ if $u : \Delta(X) \to \mathbb{R}$ is affine and for all $x, y \in \Delta(X)$,

$$P \succ Q \Longrightarrow \sum_x P(x)u(x) > \sum_x Q(x)u(x).$$

A.2.1 Proof for Latent Expected Utility

We denote the elements of $\Delta(X)$ by x, y. We denote for $\alpha \in [0,1]$, $x\alpha y = \alpha x + (1-\alpha)y$. We start with simple lemmas. The first shows that there are strictly middle points of preferences in convex combinations.

Lemma A.7 Assume that $\succsim$ satisfies Axioms 1.1, 1.2. If $x \succ y$, then there exist $\alpha^1, \alpha^2 \in (0,1)$ such that

$$x \succ x\alpha^1 y, x\alpha^2 y \succ y$$

and for all $\alpha \in (\alpha^1, 1], \alpha' \in [0, \alpha^2)$,

$$x\alpha y \succ x\alpha^1 y \text{ and } x\alpha^2 y \succ x\alpha' y.$$

Proof. Let $x \succ y$. By Axiom 1.2, the set

$$\{\alpha \in [0,1] | x \succ \alpha x + (1-\alpha)y\} \text{ is open and } \{\alpha \in [0,1] | y \succsim \alpha x + (1-\alpha)y\} \text{ is closed.}$$

Thus there exists $\alpha^0 \in (0,1)$ such that $\alpha^0 \in \{\alpha \in [0,1] | x \succ \alpha x + (1-\alpha)y\}$ and $\alpha^0 \notin \{\alpha \in [0,1] | y \succsim \alpha x + (1-\alpha)y\}$. Thus $x \succ x\alpha^0 y \succ y$. Next denote

$$\alpha^1 = \max\{\alpha \in [0,1] | (x\alpha^0 y) \succsim \alpha x + (1-\alpha)(x\alpha^0 y)\},$$

$$\alpha^2 = \max\{\alpha \in [0,1] | \alpha y + (1-\alpha)(x\alpha^0 y) \succsim (x\alpha^0 y)\}.$$

Now define $\alpha^3 = \alpha^1 + (1-\alpha^1)\alpha^0$ and $\alpha^4 = (1-\alpha^2)\alpha^0$. Then we have

$$x\alpha^0 y \sim x\alpha^3 y \sim x\alpha^4 y$$

and for all $\alpha \in (\alpha^3, 1), \alpha' \in (0, \alpha^4)$,

$$x\alpha y \succ x\alpha^3 y \text{ and } x\alpha^4 y \succ x\alpha' y.$$

□

The next two lemmas show that indifference sets are convex and strict upper and lower contour sets are convex.

Lemma A.8 Assume that $\succsim$ satisfies Axioms 1.1, 1.2, 4′. If $x \sim y$, then for all $\alpha \in (0,1)$, $x \sim \alpha x + (1-\alpha)y$.

Proof. Assume, per contra, there exist $x, y \in \Delta(X), \alpha^0 \in (0,1)$ such that $x \sim y$ and $x \not\sim \alpha^0 x + (1-\alpha^0)y$. Assume w.l.o.g. $\alpha^0 x + (1-\alpha^0)y \succ x$. By Lemma A.7, there exists $\alpha^1 \in (\alpha^0, 1)$ such that for all $\alpha \in (\alpha^1, 1)$,

$$x\alpha^1 y \succ x\alpha y. \tag{A.1}$$

Now $x\alpha^0 y \succ y$ and thus by Weak Independence,

$$x\alpha^1 y \precsim x\alpha^1(x\alpha^0 y) \equiv x\big(\alpha^1 + (1-\alpha^1)\alpha^0\big)y \overset{(A.1)}{\prec} x\alpha^1 y,$$

which contradicts transitivity. □

Lemma A.9 Assume that $\succsim$ satisfies Axioms 1.1, 1.2, 4′. If $x \succ z$ and $y \succ z$, then for all $\alpha \in (0,1)$, $\alpha x + (1-\alpha)y \succ z$. If $z \succ x$ and $z \succ y$, then for all $\alpha \in (0,1)$, $z \succ \alpha x + (1-\alpha)y$.

Proof. Assume, per contra, there exist $x, y, z \in \Delta(X), \alpha^0 \in (0,1)$ such that $x \succ z$,$y \succ z$ and $\alpha^0 x + (1-\alpha^0)y \sim z$. By Lemma A.7, there exists $\alpha^1 \in (\alpha^0, 1)$ such that for all $\alpha \in (\alpha^1, 1)$,

$$x\alpha y \succ x\alpha^1 y. \tag{A.2}$$

Now $y \succ x\alpha^0 y$ and thus by Weak Independence,

$$x\alpha^1 y \succsim x\alpha^1(x\alpha^0 y) \equiv x\big(\alpha^1 + (1-\alpha^1)\alpha^0\big)y \overset{(A.2)}{\succ} x\alpha^1 y,$$

which contradicts transitivity. The other case follows symmetrically. □

Next, we show the existence of an affine utility that gives a weak representation for the preferences. This follows similarly to vNM theorem but in here we define utility at a point where preferences are strictly increasing.

Proposition A.10 Assume that $\succsim$ satisfies Axioms 1.1, 1.2, 4′. Then there exists $u : \Delta(X) \to \mathbb{R}$ that is a weak affine representation for $\succsim$.

Proof. By nontriviality, there exist $x^*, x_* \in \Delta(X)$ such that $x^* \succ x_*$. By applying Lemma A.7 twice, there exist $\alpha^*, \alpha_* \in (0,1)$ such that $\alpha^* > \alpha_*$ and for all $\alpha \in (\alpha^*, 1], \alpha' \in [0, \alpha_*)$,

$$x^*\alpha x_* \succ x^*\alpha^* x_* \succ x^*\alpha_* x_* \succ x^*\alpha' x_*.$$

Next, we define the representation by considering cases. If $x \succ x^*\alpha^* x_*$,

$$\alpha^x = \min\{\alpha \in [0,1] | \alpha x + (1-\alpha)(x^*\alpha_* x_*) \succ x^*\alpha^* x_*\}$$

and define

$$u(x) \coloneqq \frac{1}{\alpha^x}.$$

Now especially $u(x) > 1$. If $x^*\alpha_* x_* \succ x$,

$$\alpha^x = \max\{\alpha \in [0,1] | x^*\alpha_* x_* \succ \alpha(x^*\alpha^* x_*) + (1-\alpha)x\}$$

and define

$$u(x) \coloneqq \frac{-\alpha^x}{1-\alpha^x}.$$

Now especially $u(x) < -1$. If $x^*\alpha^* x_* \succsim x \succsim x^*\alpha_* x_*$,

$$\alpha^x = \max\{\alpha \in [0,1] | x^*\alpha^* x_* \succsim \alpha x^* + (1-\alpha)x\}$$

and define

$$u(x) \coloneqq \frac{1-u(x^*)\alpha^x}{1-\alpha^x}.$$

Now especially $1 \leq u(x) \leq -1$. We show that u is a weak representation for $\succsim$. Let $x, y \in \Delta(X)$ be such that $u(x) \geq u(y)$. Assume, per contra, $y \succ x$. If $y \succ x^*\alpha^* x_* \succsim x$ or $y \succsim x^*\alpha_* x_* \succ x$, then $u(x) > u(y)$. So we have three cases to consider. 1) $y \succ x \succ x^*\alpha^* x_*$. Now there exists $\alpha^0 \in (0,1)$ such that $y\alpha^0(x^*\alpha_* x_*) \succ x$. Then by Weak Risk Independence and Lemma A.8, for all $\alpha \in (0, \alpha^x)$,

$$\alpha(x^*\alpha_* x_*) + (1-\alpha)y(\alpha^0(x^*\alpha_* x_*)) \succsim \alpha(x^*\alpha_* x_*) + (1-\alpha)x \succ x^*\alpha^* x_*.$$

Thus $\alpha^y \geq \alpha^x + (1-\alpha^x)(1-\alpha^0)$ which contradicts $u(x) \geq u(y)$. 2) $x^*\alpha^* x_* \succ y \succ x$ follows symmetrically since there exists $\alpha^0 \in (0,1)$ such that $y \succ x\alpha^0(x^*\alpha^* x_*)$. 3) $x^*\alpha^* x_* \succsim y \succ x \succsim x^*\alpha_* x_*$ follows symmetrically since there exists $\alpha^0 \in (0,1)$ such that $y \succ x\alpha^0 x^*$.

We show the linearity of u within each of the three above areas. 1) Let $y, x \succ x^*\alpha^* x_*, \alpha \in (0,1)$ and $z \coloneqq \alpha x + (1-\alpha)y$. Next, we show that $\alpha^z = \frac{\alpha^x\alpha^y}{\alpha^x(1-\alpha)+\alpha\alpha^y}$. First, by Lemma A.8,

$$\begin{aligned}
&\frac{\alpha\alpha^y}{\alpha^x(1-\alpha)+\alpha^y\alpha}\big(\alpha^x x + (1-\alpha^x)(x^*\alpha_* x_*)\big) + \frac{\alpha^x(1-\alpha)}{\alpha^x(1-\alpha)+\alpha^y\alpha}\big(\alpha^y y + (1-\alpha^y)(x^*\alpha_* x_*)\big) \\
\equiv &\frac{\alpha^x\alpha^y}{\alpha^x(1-\alpha)+\alpha^y\alpha} z + \Big(1 - \frac{\alpha^x\alpha^y}{\alpha^x(1-\alpha)+\alpha^y\alpha}\Big)(x^*\alpha_* x_*) \sim x^*\alpha^* x_*.
\end{aligned}$$

Additionally, for all $\alpha > \frac{\alpha^x\alpha^y}{\alpha^x(1-\alpha)+\alpha\alpha^y}$, by Lemma A.9,

$$\alpha z + (1-\alpha)(x^*\alpha_* x_*) \succ x^*\alpha^* x_*.$$

Then the linearity follows from the definition of u. The cases $x^*\alpha^* x_* \succsim x, y \succsim x^*\alpha_* x_*$ and $x^*\alpha^* x_* \succ x, y$ follow similarly.

Next we show that u is linear between x^* and x_*. By the choice of α^* and α_*, we have for all $\alpha \in (\alpha^*, 1]$, $x^*\alpha x_* \succ x^*\alpha^* x_*$ and

$$\alpha^{x^*\alpha x_*} = \frac{\alpha^* - \alpha_*}{\alpha - \alpha_*}.$$

Thus

$$u(x^*\alpha x_*) = \frac{\alpha - \alpha_x}{\alpha^* - \alpha_*}.$$

We have for all $\alpha \in [\alpha_*, \alpha^*]$, $x^*\alpha^* x_* \succsim x^*\alpha x_* \succsim x^*\alpha_* x_*$ and

$$\alpha^{x^*\alpha x_*} = \frac{\alpha^* - \alpha}{1-\alpha}.$$

Thus

$$u(x^*\alpha x_*) = \frac{1 - \frac{1-\alpha_*}{\alpha^*-\alpha_*}\frac{\alpha^*-\alpha}{1-\alpha}}{1 - \frac{\alpha^*-\alpha}{1-\alpha}} = \frac{\alpha - \alpha_*}{\alpha^* - \alpha_*}$$

We have for all $\alpha \in [0, \alpha_*)$, $x^*\alpha_* x_* \succ x^*\alpha x_*$ and

$$\alpha^{x^*\alpha x_*} = \frac{\alpha_* - \alpha}{\alpha^* - \alpha}.$$

Thus

$$u(x^*\alpha x_*) = \frac{-\frac{\alpha_* - \alpha}{\alpha^* - \alpha}}{1 - \frac{\alpha_* - \alpha}{\alpha^* - \alpha}} = \frac{\alpha - \alpha_*}{\alpha^* - \alpha_*}.$$

Thus the utility is linear between x^* and x_*. Now let $x, y \in \Delta(X)$ be such that $x \succ x^*\alpha^* x_* \succsim y \succsim x^*\alpha_* x_*$ and we show that u is affine between x and y. Let $\alpha^1 \in (\alpha^*, 1]$ be such that $x \succ x^*\alpha^1 x_*$. Then there exist $\alpha^2 \in (0,1)$ such that $u(x\alpha^2 y) = u(x^*\alpha^1 x_*)$ and $\alpha^3 \in [\alpha_*, \alpha^*]$ such that $u(y) = u(x^*\alpha^3 x_*)$. Now since u is a weak representation, we have by the linearity, for all $z, z' \in \Delta(X)$ such that $z, z' \succ x^*\alpha^* x_*$ and for all $\alpha \in (\frac{\alpha^* - \alpha^3}{\alpha^1 - \alpha^3}, 1]$,

$$u(x\alpha^2\alpha y) = u\Big(x^*\big(\alpha^1\alpha + \alpha^3(1-\alpha)\big)x_*\Big).$$

And by the linearity, for all $z, z' \in \Delta(X)$ such that $x^*\alpha^* x_* \succsim z, z' \succsim x^*\alpha^* x_*$ and for all $\alpha \in [0, \frac{\alpha^* - \alpha^3}{\alpha^1 - \alpha^3}]$,

$$u(x\alpha^2\alpha y) = u\Big(x^*\big(\alpha^1\alpha + \alpha^3(1-\alpha)\big)x_*\Big).$$

Thus by the linearity of u between x^* and x_*, u is linear between $x\alpha^2 y$ and y. Additionally, by the linearity, for all $z, z' \in \Delta(X)$ such that $z, z' \succ x^*\alpha^* x_*$, u is linear between x and $x\Big(\alpha^2 \frac{1/2(\alpha^1 + \alpha^*) - \alpha^3}{\alpha^1 - \alpha^3}\Big)y$. Thus u is linear between x and y.

Symmetrically if $x, y \in \Delta(X)$ are such that $x^*\alpha^* x_* \succsim x \succsim x^*\alpha_* x_* \succ y$, then u is affine between x and y. Thus finally for all $x, y \in \Delta(X)$, u is affine between x and y. This shows the claim. □

The next result shows the necessity of Weak Independence for an affine weak representation.

Lemma A.11 Assume that $\succsim$ satisfies Axioms 1.1, 1.2. If there exists $u : \Delta(X) \to \mathbb{R}$ that is a weak affine representation for $\succsim$, then $\succsim$ satisfies Axiom 4′.

Proof. Assume, per contra, there exist $x, y, z \in \Delta(X), \alpha \in (0,1)$ such that $x \succ y$ and $\alpha y + (1-\alpha)z \succ \alpha x + (1-\alpha)z$. Then since u is a weak representation for $\succsim$ and affine, we have

$$u(x) > u(y) \text{ and } \alpha u(y) + (1-\alpha)u(z) = u(\alpha y + (1-\alpha)z) > u(\alpha x + (1-\alpha)z) = \alpha u(x) + (1-\alpha)u(z)$$

which is a contradiction. □

The next result shows the uniqueness of the affine weak representation.

Proposition A.12 Assume that $\succsim$ satisfies Axioms 1.1, 1.2, 4′. If $u : \Delta(X) \to \mathbb{R}$ and $\widetilde{u} : \Delta(X) \to \mathbb{R}$ are weak affine representations for $\succsim$, then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that

$$\widetilde{u} = \alpha u + \beta.$$

Proof. First, we show that for all $x, y \in \Delta(X)$,

$$u(x) \geq u(y) \iff \widetilde{u}(x) \geq \widetilde{u}(y).$$

Assume, per contra, there exist $\widetilde{u}(x) > \widetilde{u}(y)$ and $u(x) \leq u(y)$. By Axiom 1.1, there exist $x^*, x_* \in \Delta(X)$ such that $x^* \succ x_*$. By applying Lemma A.7 twice, there exist $\alpha^*, \alpha_* \in (0,1)$ such that $\alpha^* > \alpha_*$ and for all $\alpha \in (\alpha^*, 1], \alpha' \in [0, \alpha_*)$,

$$x^* \alpha x_* \succ x^* \alpha^* x_* \succ x^* \alpha_* x_* \succ x^* \alpha' x_*. \tag{A.3}$$

Now since u is a weak representation and so

$$u(x^*) > u(x_*), \tag{A.4}$$

and either $u(x^* \alpha^* x_*) \geq u(x)$ or $u(x^* \alpha_* x_*) \leq u(y)$. Assume w.l.o.g. $u(x^* \alpha^* x_*) \geq u(x)$. By the affinity of u and $\widetilde{u}$, there exist α^x, α^y such that

$$\widetilde{u}(x \alpha^x x_*) > \widetilde{u}(y \alpha^y x^*) \tag{A.5}$$

and thus $u(x \alpha^x x_*) < u(y \alpha^y x^*)$. Denote $\widetilde{x} \coloneqq x \alpha^x x_*$ and $\widetilde{y} \coloneqq y \alpha^y x^*$. By (A.4), there exists $\alpha^0 \in (0,1]$ such that $u(x^* \alpha^* x_*) = u(\widetilde{x} \alpha^0 x^*)$. Thus especially by the affinity of u, we have $u(\widetilde{y} \alpha^0 x^*) > u(x^* \alpha^* x_*)$. By affinity of u, there exists $\alpha^1 \in (\alpha^*, 1]$ such that $u(\widetilde{y} \alpha^0 x^*) > u(x^* \alpha^1 x_*)$. Thus since u is a weak representation for $\succsim$, we have

$$\widetilde{y} \alpha^0 x^* \succsim x^* \alpha^1 x_* \overset{(A.3)}{\succ} x^* \alpha^* x_* \sim \widetilde{x} \alpha^0 x^*.$$

Thus especially since $\widetilde{u}$ is a weak representation for $\succsim$

$$\alpha^0 \widetilde{u}(\widetilde{y}) + (1 - \alpha^0) u(x^*) = \widetilde{u}(\widetilde{y} \alpha^0 x^*) > \widetilde{u}(\widetilde{x} \alpha^0 x^*) = \alpha^0 \widetilde{u}(\widetilde{x}) + (1 - \alpha^0) u(x^*).$$

However, this contradicts (A.5).

Thus for all $x, y \in \Delta(X)$,

$$u(x) \geq u(y) \iff \widetilde{u}(x) \geq \widetilde{u}(y).$$

Define $\trianglerighteq \subseteq \Delta(X) \times \Delta(X)$ by

$$x \trianglerighteq y \iff u(x) \geq u(y).$$

Now u and $\widetilde{u}$ are affine representations for $\trianglerighteq$ and hence by Herstein and Milnor (1953) and Fishburn (1970) there exist $\alpha > 0, \beta \in \mathbb{R}$ such that $\widetilde{u} = \alpha u + \beta$. □

Finally, the weak affine representation gives us a full representation with a censoring function.

Theorem A.6 (Latent Expected Utility) $\succsim$ satisfies Axioms 1.1, 1.2, $4'$ iff there exist affine $u : \Delta(X) \to \mathbb{R}$ and continuous and weakly increasing $H : u(\Delta(X)) \to \mathbb{R}$ such that for all $P, Q \in \Delta(X)$,

$$P \succsim Q \iff H\Big(\sum_x P(x)u(x)\Big) \geq H\Big(\sum_x Q(x)u(x)\Big).$$

Additionally, if $(\widetilde{u}, \widetilde{H})$ is another censored expected utility for $\succsim$, then there exist $a > 0, b \in \mathbb{R}$ such that $\widetilde{u} = au + b$ and a strictly increasing $\varphi : H(u(\Delta(X))) \to \mathbb{R}$ such that for all $v \in u(\Delta(X))$,

$$\widetilde{H}(av + b) = \varphi(H(v)).$$

Proof. We first show the sufficiency of the axioms. By Proposition A.10, there exists $u : \Delta(X) \to \mathbb{R}$ that is affine and a weak representation for $\succsim$. Define $\trianglerighteq \subseteq u(\Delta(X)) \times u(\Delta(X))$ by for all $x, y \in \Delta(X)$,

$$u(x) \trianglerighteq u(y) \iff x \succsim y.$$

First, $u(\Delta(X))$ is an interval since u is affine. Second, $\trianglerighteq$ is well-defined that is there does not exist $x, y \in \Delta(X)$ such that $u(x) = u(y)$ and $x \not\sim y$. If $x \not\sim y$, then since u is a weak representation for $\succsim$ and $\succsim$ is complete, we have $x \succ y$ or $y \succ x$ and so $u(x) > u(y)$ or $u(y) > u(x)$ respectively.

By Axiom 1.1, $\trianglerighteq$ is complete and transitive.

We show that $\trianglerighteq$ is continuous. We show that upper contour sets of preferences are closed. The lower contour sets follow symmetrically. Let $(x_i)_{i=1}^{\infty} \subseteq \Delta(X), x, y \in \Delta(X)$ be such that $u(x_i) \to u(x)$ as $i \to \infty$. Assume that for all $i \in \mathbb{N}$, $u(x_i) \trianglerighteq u(y)$. We show that $u(x) \trianglerighteq u(y)$. First, we define $x^*, x_* \in \Delta(X)$ such that for all $i \in \mathbb{N}$, $u(x^*) \geq u(x_i) \geq u(x_*)$. We only define x^* since x_* follows symmetrically. If $\sup u(\Delta(X)) \in u(\Delta(X))$, then define $x^* = \sup u(\Delta(X))$. So assume that $\sup u(\Delta(X)) \notin u(\Delta(X))$. Since $u(\Delta(X))$ is an interval, there exists $\varepsilon > 0$ such that $u(x) + \varepsilon \in u(\Delta(X))$. Since $u(x^i) \to u(x)$ as $i \to \infty$, there exists $i^\varepsilon \in \mathbb{N}$ such that for all $i > i^\varepsilon$, $u(x) + \varepsilon > u(x_i)$. Let

$$u^* = \max\{u(x^i) | 1 \leq i \leq i^\varepsilon\} \cup \{u(x) + \varepsilon\}.$$

Since u^* is a maximum over a finite set, it exists and so there exists $x^* \in \Delta(X)$ such that $u^* = u(x^*)$.

So there exist $x^*, x_* \in \Delta(X)$ such that for all $i \in \mathbb{N}$, $u(x^*) \geq u(x_i) \geq u(x_*)$ and since $u(\Delta(X))$ is not a singleton, $u(x^*) > u(x_*)$. Thus for each $i \in \mathbb{N}$, there exists a unique $\alpha^i \in [0, 1]$ such that $\alpha^i u(x^*) + (1 - \alpha^i)u(x_*) = u(x_i)$. Now since $u(x_i) \to u(x)$ as $i \to \infty$ there exists $\alpha \in [0, 1]$ such that $\alpha^i \to \alpha$ as $i \to \infty$ and $\alpha u(x^*) + (1 - \alpha)u(x_*) = u(x)$. Now we have for all $i \in \mathbb{N}$, $u(x_i) \trianglerighteq u(y)$ and so $x_i \succsim y$. By above, since $\trianglerighteq$ is well-defined and u is affine, we have for all $i \in \mathbb{N}$, $x_i \sim \alpha^i x^* + (1 - \alpha^i)x_*$. Thus by Axiom 1.2, $\alpha x^* + (1 - \alpha)x_* \succsim y$. And so $\alpha u(x^*) + (1 - \alpha)u(x_*) = u(x) \trianglerighteq u(y)$. This shows the claim.

Now $u(\Delta(X))$ is connected and separable. By Debreu (1954), there exists a continuous function $H : u(\Delta(X)) \to \mathbb{R}$ such that for all $x, y \in \Delta(X)$,

$$H(u(x)) \geq H(u(y)) \iff u(x) \trianglerighteq u(y) \iff x \succsim y.$$

Since u is affine, we have for all (simple) lotteries $P \in \Delta(X)$,

$$H\big(u(P)\big) = H\Big(\sum_{x \in X} P(x)u(x)\Big).$$

Finally, we show that H is weakly increasing. We have for all $x, y \in \Delta(X)$, if $H(u(x)) > H(u(y))$, then $x \succ y$ and since u is a weak representation $u(x) > u(y)$. This shows that H is weakly increasing.

Next, we show the necessity of the axioms. Assume that there exist affine $u : \Delta(X) \to \mathbb{R}$ and continuous and weakly increasing $H : u(\Delta(X)) \to \mathbb{R}$ such that for all $x, y \in \Delta(X)$,

$$x \succsim y \iff H(u(x)) \geq H(u(y)).$$

First, since $\succsim$ has a representation it is complete and transitive. Second, since u is affine and H is continuous, $\succsim$ satisfies mixture continuity, Axiom 1.2.

Third, we show Weak Independence. Let $x, y, z \in \Delta(X)$, $\alpha \in (0, 1)$ and $H(u(x)) > H(u(y))$. Since H is weakly increasing, $u(x) > u(y)$. Since u is affine, $u(\alpha x + (1-\alpha)z) > u(\alpha y + (1-\alpha)z)$. Since H is weakly increasing,

$$H(u(\alpha x + (1-\alpha)z)) \geq H(u(\alpha y + (1-\alpha)z))$$

that shows the claim.

Next we move on to the uniqueness. By Proposition A.12, there exist $a > 0, b \in \mathbb{R}$ such that $\widetilde{u} = au + b$. Now $H \circ u$ and $\widetilde{H} \circ \widetilde{u}$ are representations for $\succsim$. By the ordinal uniqueness of the representation, there exists a strictly increasing function $\varphi : H(u(\Delta(X))) \to \mathbb{R}$ such that for all $x \in \Delta(X)$,

$$\varphi(H(u(x))) = \widetilde{H}(\widetilde{u}(x)).$$

Thus for all $v \in u(\Delta(X))$, we have $av + b \in \widetilde{u}(\Delta(X))$ and by the uniqueness of u,

$$\widetilde{H}(av + b) = \varphi(H(v))$$

□

A.3 Existence of State Dependent Dual-Self Expected Utility and Dual-Self Variational Expected Utility

For acts $f, g \in H$ and $\alpha \in [0,1]$, denote $f \alpha g \coloneqq \alpha f + (1-\alpha)g$.

For $x \in \mathbb{R}$, $\bar{x}_S \in \mathbb{R}^S$ denotes the constant vector of x. Denote $h^* = 1/2 f^* + 1/2 g^*$ for equally crisp acts f^*, g^* from Axiom 1.5.

We start off with a simple observation.

Lemma A.13 Let $\succsim$ satisfy transitivity. If $f, g \in H$ s.t. for all $s \in S$, $f_s \succsim_s g_s$, then $f \succsim g$.

Proof. Since S is finite we can enumerate $S = \{s_1, \ldots, s_{|S|}\}$. Now we have by transitivity of $\succsim$ and the definition of $\succsim_s$,

$$f = (f_{s_1}, f_{s_2}, \ldots, f_{s_{|S|}}) \succsim (g_{s_1}, f_{s_2}, \ldots, f_{s_{|S|}}) \succsim \cdots$$
$$\succsim (g_{s_1}, g_{s_2}, \ldots, g_{s_{|S|-1}}, f_{s_{|S|}}) \succsim (g_{s_1}, g_{s_2}, \ldots, g_{s_{|S|}}) \equiv g.$$

Thus by transitivity, $f \succsim g$. □

A.3.1 Statewise Weak Affine Representations

Next we show that for each $s \in S$, $\succsim_s$ has a weak affine representation as defined in Appendix A.2. We show that $\succsim_s$ is weak order that is mixture continuous and satisfy Weak Independence by assumption. First, we show weak order part.

Lemma A.14 Let $s \in S$. If $\succsim$ satisfies Axioms 1-3, then $\succsim_s$ satisfies

1. completeness, i.e., for all $x_s, y_s \in \Delta(X_s)$, $x_s \succsim_s y_s$ or $y_s \succsim_s x_s$,
2. transitivity, i.e., for all $x_s, y_s, z_s \in \Delta(X_s)$, if $x_s \succsim_s y_s$ and $y_s \succsim_s z_s$, then $x_s \succsim_s z_s$, and

Proof. By the definition of $\succsim_s$, for all $x_s, y_s \in \Delta(X_s)$,

$$x_s \succsim_s y_s \iff \forall f \in H, (x_s, f_{-s}) \succsim (y_s, f_{-s}) \tag{A.6}$$

and by Axiom 1.3,

$$x_s \succ_s y_s \iff \exists f \in H, (x_s, f_{-s}) \succ (y_s, f_{-s}). \tag{A.7}$$

Complete: Let $x_s, y_s \in \Delta(X_s)$. If there exists $f \in H$ such that $(x_s, f_{-s}) \succ (y_s, f_{-s})$, then by (A.7), $x_s \succ_s y_s$ and so especially by the definition of $\succ_s$, $x_s \succsim_s y_s$. If the above $f \in H$ does not exist, then by the completeness of $\succsim$, for all $f \in H$, $(y_s, f_{-s}) \succsim (x_s, f_{-s})$ and so by (A.6) $y_s \succsim_s x_s$.

Transitive: Let $x_s, y_s, z_s \in \Delta(X_s)$ such that $x_s \succsim_s y_s \succsim_s z_s$, then $\forall\, f \in H$, $(x_s, f_{-s}) \succsim (y_s, f_{-s}) \succsim (z_s, f_{-s})$ and so by Axiom 1.1, $\forall\, f \in H$, $(x_s, f_{-s}) \succsim (z_s, f_{-s})$. Thus by (A.6) $x_s \succsim_s z_s$.

□

Next, we show mixture continuity for $\succsim_s$.

Lemma A.15 Let $s \in S$. If $\succsim$ satisfies Axioms 1-3, then for all $x_s, y_s, z_s \in \Delta(X_s)$, the sets $\{\alpha \in [0,1] | \alpha x_s + (1-\alpha) y_s \succsim_s z_s\}$ and $\{\alpha \in [0,1] | z_s \succsim_s \alpha x_s + (1-\alpha) y_s\}$ are closed in $\mathbb{R}$.

Proof. Let $s \in S$ and $x_s, y_s, z_s \in \Delta(X_s)$. We show that the sets $\{\alpha \in [0,1] | \alpha x_s + (1-\alpha) y_s \succ_s z_s\}$ and $\{\alpha \in [0,1] | z_s \succ_s \alpha x_s + (1-\alpha) y_s\}$ are open in $\mathbb{R}$ which shows the claim since by Lemma A.14, $\succsim_s$ is complete. We show that $\{\alpha \in [0,1] | \alpha x_s + (1-\alpha) y_s \succ_s z_s\}$ is open since $\{\alpha \in [0,1] | z_s \succ_s \alpha x_s + (1-\alpha) y_s\}$ is open by a symmetric argument. Let $\alpha^0 \in \{\alpha \in [0,1] | \alpha x_s + (1-\alpha) y_s \succ_s z_s\}$. Thus $\alpha^0 x_s + (1-\alpha^0) y_s \succ_s z_s$. By the definition of $\succ_s$, there exists $f \in H$ such that $(\alpha^0 x_s + (1-\alpha^0) y_s, f_{-s}) \succ (z_s, f_{-s})$. Thus $\alpha^0 \in \left\{\alpha \in [0,1] \middle| (\alpha^0 x_s + (1-\alpha^0) y_s, f_{-s}) \succ (z_s, f_{-s})\right\}$. By Axiom 1.2, there exists a neighborhood V of α^0 such that for all $\alpha \in V \cap [0,1]$, $(\alpha x_s + (1-\alpha) y_s, f_{-s}) \succ (z_s, f_{-s})$. Thus for all $\alpha \in V \cap [0,1]$, $\alpha x_s + (1-\alpha) y_s \succ_s z_s$. Hence $V \cap [0,1] \subseteq \{\alpha \in [0,1] | z_s \succ_s \alpha x_s + (1-\alpha) y_s\}$ and since V is a neighborhood of α^0 this shows the claim. □

Next, we put the above together for weak affine representation.

Lemma A.16 Let $s \in S$. If $\succsim$ satisfies Axioms 1-1.4, then there exists an affine $u_s : \Delta(X_s) \to \mathbb{R}$ such that u_s is a weak affine representation for $\succsim_s$.

Proof. By Lemmas A.14 and A.15, $\succsim_s$ is a nontrivial weak order and satisfies continuity. By Axiom 1.4, $\succsim$ satisfies Weak Independence. By Proposition A.10, there exists an affine $u_s : \Delta(X_s) \to \mathbb{R}$ such that for all $x_s, y_s \in \Delta(X_s)$,

$$x_s \succ_s y_s \Rightarrow u_s(x_s) > u_s(y_s).$$

□

The next result shows statewise monotonicity results with the weak affine representation.

Lemma A.17 Assume that for each $s \in S$, $\succsim_s$ has a weak affine representation with u_s. Let $s \in S$. Then

1. For all $x_s, y_s \in \Delta(X_s)$, $u_s(x_s) \geq u_s(y_s) \Longrightarrow x_s \succsim_s y_s$.
2. For all $f, g \in H$, if for all $s \in S^{\mathcal{P}}$, $u_s(f_s) \geq u_s(g_s)$, then $f \succsim g$.

Proof. The first one follows from the negation of the definition of weak representation. The second one follows from the first one and Lemma A.13. □

A.3.2 Preferences Monotonic Between Equally Crisp Acts

Next we show that $\succsim$ is monotonic when moving from one crisp act to another one.

Lemma A.18 Let $\succsim$ satisfy Axioms 1-5. Let for all $s \in S^{\mathcal{P}}$, $u_s : \Delta(X_s) \to \mathbb{R}$ be an affine weak representation for $\succsim_s$. Let f^*, g^* be equally crisp acts such that for all $s \in S^{\mathcal{P}}$,

$$u_s(f_s^*) > u_s(g_s^*). \tag{A.8}$$

Then for all $\alpha, \alpha' \in [0,1]$,

$$\alpha \geq \alpha' \iff f^* \alpha g^* \succsim f^* \alpha' g^*.$$

Proof. Now $u(f^{*1/2}g^*) \in \operatorname{int} u(H)$. Assume, per contra, that there exist $1 > \alpha^* > \alpha_* > 0$ such that

$$f^* \alpha g^* \sim f^* \alpha' g^* \tag{A.9}$$

By the nontriviality, there exists $f^0 \not\sim f^* \alpha g^*$. W.l.o.g. assume that $f^* \alpha g^* \succ f^0$. By continuity, there exists $\alpha^1 \in (0,1)$ such that $f^1 = \alpha^1 f^0 + (1-\alpha^1)(f^{*1/2}g^*)$ and $f^* \alpha g^* \succ f^1$. Now $u(f^1) \in \operatorname{int} u(H)$. Let

$$\alpha^f = \inf\{\alpha \in [0,1] | \alpha f^1 + (1-\alpha) f^* \alpha^* g^* \succsim f^* \alpha^* g^*\}. \tag{A.10}$$

By the continuity and the assumption, $\alpha^* \in [0,1]$ and $\alpha^f f^0 + (1-\alpha^f) f^* \alpha^* g^* \sim f^* \alpha^* g^*$. Now by Corollary A.42, $f^* \alpha^* g^*$ and $f^* \alpha_* g^*$ are equally crisp. Thus

$$\alpha^f f^0 + (1-\alpha^f) f^* \alpha^* g^* \sim \alpha^f f^* \alpha^* g^* + (1-\alpha^f) f^* \alpha^* g^*$$

and so

$$\alpha^f f^0 + (1-\alpha^f) f^* \alpha_* g^* \sim \alpha^f f^* \alpha^* g^* + (1-\alpha^f) f^* \alpha_* g^* = f^*(\alpha^f \alpha^* + (1-\alpha^f)\alpha_*) g^*. \tag{A.11}$$

Let $0 < \varepsilon < 1 - \alpha^f$ and

$$0 < \varepsilon < \frac{(1-\alpha^f)(\alpha^* - \alpha_*) \min_{s \in S^{\mathcal{P}}} |u_s(f_s^*) - u(g_s^*)|}{\|u(f^0) - u(f^* \alpha^* g^*)\|_\infty}. \tag{A.12}$$

Then we have for all $s \in S^{\mathcal{P}}$,

$$
\begin{aligned}
& u_s((\alpha^f+\varepsilon)f_s^0 + (1-\alpha^f-\varepsilon)f_s^*\alpha^*g_s^*) \qquad (A.13)\\
&= \alpha^f u_s(f_s^0) + (1-\alpha^f)\big(u_s(g_s^*) + \alpha^*(u_s(f_s^*) - u_s(g_s^*))\big) + \varepsilon(u_s(f_s^0) - u_s(f_s^*\alpha^*g_s^*))\\
&\geq \alpha^f u_s(f_s^0) + (1-\alpha^f)u_s(g_s^*) + (1-\alpha^f)\alpha^*(u_s(f_s^*) - u_s(g_s^*)) - \varepsilon\Big\|u(f^0) - u(f^*\alpha^*g^*)\Big\|\\
&\overset{(A.12)}{>} \alpha^f u_s(f_s^0) + (1-\alpha^f)u_s(g_s^*) + (1-\alpha^f)\alpha^*(u_s(f_s^*) - u_s(g_s^*))\\
&\qquad\qquad -(1-\alpha^f)(\alpha^*-\alpha_*)\min_{\widetilde{s}\in S^{\mathcal{P}}}\Big|u_{\widetilde{s}}(f_{\widetilde{s}}^*) - u(g_{\widetilde{s}}^*)\Big|\\
&\geq \alpha^f u_s(f_s^0) + (1-\alpha^f)u_s(g_s^*) + (1-\alpha^f)\alpha^*(u_s(f_s^*) - u_s(g_s^*)) - (1-\alpha^f)(\alpha^*-\alpha_*)(u_s(f_s^*) - u(g_s^*))\\
&= \alpha^f u_s(f_s^0) + (1-\alpha^f)u_s(g_s^*) + (1-\alpha^f)\alpha_*(u_s(f_s^*) - u_s(g_s^*))\\
&= u_s(\alpha^f f_s^0 + (1-\alpha^f)f_s^*\alpha_*g_s^*).
\end{aligned}
$$

Thus by Lemma A.17,

$$
\begin{aligned}
(\alpha^f+\varepsilon)f^0 + (1-\alpha^f-\varepsilon)f^*\alpha^*g^* &\overset{(A.13)}{\succsim} \alpha^f f^0 + (1-\alpha^f)f^*\alpha_*g^*\\
&\overset{(A.11)}{\sim} f^*(\alpha^f\alpha^* + (1-\alpha^f)\alpha_*)g^* \overset{(A.8)}{\succsim} f^*\alpha_*g^* \overset{(A.9)}{\sim} f^*\alpha^*g^*.
\end{aligned}
$$

Thus for all $0<\varepsilon<1-\alpha^f$ that satisfy (A.12), $(\alpha^f+\varepsilon)f^0+(1-\alpha^f-\varepsilon)f^*\alpha^*g^* \succsim f^*\alpha^*g^*$. However, this contradicts the definition of α^f in (A.10). Thus if $1>\alpha^*>\alpha_*>0$, then by (A.8) and Lemma A.17 $f^*\alpha^*g^* \succ f^*\alpha_*g^*$. Finally, if $1\geq\alpha^*>\alpha_*\geq 0$, then by above, (A.8) and Lemma A.17,

$$
f^*\alpha^*g^* \succsim f^*(3/4\alpha^* + 1/4\alpha^*)g^* \succsim f^*(1/4\alpha^* + 3/4\alpha^*)g^* \succsim f^*\alpha_*g^*.
$$

□

The next result generalizes the previous monotonicity when the equally crisp acts are mixed with another act.

Lemma A.19 Let $\succsim$ satisfy Axioms 1-5 and let for all $s \in S^{\mathcal{P}}$, $u_s : \Delta(X_S) \to \mathbb{R}$ be a weak affine representation for $\succsim_s$. Let f^*, g^* be equally crisp acts such that for all $s \in S^{\mathcal{P}}$, $u_s(f_s^*) > u_s(g_s^*)$. Then for all $f \in H, \beta \in [0,1)$,

$$
\alpha \geq \alpha' \iff \beta f + (1-\beta)(f^*\alpha g^*) \succsim \beta f + (1-\beta)(f^*\alpha' g^*).
$$

Proof. Let $f \in H$ and $\beta \in (0,1)$. By Corollary A.43, $\beta f + (1-\beta)f^*, \beta f + (1-\beta)g^*$ are equally crisp acts and for all $s \in S^{\mathcal{P}}$, by the affinity of u_s, $u_s(\beta f + (1-\beta)f^*) > u_s(\beta f + (1-\beta)g^*)$. Thus the claim follows from Lemma A.18. □

A.3.3 Transfinite Induction

The next steps show the existence of C-additive and monotonic representation. First, we define a representation between the two equally crisp acts. Second, we extend this representation using the equally crisp acts in a manner that maintains C-additivity. This is done using transfinite induction in multiple steps.

The next lemma shows, that extending the starting representation using equally crisp acts and C-additivity gives a unique way to extend the representation.

Lemma A.20 Assume that $S = S^{\mathcal{P}}$. Let $\succsim$ satisfy Axioms 1-1.4,1.6, and for all $s \in S^{\mathcal{P}}$, let $u_s : \Delta(X_s) \to \mathbb{R}$ be a weak affine representation for $\succsim_s$ such that

$$u_s(f_s^*) - u_s(g_s^*) = 1. \tag{A.14}$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1, a_2 \in \mathbb{R} \cup \{-\infty, \infty\}, a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \operatorname{int} u(H)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$, if there exist $f', g' \in H$ such that $u(f'), u(g') \in \operatorname{dom} \widetilde{I}$ and $f' \succsim f \succsim g'$, then $u(f) \in \operatorname{dom} \widetilde{I}$.
3. For all $\alpha \in (0,1)$, $u\big(\alpha f^* + (1-\alpha) g^*\big) \in \operatorname{dom} \widetilde{I}$ and
$$\widetilde{I} \circ u\big(\alpha f^* + (1-\alpha) g^*\big) = \alpha.$$
4. If $\varphi \in \operatorname{dom} \widetilde{I}$ and $\alpha^1 > 0$ are such that $\varphi + \alpha^1 \bar{1} \in \operatorname{dom} \widetilde{I}$, then $\widetilde{I}(\varphi + \alpha^1 \bar{1}) = \widetilde{I}(\varphi) + \alpha^1$.
5. For all $f, g \in H$ with $u(f), u(g) \in \operatorname{dom} \widetilde{I}$, $f \succsim g$ iff $\widetilde{I}(u(f)) \geq \widetilde{I}(u(g))$.

Let $A_1, A_2 \in \mathcal{A}$, $A_1 \subseteq A_2$ and I_1, I_2 be the associated functions. Then $\operatorname{dom} I_1 \subseteq \operatorname{dom} I_2$ and for all $P \in \operatorname{dom} I_1$, $I_1(P) = I_2(P)$.

Proof. Denote $f^* \frac{1}{2} g^* = \frac{1}{2} f^* + \frac{1}{2} g^*$. We will show the claim for all $\varphi \in \operatorname{dom} I_1$ such that $I_1(\varphi) \geq \frac{1}{2}$. The other case follows symmetrically. Denote

$$m \coloneqq \inf \{a \in [1/2, \infty) | \exists \psi \in \operatorname{dom} I_1, I_1(\psi) = a, [\psi \notin \operatorname{dom} I_2 \text{ or } I_1(\psi) \neq I_2(\psi)]\}. \tag{A.15}$$

We will show that $m = \infty$. We will first show that $m > 0$. Let $\varphi \in \operatorname{dom} I_1$, be such that $\frac{1}{2} \leq I_1(\varphi) < 1$. Let $f \in H$ be such that $u(f) = \varphi$ and $\alpha^{\varphi}\big(I_1(\varphi), 1\big)$. Now we have by Condition 3,

$$u\big(f^* \frac{1}{2} g^*\big), u\big(\alpha^{\varphi} f^* + (1-\alpha^{\varphi}) g^*\big), u\Big(I_1(\varphi) f^* + \big(1 - I_1(\varphi)\big) g^*\Big) \in \operatorname{dom} I_1, \operatorname{dom} I_2.$$

By Conditions 3 and 5 for I_1,

$$\alpha^{\varphi} f^* + (1-\alpha^{\varphi}) g^* \succsim f \sim I_1(\varphi) f^* + \big(1 - I_1(\varphi)\big) g^* \succsim f^* \frac{1}{2} g^*.$$

By Condition 2 for I_2, $u(f) = \varphi \in \operatorname{dom} I_2$. And by Conditions 3 and 5 for I_2,

$$I_2(\varphi) = I_2(u(f)) = I_2 \circ u\Big(I_1(\varphi)f^* + \big(1 - I_1(\varphi)\big)g^*\Big) = I_1(\varphi).$$

Thus $m > \frac{1}{2}$.

Assume, per contra, $m < \infty$. By Condition 1, there exists $\varphi \in \operatorname{dom} I_1$ and $I_1 \circ u(\varphi) = m$. Since A_1 is open, there exists $\bar{f}^1 \in H$ such that $u(\bar{f}^1) \in \operatorname{dom} I_1$ and $I_1 \circ u(f^m) < I_1 \circ u(\bar{f}^1)$. Since $A_2 \supseteq A_1$, there exists $\bar{f}^2 \in H$ such that $u(\bar{f}^2) \in \operatorname{dom} I_2$ and $I_2 \circ u(\bar{f}^2) = I_1 \circ u(\bar{f}^1)$. Since $\varphi \in \operatorname{int} u(H)$, there exists $\alpha^* \in (0, 1)$ such that

$$\frac{1}{\alpha^*}\varphi - \frac{1 - \alpha^*}{\alpha^*}u(f^{*1/2}g^*) \in \operatorname{int} u(H).$$

Thus there exists $f^m \in H$ such that

$$u(f^m) = \frac{1}{\alpha^*}\varphi - \frac{1 - \alpha^*}{\alpha^*}u(f^{*1/2}g^*).$$

Now

$$\alpha^* u(f^m) + (1 - \alpha^*)u(f^{*1/2}g^*) = \varphi.$$

By Condition 5 for I_1,

$$\bar{f}^1 \succ \alpha^* u(f^m) + (1 - \alpha^*)u(f^{*1/2}g^*) \succ f^{*1/2}g^*.$$

By Axiom 1.2, there exists $\alpha^\dagger \in (0, 1/2)$ such that for all $\alpha \in (1/2 - \alpha^\dagger, 1/2 + \alpha^\dagger)$,

$$\bar{f}^1 \succ \alpha^* f^m + (1 - \alpha^*)f^* \alpha g^* \succ f^{*1/2}g^*.$$

By Condition 2, for all $\alpha \in (1/2 - \alpha^\dagger, 1/2 + \alpha^\dagger)$,

$$\alpha^* u(f^m) + (1 - \alpha^*)u(f^* \alpha g^*) \in \operatorname{dom} I_1.$$

By (A.14) and Condition 4, for all $\alpha \in (1/2 - \alpha^\dagger, 1/2 + \alpha^\dagger)$,

$$I_1\Big(\alpha^* u(f^m) + (1 - \alpha^*)u(f^* \alpha g^*)\Big) = I_1\Big(\alpha^* u(f^m) + (1 - \alpha^*)u(f^{*1/2}g^*)\Big) + (1 - \alpha^*)(\alpha - 1/2).$$

Thus by (A.15) for all $\alpha \in (1/2 - \alpha^\dagger, 1/2)$,

$$\alpha^* u(f^m) + (1 - \alpha^*)u(f^* \alpha g^*) \operatorname{dom} I_2$$

and

$$I_2(\alpha^* u(f^m) + (1 - \alpha^*)u(f^* \alpha g^*)) = I_1(\alpha^* u(f^m) + (1 - \alpha^*)u(f^* \alpha g^*)) < I_1(\varphi) < I_2 \circ u(\bar{f}^2).$$

By Condition 5 for I_2, for all $\alpha \in (1/2 - \alpha^\dagger, 1/2)$,

$$\bar{f}^2 \succ \alpha^* f^m + (1 - \alpha^*)(f^* \alpha g^*)$$

and so by Axiom 1.2,

$$\bar{f}^2 \succsim \alpha^* f^m + (1-\alpha^*)(f^{*1/2}g^*).$$

By Condition 2,

$$\alpha^* u(f^m) + (1-\alpha^*)u(f^{*1/2}g^*) \,\mathrm{dom}\, I_2$$

and by Condition 4,

$$I_2\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^{*1/2}g^*)\big) = I_1\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^{*1/2}g^*)\big). \tag{A.16}$$

By Condition 5, $\bar{f}^2 \succ \alpha^* f^m + (1-\alpha^*)(f^{*1/2}g^*)$. By Axiom 1.2, there exists $\alpha^\ddagger \in (0, \alpha^\dagger)$ such that for all $\alpha \in [1/2, 1/2 + \alpha^\ddagger]$,

$$\bar{f}^2 \succ \alpha^* f^m + (1-\alpha^*) f^* \alpha g^*.$$

By Condition 2, for all $\alpha \in [1/2, 1/2 + \alpha^\ddagger]$,

$$\alpha^* u(f^m) + (1-\alpha^*)u(f^*\alpha g^*) \,\mathrm{dom}\, I_1 \cap \mathrm{dom}\, I_2. \tag{A.17}$$

By Condition 4 and (A.16), for all $\alpha \in [1/2, 1/2 + \alpha^\ddagger]$,

$$I_1\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^*\alpha g^*)\big) = I_2\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^*\alpha g^*)\big). \tag{A.18}$$

Next, let $h \in H$ and $u(h) \in \mathrm{dom}\, I_1$ be such that

$$I_1\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^*\alpha^\ddagger g^*)\big) \geq I_1 \circ u(h) \geq I_1\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^{*1/2}g^*)\big).$$

Let We show that $u(h) \in \mathrm{dom}\, I_2$ and $I_2 \circ u(h) = I_1 \circ u(h)$. By Condition 5 for I_1,

$$\alpha^* f^m + (1-\alpha^*)(f^*\alpha^\ddagger g^*) \precsim h \precsim \alpha^* f^m + (1-\alpha^*)(f^{*1/2}g^*).$$

By Condition 2 and (A.17), $u(h) \in \mathrm{dom}\, I_2$. By Axiom 1.2, there exists $\alpha^h \in [1/2, \alpha^\ddagger]$ such that

$$h \sim \alpha^* f^m + (1-\alpha^*)(f^*\alpha^h g^*).$$

By Condition 5 and (A.17,A.18),

$$I_1(h) = I_1\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^*\alpha^h g^*)\big) = I_2\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^*\alpha^h g^*)\big) = I_2(h).$$

Thus for all $\psi \in \mathrm{dom}\, I_1$ such that $I_1\big(\alpha^* u(f^m) + (1-\alpha^*)u(f^*\alpha^\ddagger g^*)\big) \geq I_1(\psi) \geq I_1(\varphi)$, we have $\psi \in \mathrm{dom}\, I_2$ and $I_2(\psi) = I_1(\psi)$. This contradicts the choice of φ in (A.15). Thus $m = \infty$. □

The next lemma shows that if we measure the utility of other acts by the two equally crisp acts, we maintain C-additivity.

Lemma A.21 Assume that $S = S^{\mathcal{P}}$. Let $\succsim$ satisfy Axioms 1-1.4,1.6, and for all $s \in S^{\mathcal{P}}$, let $u_s : \Delta(X_s) \to \mathbb{R}$ be a weak affine representation for $\succsim_s$ such that

$$u_s(f_s^*) - u_s(g_s^*) = 1.$$

Let $\beta^* \in [0,1)$, $f^\dagger \in H, c \in \mathbb{R}$ and denote

$$B = \{\varphi \in \operatorname{int} u(H) \big| \, \exists\, f \in H, \alpha \in (0,1), u(f) = \varphi, f \sim \beta^* f^\dagger + (1-\beta^*)(f^* \alpha g^*)\}.$$

Define $I : B \to \mathbb{R}$ by the following. For each $\varphi \in B$, let $f^\varphi, \alpha^\varphi \in (0,1) \in H$ be such that $u(f) = \varphi$ and $f \sim \beta^* f^\dagger + (1-\beta^*)(f^* \alpha g^*)$, then define

$$I(\varphi) = (1-\beta^*)\alpha + c. \tag{A.19}$$

If $\varphi \in B$ and $\alpha^1 > 0$ are such that $\varphi + \alpha^1 \bar{1} \in \operatorname{dom} B$, then $I(\varphi + \alpha^1 \bar{1}) = \widetilde{I}(\varphi) + \alpha^1$.

Proof. First, for all $f, g \in H$ such that $u(f), u(g) \in B$ and $f \sim g$, we have by Lemma A.19,

$$I\big(u(f)\big) = I\big(u(g)\big). \tag{A.20}$$

Let $\varphi \in B$ and $\alpha^0 > 0$ such that $\varphi + \alpha^0 \bar{1} \in B$. Let $\varphi^{-1}, f^0 \in H$ be such that $\varphi = u(\varphi^{-1}), \varphi + \alpha^0 \bar{1} = u(f^0)$. Consider the mapping $J : [0, \alpha^0] \to \mathbb{R}$ defined by for all $\alpha \in [0, \alpha^0]$, $J(\alpha) = I(\varphi + \alpha \bar{1})$. We show that this is an affine function with derivative 1. By Lemma A.19 and the definition of B,

$$\beta^* f^\dagger + (1-\beta^*) f^* \succ f^0, \varphi^{-1} \succ \beta^* f^\dagger + (1-\beta^*) g^*. \tag{A.21}$$

Let $\alpha^1 \in [0, \alpha^0]$. Since $\varphi, \varphi + \alpha^1 \bar{0} \in \operatorname{int} u(H)$, $\varphi + \alpha^1 \bar{1} \in \operatorname{int} u(H)$. Let $f^1 \in H$ be such that $u(f^1) = \varphi + \alpha^1 \bar{1}$. By Lemma A.17,

$$\beta^* f^\dagger + (1-\beta^*) f^* \succ^0 \gtrsim f^1 \gtrsim \varphi^{-1} \succ \beta^* f^\dagger + (1-\beta^*) g^*.$$

Thus there exists $\alpha^2 \in (0,1)$ such that $\beta^* f^\dagger + (1-\beta^*)(f^* \alpha^2 g^*) \sim f^1$. Since $u(\varphi + \alpha^1 \bar{1}) \in \operatorname{int} u(H)$, there exists $\varepsilon^\varphi > 0$ such that $B_\infty(\varphi + \alpha^1 \bar{1}, \varepsilon^\varphi) \subseteq u(H)$.

Let $\beta^\varphi \in (\beta^*, 1)$, which exists since $\beta^* < 1$, be such that

$$\frac{1-\beta^\varphi}{\beta^\varphi}\Big(\big\|\varphi + \alpha^1 \bar{1}\big\|_\infty + \big\|u(f^* \alpha^2 g^*)\big\|_\infty\Big) < \varepsilon^\varphi.$$

Thus by the choice of β^φ, there exists $f^\varphi \in H$ such that

$$u(f^\varphi) = \varphi + \alpha^1 \bar{1} + \frac{1-\beta^\varphi}{\beta^\varphi}(\varphi + \alpha^1 \bar{1}) - \frac{1-\beta^\varphi}{\beta^\varphi} u(f^* \alpha^2 g^*).$$

Especially

$$\beta^\varphi u(f^\varphi) + (1-\beta^\varphi) u(f^* \alpha^2 g^*) = \varphi + \alpha^1 \bar{1} = u(f^1).$$

Now by Lemma A.17 and since $\beta^\varphi > \beta^*$,

$$\begin{aligned}&\beta^\varphi f^\varphi + (1-\beta^\varphi)(f^*\alpha^2 g^*) \sim f^1 \sim \beta^* f^\dagger + (1-\beta^*)(f^*\alpha^2 g^*)\\ &\equiv \beta^\varphi\Big(\frac{\beta^*}{\beta^\varphi} f^\dagger + \frac{\beta^\varphi - \beta^*}{\beta^\varphi}(f^*\alpha^2 g^*)\Big) + (1-\beta^\varphi)(f^*\alpha^2 g^*).\end{aligned}$$

Since f^*, g^* are equally crisp, we have for all $\gamma \in [0,1]$, by Corollary A.41,

$$\beta^\varphi f^\varphi + (1-\beta^\varphi)(f^*\gamma g^*) \sim \beta^\varphi\Big(\frac{\beta^*}{\beta^\varphi} f^\dagger + \frac{\beta^\varphi - \beta^*}{\beta^\varphi}(f^*\alpha^2 g^*)\Big) + (1-\beta^\varphi)(f^*\gamma g^*). \tag{A.22}$$

By the linearity of u, we have

$$\begin{aligned}&u\big(\beta^\varphi f^\varphi + (1-\beta^\varphi)(f^*\gamma g^*)\big) \qquad\qquad (A.23)\\ &= u(f^1) + (1-\beta^\varphi)\big(u(f^*\gamma g^*) - u(f^*\alpha^2 g^*)\big) = \varphi + \big(\alpha^1 + (1-\beta^\varphi)(\gamma - \alpha^2)\big)\bar{1}\end{aligned}$$

and since $1-\beta^* > 1-\beta^\varphi$

$$\begin{aligned}&\beta^\varphi u\Big(\frac{\beta^*}{\beta^\varphi} f^\dagger + \frac{\beta^\varphi - \beta^*}{\beta^\varphi}(f^*\alpha^2 g^*)\Big) + (1-\beta^\varphi)u(f^*\gamma g^*) \qquad\qquad (A.24)\\ &= \beta^* u(f^\dagger) + \big((\beta^\varphi - \beta^*)\alpha^2 + (1-\beta^\varphi)\gamma\big)u(f^*) + \big((\beta^\varphi - \beta^*)(1-\alpha^2) + (1-\beta^\varphi)(1-\gamma)\big)u(g^*)\\ &= \beta^* u(f^\dagger) + (1-\beta^*)u\Big(f^*\big(\alpha^2 + \frac{1-\beta^\varphi}{1-\beta^*}(\gamma - \alpha^2)\big)g^*\Big).\end{aligned}$$

We have for all $\gamma \in [0,1]$,

$$\begin{aligned}&J\big(\alpha^1 + (1-\beta^\varphi)(\gamma - \alpha^2)\big)\\ &= I\Big(\varphi + \big(\alpha^1 + (1-\beta^\varphi)(\gamma - \alpha^2)\big)\bar{1}\Big) \overset{(A.23)}{=} I\Big(u\big(\beta^\varphi f^\varphi + (1-\beta^\varphi)(f^*\gamma g^*)\big)\Big)\\ &\overset{(A.20, A.22)}{=} I\Big(\beta^\varphi u\Big(\frac{\beta^*}{\beta^\varphi} f^\dagger + \frac{\beta^\varphi - \beta^*}{\beta^\varphi}(f^*\alpha^2 g^*)\Big) + (1-\beta^\varphi)u(f^*\gamma g^*)\Big)\\ &\overset{(A.24)}{=} I\Big(\beta^* u(f^\dagger) + (1-\beta^*)u\Big(f^*\big(\alpha^2 + \frac{1-\beta^\varphi}{1-\beta^*}(\gamma - \alpha^2)\big)g^*\Big)\Big)\\ &\overset{(A.19)}{=} (1-\beta^*)\Big(\alpha^2 + \frac{1-\beta^\varphi}{1-\beta^*}(\gamma - \alpha^2)\Big) + c = (1-\beta^*)\alpha^2 + (1-\beta^\varphi)(\gamma - \alpha^2) + c.\end{aligned}$$

Thus for all $x \in \big(\alpha^1 - (1-\beta^\varphi)\alpha^2, \alpha^1 + (1-\beta^\varphi)(1-\alpha^2)\big) \cap [0, \alpha^1]$, first $\frac{(x-\alpha^1)}{1-\beta^\varphi} + \alpha^2 \in (0,1)$ and so by above,

$$\begin{aligned}J(x) &= J\Big(\alpha^1 + (1-\beta^\varphi)\Big(\frac{(x-\alpha^1)}{1-\beta^\varphi} + \alpha^2 - \alpha^2\Big)\Big)\\ &= (1-\beta^*)\alpha^2 + (1-\beta^\varphi)\Big(\frac{(x-\alpha^1)}{1-\beta^\varphi} + \alpha^2 - \alpha^2\Big) + c = (1-\beta^*)\alpha^2 + x + c.\end{aligned}$$

So J is affine in the neighborhood $\big(\alpha^1 - (1-\beta^\varphi)\alpha^2, \alpha^1 + (1-\beta^\varphi)(1-\alpha^2)\big) \cap [0, \alpha^1]$ of α^1 and the derivative is 1. Since the point $\alpha^1 \in [0, \alpha^0]$ was arbitrary, the derivative of J is locally constant at 1 and J is affine locally at 0 and α^0. Since the interval $[0, \alpha^0]$ is connected, by (Viro et al., 2008, Problem 12.2x) the derivative of J is constant at 1 on the set $(0, \alpha^0)$. Thus J is affine function on

the set $[0, \alpha^0]$. Thus by the Fundamental theorem of calculus and since J is locally affine at 0 and α^0,

$$J(\alpha^0) - J(0) = J(\alpha^0) - \lim_{\alpha \nearrow \alpha^0} J(\alpha) - J(0) + \lim_{\alpha \searrow 0} J(\alpha) + \int_0^{\alpha^0} \nabla J(\alpha) d\alpha = \int_0^{\alpha^0} 1 d\alpha = \alpha^0.$$

Hence

$$I(\varphi + \alpha^0 \bar{1}) = J(\alpha^0) = J(0) + \alpha^0 = I(\varphi) + \alpha^0.$$

□

The next lemma is the first step to using Zorn's lemma for the extension and shows that the set of extensions is non-empty.

Lemma A.22 Assume that $S = S^{\mathcal{P}}$. Let $\succsim$ satisfy Axioms 1-1.4,1.6, and for all $s \in S^{\mathcal{P}}$, let $u_s : \Delta(X_s) \to \mathbb{R}$ be a weak affine representation for $\succsim_s$ such that

$$u_s(f_s^*) - u_s(g_s^*) = 1.$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1, a_2 \in \mathbb{R} \cup \{-\infty, \infty\}, a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \operatorname{int} u(H)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$, if there exist $f', g' \in H$ such that $u(f'), u(g') \in \operatorname{dom} \widetilde{I}$ and $f' \succsim f \succsim g'$, then $u(f) \in \operatorname{dom} \widetilde{I}$.
3. For all $\alpha \in (0,1)$, $u\big(\alpha f^* + (1-\alpha) g^*\big) \in \operatorname{dom} \widetilde{I}$ and
$$\widetilde{I} \circ u\big(\alpha f^* + (1-\alpha) g^*\big) = \alpha.$$
4. If $\varphi \in \operatorname{dom} \widetilde{I}$ and $\alpha^1 > 0$ are such that $\varphi + \alpha^1 \bar{1} \in \operatorname{dom} \widetilde{I}$, then $\widetilde{I}(\varphi + \alpha^1 \bar{1}) = \widetilde{I}(\varphi) + \alpha^1$.
5. For all $f, g \in H$ with $u(f), u(g) \in \operatorname{dom} \widetilde{I}$, $f \succsim g$ iff $\widetilde{I}(u(f)) \geq \widetilde{I}(u(g))$.

Then $\mathcal{A} \neq \varnothing$.

Proof. We show that $(0,1) \in \mathcal{A}$. Denote

$$B = \{\varphi \in \operatorname{int} u(H) | \exists f \in H, u(f) = \varphi, f^* \succ f \succ g^*\}.$$

Define $I^0 : B \to \mathbb{R}$ by the following. Let $\varphi \in B$ and $f \in H$ such that $u(f) = \varphi, f^* \succ f \succ g^*$. By Axiom 1.2, there exists $\alpha^f \in (0,1)$ such that $f^* \alpha^f + (1 - \alpha^f) g^* \sim f$. Define $I^0(\varphi) = \alpha^f$. By Lemmas A.17 and A.18, this is well-defined function. We show that $(0,1)$ and with the function I^0 satisfies Conditions 1-5.

Conditions 1-3 follow from the definition of I^0.

Condition 4: Follows from Lemma A.21.

Condition 5: Since for all $h \in \operatorname{dom} I^0$, $f^* \succ h \succ g^*$, there exist $\alpha^f, \alpha^g \in (0,1)$ such that $f \sim \alpha^f f^* + (1-\alpha^f)g^*$ and $g \sim \alpha^g f^* + (1-\alpha^g)g^*$. By the definition of I^0 and Lemma A.19, we have

$$I^0(u(f)) \geq I^0(u(g)) \Leftrightarrow \alpha^f \geq \alpha^g \Leftrightarrow f^*\alpha^f g^* \succsim f^*\alpha^g g^* \Leftrightarrow f \succsim g.$$

Thus $(0,1) \in \mathcal{A}$. □

The next lemma is the second step to using Zorn's lemma and shows that any chain of extensions has an upper bound.

Lemma A.23 Assume that $S = S^{\mathcal{P}}$. Let $\succsim$ satisfy Axioms 1-1.4,1.6, and for all $s \in S^{\mathcal{P}}$, let $u_s : \Delta(X_s) \to \mathbb{R}$ be a weak affine representation for $\succsim_s$ such that

$$u_s(f_s^*) - u_s(g_s^*) = 1.$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1, a_2 \in \mathbb{R} \cup \{-\infty, \infty\}, a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \operatorname{int} u(H)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$, if there exist $f', g' \in H$ such that $u(f'), u(g') \in \operatorname{dom} \widetilde{I}$ and $f' \succsim f \succsim g'$, then $u(f) \in \operatorname{dom} \widetilde{I}$.
3. For all $\alpha \in (0,1)$, $u\big(\alpha f^* + (1-\alpha)g^*\big) \in \operatorname{dom} \widetilde{I}$ and
$$\widetilde{I} \circ u\big(\alpha f^* + (1-\alpha)g^*\big) = \alpha.$$
4. If $\varphi \in \operatorname{dom} \widetilde{I}$ and $\alpha^1 > 0$ are such that $\varphi + \alpha^1 \bar{1} \in \operatorname{dom} \widetilde{I}$, then $\widetilde{I}(\varphi + \alpha^1 \bar{1}) = \widetilde{I}(\varphi) + \alpha^1$.
5. For all $f, g \in H$ with $u(f), u(g) \in \operatorname{dom} \widetilde{I}$, $f \succsim g$ iff $\widetilde{I}(u(f)) \geq \widetilde{I}(u(g))$.

Let $\mathcal{B} \subseteq \mathcal{A}$ be such that for each $B_1, B_2 \in \mathcal{B}$, $B_1 \subseteq B_2$ or $B_2 \subseteq B_1$. Then $\bigcup_{B \in \mathcal{B}} B \in \mathcal{A}$.

Proof. Denote $B^* \coloneqq \bigcup_{B \in \mathcal{B}} B$. For each $B \in \mathcal{B}$ denote by I^B the associated function. Since each $\mathcal{B}$ consists of nested open intervals, B^* is an open interval. Next, define the associated extension I^*. For all $\varphi \in \operatorname{int} u(H)$, if there exists $B \in \mathcal{B}$ such that $\varphi \in \operatorname{dom} I^B$, define $I^*(\varphi) = I^B(\varphi)$. Since $\mathcal{B}$ is a chain by Lemma A.20, I^* is well-defined. We will show that I^* and B^* satisfies Conditions 1-5.

Condition 1 follows from the definition of I^*.

Condition 2: Let $f \in H$ be such that $u(f) \in \operatorname{int} u(H)$ and $f', g' \in H$ such that $u(f'), u(g') \in \operatorname{dom} I^*$ and $f' \succsim f \succsim g'$. Now there exist B_1, B_2 such that $f' \in \operatorname{dom} I^{B_1}, g' \in \operatorname{dom} I^{B_2}$ and w.l.o.g. $B_1 \subseteq B_2$. By Lemma A.20, $f' \in \operatorname{dom} I^{B_2}$. By Condition 2 for B_2, $u(f) \in \operatorname{dom} I^{B_2}$.

Condition 3 follows from the definition of I^*.

Conditions 4 and 5 follow similarly as 2 since $\mathcal{B}$ is a chain.

□

The next lemma is the last step to using Zorn's lemma for the extension and shows that any extension can be extended further unless they are the desired representation defined on the interior of $u(H)$.

Lemma A.24 Assume that $S = S^{\mathcal{P}}$. Let $\succsim$ satisfy Axioms 1-1.4,1.6, and for all $s \in S^{\mathcal{P}}$, let $u_s : \Delta(X_s) \to \mathbb{R}$ be a weak affine representation for $\succsim_s$ such that

$$u_s(f_s^*) - u_s(g_s^*) = 1.$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1, a_2 \in \mathbb{R} \cup \{-\infty, \infty\}, a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \operatorname{int} u(H)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$, if there exist $f', g' \in H$ such that $u(f'), u(g') \in \operatorname{dom} \widetilde{I}$ and $f' \succsim f \succsim g'$, then $u(f) \in \operatorname{dom} \widetilde{I}$.
3. For all $\alpha \in (0,1)$, $u\big(\alpha f^* + (1-\alpha)g^*\big) \in \operatorname{dom} \widetilde{I}$ and
$$\widetilde{I} \circ u\big(\alpha f^* + (1-\alpha)g^*\big) = \alpha.$$
4. If $\varphi \in \operatorname{dom} \widetilde{I}$ and $\alpha^1 > 0$ are such that $\varphi + \alpha^1 \bar{1} \in \operatorname{dom} \widetilde{I}$, then $\widetilde{I}(\varphi + \alpha^1 \bar{1}) = \widetilde{I}(\varphi) + \alpha^1$.
5. For all $f, g \in H$ with $u(f), u(g) \in \operatorname{dom} \widetilde{I}$, $f \succsim g$ iff $\widetilde{I}(u(f)) \geq \widetilde{I}(u(g))$.

If $A \in \mathcal{A}$ with an associated function I^A and there exists $\varphi \in \operatorname{int} u(H)$ such that $\varphi \notin \operatorname{dom} I^A$, then there exists $A^* \supsetneq A$ such that $A^* \in \mathcal{A}$.

Proof. Let $a_1, a_2 \in \mathbb{R} \cup \{\infty, -\infty\}$ be such that $A = (a_1, a_2)$. Assume that there exists $f^1 \in H$ such that $u(f^1) \in \operatorname{int} u(H)$ and $u(f^1) \notin \operatorname{dom} I^A$. Assume w.l.o.g. $f^1 \succsim h^*$. Let

$$\alpha^2 \coloneqq \sup\{\alpha \in [0,1] | u(\alpha f^1 + (1-\alpha)h^*) \notin \operatorname{dom} I^A\}. \tag{A.25}$$

Since $u(f^1) \notin \operatorname{dom} I^A$, $a^2 \in [0,1]$. Denote

$$f^2 \coloneqq \alpha^2 f^1 + (1-\alpha^2)h^*. \tag{A.26}$$

Now $u(f^2) \in \operatorname{int} u(H)$ and thus there exists $\varepsilon > 0$ such that $B_\infty(u(f^2), \varepsilon) \subseteq u(H)$. Let $\beta^* \in (1/2, 1)$ be such that

$$\big(\frac{1-\beta^*}{\beta^*}\big)\Big(\big\|u(f^2)\big\|_\infty + \|u(h^*)\|_\infty\Big) < \varepsilon.$$

By the choice of β^*, there exists $f^3 \in H$ such that

$$u(f^3) = u(f^2) + \big(\frac{1-\beta^*}{\beta^*}\big)u(f^2) - \frac{1-\beta^*}{\beta^*}u(h^*).$$

Especially, $u(f^3) \in \operatorname{int} u(H)$ and

$$\beta^* u(f^3) + (1-\beta^*)u(h^*) = u(f^2). \tag{A.27}$$

Next, we show that $f^2 \succsim f^*$: Assume, per contra, $f^* \succ f^2$. By applying Axiom 1.2 three times and by Lemma A.18, there exist $\alpha, \hat{\alpha} \in (0,1)$ such that for all $\tilde{\alpha} \in [0, \alpha)$,

$$\hat{\alpha} f^* + (1-\hat{\alpha}) g^* \succ \tilde{\alpha} f^1 + (1-\tilde{\alpha}) f^2 = \big(\tilde{\alpha} + (1-\tilde{\alpha})\alpha^2\big) f^1 + (1-\alpha^2)(1-\tilde{\alpha}) h^* \succ \hat{\alpha} g^* + (1-\hat{\alpha}) f^*.$$

Thus by Conditions 2 and 3, for all $\tilde{\alpha} \in [0, \alpha)$,

$$u\Big(\big(\tilde{\alpha} + (1-\tilde{\alpha})\alpha^2\big) f^1 + (1-\alpha^2)(1-\tilde{\alpha}) h^*\Big) \in \operatorname{dom} I^A,$$

which contradicts (A.25). Thus $f^2 \succsim f^*$. Especially, by Axiom 1.2,

$$\alpha^2 > 0. \tag{A.28}$$

Additionally, we have by Lemma A.17,

$$\beta^* f^3 + (1-\beta^*) h^* \sim f^2 \succsim f^* \succ h^*.$$

Since by Lemma A.40, h^* and g^* are equally crisp and by Lemma A.19, we have for all $\alpha \in [0,1]$,

$$\beta^* f^3 + (1-\beta^*)(f^* \alpha g^*) \succsim \beta^* f^3 + (1-\beta^*) g^* \succ \beta^* h^* + (1-\beta^*) g^* \succ g^*. \tag{A.29}$$

We show next that for all $\alpha \in [0, 1/2)$,

$$\beta^* u(f^3) + (1-\beta^*) u(f^* \alpha g^*) \in \operatorname{dom} I^A. \tag{A.30}$$

Let $\alpha \in [0, 1/2)$. Let $0 < \varepsilon^\alpha < \alpha^2$ be such that

$$0 < \varepsilon^\alpha < \frac{(1-\beta^*)(1/2 - \alpha) \min_{s \in S^{\mathcal{P}}} |u_s(f_s^*) - u(g_s^*)|}{\|u(h^*) - u(f^1)\|_\infty} \tag{A.31}$$

and $u((\alpha^2 - \varepsilon^\alpha) f^1 + (1 - \alpha^2 + \varepsilon^\alpha) h^*) \in \operatorname{dom} I^A$ that exists by (A.25,A.28) and since $u(h^*) \neq u(f^1)$ by Condition 3. Then we have for all $s \in S^{\mathcal{P}}$,

$$\begin{aligned}
&u_s((\alpha^2 - \varepsilon^\alpha) f_s^1 + (1 - \alpha^2 + \varepsilon^\alpha) h_s^*) \\
&\geq \alpha^2 u_s(f_s^1) + (1-\alpha^2) u_s(h_s^*) - \varepsilon^\alpha \Big\| u(h^*) - u(f^1) \Big\|_\infty \\
&\overset{(A.26,A.31)}{>} u_s(f^2) - \beta^*(1/2 - \alpha) \min_{\tilde{s} \in S^{\mathcal{P}}} \Big| u_{\tilde{s}}(f_{\tilde{s}}^*) - u_{\tilde{s}}(g_{\tilde{s}}^*) \Big| \\
&\overset{(A.27)}{\geq} u_s\big(\beta^* f_s^3 + (1-\beta^*) h_s^*\big) - \beta^* \Big(\frac{1}{2} - \alpha\Big) \big(u_s(f_s^*) - u_s(g_s^*)\big) \\
&= \beta^* u_s(f_s^3) + (1-\beta^*) \Big(\frac{1}{2} u_s(f_s^*) + \frac{1}{2} u_s(g_s^*)\Big) + \beta^* \Big(\alpha - \frac{1}{2}\Big) \big(u_s(f_s^*) - u_s(g_s^*)\big) \\
&= \beta^* u_s(f_s^3) + (1-\beta^*) \big(\alpha u_s(f_s^*) + (1-\alpha) u_s(g_s^*)\big) = u_s\big(\beta^* f_s^3 + (1-\beta^*)(f_s^* \alpha g_s^*)\big).
\end{aligned}$$

By Lemma A.17,

$$(\alpha^2-\varepsilon^\alpha)f^1+(1-\alpha^2+\varepsilon^\alpha)g^* \succsim \beta^* f^3+(1-\beta^*)(f^*\alpha g^*) \overset{(A.29)}{\succ} g^*.$$

By Condition 2,

$$\beta^* u(f^3)+(1-\beta^*)u(f^*\alpha g^*)\in \operatorname{dom} I^A$$

Next, we show that

$$I^A(u(\beta^* f^3+(1-\beta^*)g^*))+(1-\beta^*)\frac{1}{2}=a_2 \tag{A.32}$$

and so especially $a_2<\infty$. By Condition 5 and (A.30),

$$I^A(u(\beta^* f^3+(1-\beta^*)g^*))+(1-\beta^*)\frac{1}{2}\le a_2.$$

Assume, per contra, that there exists $k\in H$ such that $u(k)\in \operatorname{dom} I^A$ and

$$I^A\circ u(k) > I^A(u(\beta^* f^3+(1-\beta^*)g^*))+(1-\beta^*)\frac{1}{2}.$$

Then by (A.30) and Conditions 4 and 5, for all $\alpha\in(0,1/2)$,

$$k\succsim \beta^* f^3+(1-\beta^*)(f^*\alpha g^*).$$

Thus by Axiom 1.2,

$$k\succsim \beta^* f^3+(1-\beta^*)h^*.$$

By Condition 2, $\beta^* f^3+(1-\beta^*)h^*\in \operatorname{dom} I^A$ which is a contradiction. This shows (A.32).

Denote

$$B=\{\varphi\in \operatorname{int} u(H)\,|\,\exists\, f\in H,\alpha\in(0,1),u(f)=\varphi, f\sim \beta^* f^3+(1-\beta^*)(f^*\alpha g^*)\}.$$

We define an extension of I^A, $I^*: B\cup \operatorname{dom} I^A\to\mathbb{R}$ as follows: for all $\varphi\in\operatorname{dom} I^A$, $I^*(\varphi)=I^A(\varphi)$. For all $f\in H$ such that $u(f)\notin \operatorname{dom} I^A$, $u(f)\in \operatorname{int} u(H)$, and there exists $\alpha\in[0,1)$ such that $f\sim\beta^* f^3+(1-\beta^*)(f^*\alpha g^*)$, define $I^*(u(f))=I^A\big(u(\beta^* f^3+(1-\beta^*)g^*)\big)+(1-\beta^*)\alpha$.

First, by (A.32), $\operatorname{Im} I^*\supsetneq A$. Second, by Lemmas A.17 and A.19 and Conditions 4 and 5, I^* is well-defined. Especially, we have for all $\alpha\in[0,1/2)$, $I^A(u(\beta^* f^3+(1-\beta^*)(f^*\alpha g^*)))=I^A(u(\beta^* f^3+(1-\beta^*)g^*))+(1-\beta^*)\alpha$ by Condition 4 and (A.30). Thus for all $\alpha\in[0,1)$,

$$I^*(u(\beta^* f^3+(1-\beta^*)(f^*\alpha g^*)))=I^A(u(\beta^* f^3+(1-\beta^*)g^*))+(1-\beta^*)\alpha. \tag{A.33}$$

We show that I^* satisfies Conditions 1-5 for the interval $\big(a_1, a_2+\frac{1}{2}(1-\beta^*)\big)$.

Condition 1: For $a \in (a_1, a_2)$, there exists $\varphi \in \operatorname{dom} I^*$ such that $I^*(\varphi) = a$ since I^* is an extension of I^A. Let $a \in [a_2, a_2 + \frac{1}{2}(1-\beta^*))$. We have $\frac{(a-a_2)}{(1-\beta^*)} < \frac{1}{2}$ and so

$$I^* \circ u\Big(\beta^* f^3 + (1-\beta^*)\Big[f^*\Big(\frac{1}{2} + \frac{(a-a_2)}{(1-\beta^*)}\Big)g^*\Big]\Big)$$
$$\overset{(A.33)}{=} I^A\big(u(\beta^* f^3 + (1-\beta^*)g^*)\big) + (1-\beta^*)\Big(\frac{1}{2} + \frac{(a-a_2)}{(1-\beta^*)}\Big) \overset{(A.32)}{=} a_2 + a - a_2 = a.$$

Condition 2: Let $f \in H$ be such that $u(f) \in \operatorname{int} u(H)$ and there exist $f', g' \in H$ such that $u(f'), u(g') \in \operatorname{dom} I^*$ and $f' \succsim f \succsim g'$. If $\beta^* f^3 + (1-\beta^*)(f^{*1/2}g^*) \succ f$, then by (A.30) $u(f) \in \operatorname{dom} I^A \subseteq \operatorname{dom} I^*$. If $f \succsim \beta^* f^3 + (1-\beta^*)(f^{*1/2}g^*)$, then $u(f') \in B$. By Axiom 1.2, $u(f) \in B$.

Condition 3: Follows directly since I^A satisfies Condition 3.

Condition 4: By Lemma A.21 and Condition 4 for I^A, I^* satisfies Condition 4 in $\operatorname{dom} I^A$ and B. So the only case left by Lemma A.17 is that $\varphi \in \operatorname{dom} I^A \setminus B$ and $\varphi + \alpha^0 \in B \setminus \operatorname{dom} I^A$. Let $\varphi^{-1}, f^{\alpha^0} \in H$ be such that $u(\varphi^{-1}) = \varphi$ and $u(f^{\alpha^0}) = \varphi + \alpha^0 \bar{1}$. Since $u(f^{\alpha^0}) \notin \operatorname{dom} I^A, \varphi \notin B$, by (A.30) $f^{\alpha^0} \succsim \beta^* f^3 + (1-\beta^*)(f^{*1/4}g^*) \succ \varphi^{-1}$. Thus there exists $\alpha^1 \in (0, \alpha^0)$ such that $f^{\alpha^0}\alpha^1\varphi^{-1} \sim \beta^* f^3 + (1-\beta^*)(f^{*1/4}g^*)$. Then we have by Condition 4 in $\operatorname{dom} I^A$ and B,

$$I^*(\varphi + \alpha^0\bar{1}) - I^*(\varphi) = I^*(\varphi + \alpha^0\bar{1}) - I^*(\varphi + \alpha^1\alpha^0\bar{1}) + I^*(\varphi + \alpha^1\alpha^0\bar{1}) - I^*(\varphi)$$
$$= (1-\alpha^1)\alpha^0 + I^A(\varphi + \alpha^1\alpha^0\bar{1}) - I^A(\varphi) = (1-\alpha^1)\alpha^0 + \alpha^1\alpha^0 = \alpha^0.$$

This show Condition 4.

Condition 5: Let $f, g \in H$ be such that $u(f), u(g) \in \operatorname{dom} I^*$. If $u(f), u(g) \in \operatorname{dom} I^A$ or $u(f), u(g) \in B$, then the condition holds. So assume that $u(f) \in \operatorname{dom} I^* \setminus B$ and $u(g) \in \setminus B \setminus I^A$. Then by (A.30) and the definition of B,

$$g \succ \beta^* f^3 + (1-\beta^*)(f^{*1/4}g^*) \succ f$$

and

$$I^*\big(u(g)\big) > I^*(\beta^* u(f^3) + (1-\beta^*)u(f^{*1/4}g^*)) > I^*\big(u(f)\big).$$

□

The next result applies Zorn's lemma for the desired representation.

Proposition A.25 Assume that $S = S^{\mathcal{P}}$. Let $\succsim$ satisfy Axioms 1-1.4,1.6, and for all $s \in S^{\mathcal{P}}$, let $u_s : \Delta(X_s) \to \mathbb{R}$ be a weak affine representation for $\succsim_s$ such that $u_s(f_s^*) - u_s(g_s^*) = 1$. Then there exists $I : \operatorname{int} u(H) \to \mathbb{R}$ such that I is C-additive and for all $f, g \in H$ such that $u(f), u(g) \in \operatorname{int} u(H)$,

$$f \succsim g \iff I\big(u(f)\big) \geq I\big(u(g)\big).$$

Proof. Let f^*, g^* be equally crisp acts. By Lemma A.16, for all $s \in S^{\mathcal{P}}$, there exist an affine $u_s : \Delta(X_s) \to \mathbb{R}$ and continuous and monotonic $H_s : \operatorname{Im} u_s(\Delta(X_s)) \to \mathbb{R}$ such that for all $x_s, y_s \in \Delta(X_s)$,

$$x_s \succsim_s y_s \iff H_s(u_s(x_s)) \geq H_s(u_s(y_s))$$

and $u_s(f_s^*) - u_s(g_s^*) = 1$. We construct I that is C-additive and $I \circ u$ represents preferences. Let for all $\alpha \in [0,1]$, $I^0 \circ u(\alpha f^* + (1-\alpha)g^*) = \alpha$. For all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$ and there exists $\alpha \in [0,1]$ such that $f \sim \alpha f^* + (1-\alpha)g^*$, define $I^0 \circ u(f) = \alpha$. Next, we do a transfinite induction on the collection $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1, a_2 \in \mathbb{R} \cup \{-\infty, \infty\}, a_1 \leq a_2\}$ such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \operatorname{int} u(H)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$, if there exist $f', g' \in H$ such that $u(f'), u(g') \in \operatorname{dom} \widetilde{I}$ and $f' \succsim f \succsim g'$, then $u(f) \in \operatorname{dom} \widetilde{I}$.
3. For all $\alpha \in (0,1)$, $u\big(\alpha f^* + (1-\alpha)g^*\big) \in \operatorname{dom} \widetilde{I}$ and
$$\widetilde{I} \circ u\big(\alpha f^* + (1-\alpha)g^*\big) = \alpha.$$
4. If $\varphi \in \operatorname{dom} \widetilde{I}$ and $\alpha^1 > 0$ are such that $\varphi + \alpha^1 \bar{1} \in \operatorname{dom} \widetilde{I}$, then $\widetilde{I}(\varphi + \alpha^1 \bar{1}) = \widetilde{I}(\varphi) + \alpha^1$.
5. For all $f, g \in H$ with $u(f), u(g) \in \operatorname{dom} \widetilde{I}$, $f \succsim g$ iff $\widetilde{I}(u(f)) \geq \widetilde{I}(u(g))$.

We order $\mathcal{A}$ by $\subseteq$. By Lemma A.22, $\mathcal{A} \neq \varnothing$. By Lemma A.23, every chain has an upper bound. By Zorn's lemma, there exists a maximal element $A^* \in \mathcal{A}$ and let I^* be the associated function. By Lemma A.24 and the maximality of A^*, $\operatorname{dom} I^* = \operatorname{int} u(H)$. This shows the claim. □

The next result shows that if there exists a crisp act with utility 0, then the representation is positive homogeneous.

Proposition A.26 Assume that $S = S^{\mathcal{P}}$. Let $\succsim$ satisfy Axioms 1-1.4,1.6, and for all $s \in S^{\mathcal{P}}$, let $u_s : \Delta(X_s) \to \mathbb{R}$ be a weak affine representation for $\succsim_s$ such that $u_s(f_s^*) - u_s(g_s^*) = 1$, and $u_s(c_s) = 0$. Then there exists $I : \operatorname{int} u(H) \to \mathbb{R}$ such that I is C-additive, positive homogeneous, and for all $f, g \in H$ such that $u(f), u(g) \in \operatorname{int} u(H)$,

$$f \succsim g \iff I\big(u(f)\big) \geq I\big(u(g)\big).$$

Proof. By Proposition A.25, there exists $I : \operatorname{int} u(H) \to \mathbb{R}$ such that I is C-additive and for all $f, g \in H$ such that $u(f), u(g) \in \operatorname{int} u(H)$,

$$f \succsim g \iff I\big(u(f)\big) \geq I\big(u(g)\big)$$

and after adding a constant since $u(c)=\bar{0}\in u(H)$, $\lim_{\operatorname{int} u(H)\ni\varphi\to\bar{0}} I(\varphi)=0$. We show that I is positively homogeneous. Let $f\in H$ be such that $u(f)\in\operatorname{int} u(H)$. Now for all $\alpha\in(0,1]$, $\alpha\varphi\in\operatorname{int} u(H)$. Denote the mapping $(0,1)\ni\alpha\overset{J}{\mapsto} I\big(\alpha u(f)\big)$. Since I is 1-Lipschitz, J is Lipschitz function and hence differentiable almost everywhere. Let the differentiability domain of J be $\Omega\subseteq(0,1)$ and Ω is dense in $(0,1)$.

We show that ∇J is locally constant in Ω. If $u(f)=\bar{0}$, then the claim follows from the normalization. Thus assume $u(f)\neq\bar{0}$. Let $x\in(0,1)$. Then there exists $y\in\Omega$ such that $x<y$. Let $z\in\Omega$ such that $z<y$. We show that $\nabla J(y)=\nabla J(z)$. Since $y>0$, $u(yf+(1-y)c)\in\operatorname{int} u(H)$ and there exists $\varepsilon>0$ such that $B_\infty(u(yf+(1-y)c),\varepsilon)\subseteq\operatorname{int} u(H)$. Let $\gamma\in(0,1)$ be such that $|y-\gamma|<\frac{\varepsilon}{\|u(f)\|_\infty}<\infty$. Since I is C-additive and hence especially 1-Lipschitz, we have $|J(y)-J(\gamma)|\leq\|u(f)\|_\infty|y-\gamma|<\varepsilon$. Thus there exists $g\in H$ such that $u(g)=u(yf+(1-y)c)+\bar{1}(J(\gamma)-J(y))\in\operatorname{int} u(H)$ by C-additivity, we have

$$\begin{aligned}I\big(u(g)\big)&=I\Big(yu(f)+(1-y)u(c)+\bar{1}\big(J(\gamma)-J(y)\big)\Big)\overset{\text{C-add.}}{=}J(y)+J(\gamma)-J(y)\\&=J(\gamma)=I\big(\gamma u(f)+(1-\gamma)u(c)\big)\end{aligned}$$

By the representation and the crispness of c, for all $\alpha\in(0,1)$,

$$g\sim\gamma f+(1-\gamma)c\text{ and }\alpha g+(1-\alpha)c\sim\alpha\big(\gamma f+(1-\gamma)c\big)+(1-\alpha)c.$$

By the representation,

$$\begin{aligned}&I\big(\alpha u(g)+(1-\alpha)u(c)\big)=I\Big(\alpha yu(f)+\alpha\bar{1}\big(J(\gamma)-J(y)\big)\Big)\overset{\text{C-add.}}{=}J(\alpha y)+\alpha\big(J(\gamma)-J(y)\big)\\&=I\Big(\alpha\big(\gamma u(f)+(1-\gamma)u(c)\big)+(1-\alpha)u(c)\Big)=I\big(\alpha\gamma u(f)\big)=J(\alpha\gamma).\end{aligned}$$

We have for all $\gamma\in(0,1)$ such that $0<|y-\gamma|<\frac{\varepsilon}{\|u(f)\|_\infty}<\infty$ and $\alpha\in(0,1)$,

$$J(\alpha\gamma)=J(\alpha y)+\alpha\big(J(\gamma)-J(y)\big)\iff\frac{J(\alpha\gamma)-J(\alpha y)}{\alpha\gamma-\alpha y}=\frac{J(\gamma)-J(y)}{\gamma-y}.$$

By taking $\gamma\to y$ and $\alpha=z/y<1$, we have since $y,z\in\Omega$

$$\nabla J(y)=\lim_{\gamma\to y}\frac{J(\gamma)-J(y)}{\gamma-y}=\lim_{\gamma\to y}\frac{J(z/y\gamma)-J(z)}{z/y\gamma-z}=\nabla J(z).$$

Since $z\in\Omega$ such that $z<y$ was arbitrary, ∇J is constant in the (0,y) neighborhood of x. Since $x\in(0,1)$ was arbitrary, ∇J is locally constant. Since the interval $(0,1)$ is connected, by (Viro et al., 2008, Problem 12.2x) ∇J is constant on Ω. Since J is Lipschitz function and by the normalization $\lim_{\widetilde{\alpha}\to0}J(\widetilde{\alpha})=0$, we have for all $\alpha\in(0,1)$,

$$J(\alpha)-\lim_{\widetilde{\alpha}\to0}J(\widetilde{\alpha})=\int_{(0,\alpha)\cap\Omega}\nabla J(\alpha)d\alpha=\alpha\nabla J.$$

Thus J is a linear function and by taking $\alpha = 1$, $\nabla J = J(1) = I(u(f))$. Thus for all $\alpha \in (0, 1]$,

$$I(\alpha u(f)) = \alpha I\big(u(f)\big).$$

Since $f \in H$ such that $u(f) \in \operatorname{int} u(H)$ was arbitrary, this shows the claim. □

Finally, we put the above together and show that these representations are dual-self representations and extend the representation to the situation with non-proper states.

Proposition A.27 Assume that $S = S^{\mathcal{P}}$. If $\succsim$ satisfies Axioms 1-1.6, then there exists $(u, \mathbb{P})$ that is a dual-self representation for $\succsim$.

Proof. By Lemma A.16 and Proposition A.26, for all $s \in S$, there exist affine $u_s : \Delta(X_s) \to \mathbb{R}$ and $I : \operatorname{int} u(H) \to \mathbb{R}$ such that $I \circ u$ represents $\succsim$ in interior $u(H)$ and I is C-additive and positive homogeneous. Thus by Chandrasekher et al.'s (2020) Lemma A.5 and the proof of Theorem 1, there exists $\mathbb{P} \in \mathcal{K}(\Delta(S))$ such that for all $\varphi \in \operatorname{int} u(H)$, $I(\varphi) = \max_{P \in \mathbb{P}} \min_{p \in P} p \cdot \varphi$. Finally, the representation can be extended to $u(H)$, which gives the claim. □

Proposition A.28 Assume that $S = S^{\mathcal{P}}$. If $\succsim$ satisfies Axioms 1-1.4, 1.6, then there exists $(u, \mathbb{C})$ that is a dual-self variational representation for $\succsim$.

Proof. By Lemma A.16 and Proposition A.25, for all $s \in S$, there exist affine $u_s : \Delta(X_s) \to \mathbb{R}$ and $I : \operatorname{int} u(H) \to \mathbb{R}$ such that $I \circ u$ represents $\succsim$ in interior $u(H)$ and I is C-additive and $u(f^*) = \bar{1}, u(g^*) = \bar{0}$. Thus by Chandrasekher et al.'s (2020) Lemma A.5 and Lemma S.3.2, there exists $\mathbb{C} \subseteq \{c : \Delta(S) \to \mathbb{R} \cup \{\infty\} | c \text{ is convex}\}$ such that for all $\varphi \in \operatorname{int} u(H)$, $I(\varphi) = \max_{c \in \mathbb{C}} \min_{p \in \Delta(S)} p \cdot \varphi + c(p)$. Finally, the representation can be extended to $u(H)$, which gives the claim. □

Proposition A.29 If $\succsim$ satisfies Axioms 1-1.6, then there exists $(u, \mathbb{P})$ that is a dual-self representation for $\succsim$.

Proof. Denote $\widehat{H} \coloneqq \bigtimes_{s \in S^{\mathcal{P}}} \Delta(X_s)$ and define $\widehat{\succsim} \subseteq \widehat{H} \times \widehat{H}$ by for all $\widehat{f}, \widehat{g} \in \widehat{H}$,

$$\widehat{f} \widehat{\succsim} \widehat{g} \iff \forall\, h \in H(\widehat{f}, h_{S \setminus S^{\mathcal{P}}}) \succsim (\widehat{g}, h_{S \setminus S^{\mathcal{P}}}).$$

Now $\widehat{\succsim}$ satisfies Axioms 1-1.6 and it does not have null-states. By Proposition A.27, there exist $(\widehat{u}, \widehat{\mathbb{P}})$ that is a dual-self representation for $\widehat{\succsim}$ and a crisp act $c \in \widehat{H}$ such that for all $s, s' \in S^{\mathcal{P}}$, $u_s(c_s) = u_{s'}(c_{s'}) =: c^*$. For all $s \in S \setminus S^{\mathcal{P}}$, define $u_s : \Delta(X_s) \to \mathbb{R}$ by for all $x_s \in \Delta(X_s)$, $u_s(x_s) = c^*$ and then u_s is affine. For all $s \in S^{\mathcal{P}}$, define $u_s = \widehat{u}_s$. Additionally, define

$$\mathbb{P} \coloneqq \{\{p \in \Delta(S) | (p_s)_{s \in S^{\mathcal{P}}} \in \widehat{P}\} | \widehat{P} \in \widehat{\mathbb{P}}\}.$$

Now for all $s \notin S^{\mathcal{P}}$ and $p \in P \in \mathbb{P}$, $p_s = 0$. Additionally, by the definition of null-states and $\widehat{\succsim}$, $(u, \mathbb{P})$ is a dual-self representation for $\succsim$. □

Proposition A.30 If $\succsim$ satisfies Axioms 1-1.4, 1.6, then there exists $(u, \mathbb{C})$ that is a dual-self variational representation for $\succsim$.

Proof. Symmetrically to Proposition A.29. □

A.3.4 Necessity of the Axioms for Representations

Finally, we show the necessity of our axioms. First, for state dependent dual-self variational expected utility and second for state dependent dual-self expected utility.

Lemma A.31 Let $\succsim$ have a state dependent dual-self variational representation with $(u, \mathbb{C})$. Then u_s is a weak affine representation for $\succsim_s$.

Proof. Follows from the definition of $\succsim_s$ and the monotonicity of the dual-self representation. □

Lemma A.32 Let $\succsim$ have a state dependent dual-self variational representation with $(u, \mathbb{C})$. Then $\succsim$ satisfies Axioms 1-1.4,1.6.

Proof. Define $I : u(H) \to \mathbb{R}$ by for all $f \in H$,

$$I\big(u(f)\big) = \max_{c \in \mathbb{C}} \min_{p \in \Delta(S)} p \cdot u(f) + c(p).$$

Now I is C-additive on the non-null states and monotonic. Thus especially it is 1-Lipschitz.

Completeness and transitivity follow from representation. Nontriviality follows from the definition of dual-self variational representation: Since $\mathbb{C}$ is grounded, especially there exist $c \in \mathbb{C}$ and $p \in \Delta(S)$ such that $c(p) < \infty$. Since there exists $s \in S$ such that $p_s > 0$, by the definition of dual-self variational representation, s is not null. Thus by the definition of null states, there exist $x_s, y_s \in \Delta(X_s), h \in H$ such that $(x_s, h_{-s}) \not\sim (y_s, h_{-s})$.

Axiom 1.2: Let $f, g, h \in H$. Define the mapping $[0,1] \ni \alpha \overset{J}{\mapsto} I\big(\alpha u(f) + (1-\alpha)u(g)\big)$. Now J is continuous since I is continuous and thus the following sets are closed as preimages of closed sets of a continuous function

$$J^{-1}\Big[I\big(u(h)\big), \infty\Big) = \{\alpha \in [0,1] | \alpha f + (1-\alpha)g \succsim h\} \text{ and}$$
$$J^{-1}\Big(-\infty, I\big(u(h)\big)\Big] = \{\alpha \in [0,1] | h \succsim \alpha f + (1-\alpha)g\}.$$

Axiom 1.3: Follows from the monotonicity of I.

Axiom 1.4: Follows from Lemmas A.11 and A.31.

Axiom 1.6: For all $s \in S^{\mathcal{P}}$, by above and Lemmas A.7, A.14, and A.15, there exist $x_s^*, x_{s*} \in \Delta(X_s), \alpha^s \in (0,1)$ such that

$$x_s^* \succ_s \alpha^s x_s^* + (1-\alpha^s)x_{s*} \succ_s x_{s*}$$

and either for all $\alpha \in (\alpha^s, 1]$,

$$\alpha x_s^* + (1-\alpha)x_{s*} \succ_s \alpha^s x_s^* + (1-\alpha^s)x_{s*} \tag{A.34}$$

or for all $\alpha \in [0, \alpha_*)$,

$$\alpha^s x_s^* + (1-\alpha^s)x_{s*} \succ_s \alpha x_s^* + (1-\alpha)x_{s*}.$$

Since u_s is a weak representation for $\succsim_s$, $\alpha^s x_s^* + (1-\alpha^s)x_{s*} \in \operatorname{int} u_s\big(\Delta(X_s)\big)$. Thus there exists $\alpha^* > 0$ such that for all $s \in S^{\mathcal{P}}$, there exist f_s^*, g_s^* such that

$$u_s(f_s^*) > u_s\big(\alpha^s x_s^* + (1-\alpha^s)x_{s*}\big) > u_s(g_s^*)$$

and $u_s(f_s^*) - u_s(g_s^*) = \alpha^*$. We show that $f_s^* \succ_s g_s^*$. Assume w.l.o.g. that for all $\alpha \in (\alpha^s, 1]$,

$$\alpha x_s^* + (1-\alpha)x_{s*} \succ_s \alpha^s x_s^* + (1-\alpha^s)x_{s*}.$$

Then by the affinity of u_s, there exists $\alpha^1 \in (\alpha^s, 1]$ such that $u_s(f_s^*) > u_s\big(\alpha^1 x_s^* + (1-\alpha^1)x_{s*}\big)$. Since u_s is a weak representation for $\succsim_s$, we have

$$f_s^* \succsim \alpha^1 x_s^* + (1-\alpha^1)x_{s*} \overset{(A.34)}{\succ} \alpha^s x_s^* + (1-\alpha^s)x_{s*} \succsim g_s^*.$$

Let $\widetilde{h} \in H$ and define $\widetilde{f}^*, \widetilde{g}^* \in H$ by for all $s \in S^{\mathcal{P}}$, $\widetilde{f}_s^* = f_s^*$, $\widetilde{g}_s^* = g_s^*$ and for all $s \notin S^{\mathcal{P}}$, $\widetilde{f}_s^* = \widetilde{h}_s = \widetilde{g}_s^*$. We show that f^*, g^* are equally crisp. Let $\alpha \in (0,1], f, g \in H$, then we have since I is C-additive on the non-null states

$$I\big(u(\alpha f + (1-\alpha)\widetilde{g}^*)\big) = I\big(u(\alpha f + (1-\alpha)\widetilde{f}^*)\big) - (1-\alpha)\alpha^*$$

and

$$I\big(u(\alpha g + (1-\alpha)\widetilde{g}^*)\big) = I\big(u(\alpha g + (1-\alpha)\widetilde{f}^*)\big) - (1-\alpha)\alpha^*.$$

This shows that $\widetilde{f}^*$, $\widetilde{g}^*$ are equally crisp acts such that for all $s \in S^{\mathcal{P}}$, $\widetilde{f}^* \succ_s \widetilde{g}^*$. □

Lemma A.33 Let $\succsim$ have a state dependent dual-self representation with $(u, \mathbb{P})$ such that assume that there exists $c \in H$ such that for all $p, q \in \bigcup_{P \in \mathbb{P}} P$, $\sum_{s \in S} p_s u_s(c_s) = \sum_{s \in S} q_s u_s(c_s)$. Then $\succsim$ satisfies Axioms 1-1.6.

Proof. By Lemma A.32, we need to only show non-triviality and Axiom 1.5. Non-triviality follows from that since $\mathbb{P}$ is not empty there exists $p \in P \in \mathbb{P}$, and by the definition of dual-self representation, for all $s \in S^{\mathcal{P}}$, $p_s = 0$ and thus there exists $s \in S$ such that s is non-null. By the definition of null states, $\succsim$ is non-trivial.

We show that c from the antecedent is crisp. Denote Let $p \in \bigcup_{P\in\mathbb{P}} P$ and denote $c^* \coloneqq p \cdot u(c)$. Let $f \in H, \alpha \in (0,1)$. Then we have

$$\max_{P\in\mathbb{P}} \min_{p\in P} p \cdot u(\alpha f + (1-\alpha)c) = \max_{P\in\mathbb{P}} \min_{p\in P} \alpha p \cdot u(f) + (1-\alpha) p \cdot u(c) = \alpha \max_{P\in\mathbb{P}} \min_{p\in P} p \cdot u(f) + (1-\alpha)c^*.$$

This shows the crispness of c. □

A.3.5 State Independent Dual-Self Expected Utility

Assume state independent setting with for all $s \in S$, $X_s = X$. We start off by showing that under the conditions of Remark 1.2, all the state dependent utilities are positive affine transformations of each other. We show this by showing that there exists a state s such that the utilities for all the other states give a weak affine representation for $\succsim_s$. Our first lemma shows that if there is a constant crisp act and another consequence better than the crisp consequence, then the statewise preferences are not censored above the crisp consequence and similarly for below.

Lemma A.34 Let $\succsim$ be preferences on $\Delta(X)$ such that there exist affine $u : \Delta(X) \to \mathbb{R}$ and continuous and weakly increasing $H : u(\Delta(X)) \to \mathbb{R}$ such that for all $x, y \in \Delta(X)$,

$$x \succsim y \iff H(u(x)) \geq H(u(y)).$$

Assume that there exists $x^* \in \Delta(X)$ such that for all $x, y \in \Delta(X), \alpha \in (0,1)$,

$$x \succ y \Rightarrow \alpha x + (1-\alpha)x^* \succ \alpha y + (1-\alpha)x^*.$$

Then the following conditions hold.

1. If there exist $x, y \in \Delta(X)$ such that $x \succ y \succ x^*$, then for all $\alpha, \beta \in [u(x^*), u(y)], \alpha > \beta$, we have $H(\alpha) > H(\beta)$.
2. If there exist $x, y \in \Delta(X)$ such that $x^* \succ y \succ x$, then for all $\alpha, \beta \in [u(y), u(x^*)], \alpha > \beta$, we have $H(\alpha) > H(\beta)$.

Proof. We will show (1) since (2) follows symmetrically. Now $u(x) > u(y)$ and $H(u(x)) > H(u(y))$ By continuity and weakly increasing of H, there exists $x^\dagger \in \Delta(X)$ such that $u(x^\dagger) \leq u(x)$, $H(u(x^\dagger)) =$

$H(u(x))$ and for all $\alpha \in u(\Delta(X))$ such that $\alpha < u(x^\dagger)$, $H(\alpha) < H(u(x^\dagger))$. Now we have $x^\dagger \succ y$ and $u(x^\dagger) > u(y)$. Let $\alpha, \beta \in [u(x^*), u(y)], \alpha > \beta$. Let $x^\ddagger \in \Delta(X)$ be such that

$$u(x^\ddagger) = \beta \frac{u(x^\dagger) - u(x^*)}{\alpha - u(x^*)} + u(x^*) \frac{\alpha - u(x^\dagger)}{\alpha - u(x^*)}$$

that exists since $u(\Delta(X))$ is an interval and $u(x^\dagger) > u(x^\ddagger) \geq u(x^*)$. Now we have $x^\dagger \succ x^\ddagger$ by the representation and so by assumption,

$$\frac{\alpha - u(x^*)}{u(x^\dagger) - u(x^*)} x^\dagger + \frac{u(x^\dagger) - \alpha}{u(x^\dagger) - u(x^*)} x^* \succ \frac{\alpha - u(x^*)}{u(x^\dagger) - u(x^*)} x^\ddagger + \frac{u(x^\dagger) - \alpha}{u(x^\dagger) - u(x^*)} x^*.$$

But by the affinity of u, we have

$$u\Big(\frac{\alpha - u(x^*)}{u(x^\dagger) - u(x^*)} x^\dagger + \frac{u(x^\dagger) - \alpha}{u(x^\dagger) - u(x^*)} x^*\Big) = \alpha$$

and

$$\begin{aligned} &u\Big(\frac{\alpha - u(x^*)}{u(x^\dagger) - u(x^*)} x^\ddagger + \frac{u(x^\dagger) - \alpha}{u(x^\dagger) - u(x^*)} x^*\Big) \\ &= \frac{\alpha - u(x^*)}{u(x^\dagger) - u(x^*)} \Big(\beta \frac{u(x^\dagger) - u(x^*)}{\alpha - u(x^*)} + u(x^*) \frac{\alpha - u(x^\dagger)}{\alpha - u(x^*)}\Big) + \frac{u(x^\dagger) - \alpha}{u(x^\dagger) - u(x^*)} u(x^*) = \beta. \end{aligned}$$

Thus by the representation, $H(\alpha) > H(\beta)$.

□

The next lemma shows that if the constant crisp consequence is the best or the worst consequence in some state, then all other utilities provide a weak affine representation for the statewise preferences for this state.

Lemma A.35 Let $\succsim, \succsim'$ be preferences on $\Delta(X)$ that satisfy Axioms 1.1, 1.2, 4′ and have censored EU representations with (H, u) and (H', u') respectively. If for all $x, y \in \Delta(X)$,

$$x \succ y \Rightarrow x \succsim' y \text{ and } x \succ' y \Rightarrow x \succsim y \tag{A.35}$$

and there exists $x^* \in \Delta(X)$ such that for all $x', y' \in \Delta(X), \alpha \in (0, 1)$,

$$x' \succ y' \Rightarrow \alpha x' + (1-\alpha)x^* \succ \alpha y' + (1-\alpha)x^* \text{ and } x' \succ' y' \Rightarrow \alpha x' + (1-\alpha)x^* \succ' \alpha y' + (1-\alpha)x^* \tag{A.36}$$

and $u(x^*) = \sup u(\Delta(X))$ or $u(x^*) = \inf u(\Delta(X))$, then u' is a weak affine representation for $\succsim$.

Proof. The case $u(x^*) = \inf u(\Delta(X))$ follows symmetrically to $u(x^*) = \sup u(\Delta(X))$ so we only consider the latter case. Since $\succsim$ is non-trivial, there exist $\tilde{x}^\dagger, y^\dagger \in \Delta(X)$ such that $\tilde{x}^\dagger \succ y^\dagger$. By Axiom 1.2, there exists $x^\dagger \in \Delta(X)$ such that $\tilde{x}^\dagger \succ x^\dagger \succ y^\dagger$. Now $u(x^*) \geq u(\tilde{x}^\dagger)$ and so by the weak

affine representation, we have $x^* \succsim \tilde{x}^\dagger \succ x^\dagger \succ y^\dagger$. By Lemma A.34, for all $\alpha, \beta \in [u(x^*), u(x^\dagger)], \alpha > \beta$, we have

$$H(\alpha) > H(\beta). \tag{A.37}$$

Assume, per contra, that u' is not a weak affine representation for $\succsim$. Thus there exist $\tilde{x}^\ddagger \succ \tilde{y}^\ddagger$ and $u'(\tilde{y}^\ddagger) \geq u'(\tilde{x}^\ddagger)$. By Axiom 1.1, $\succsim'$ is non-trivial and so for some $a, b \in \Delta(X)$, $u'(a) > u'(b)$. By (A.37) and by applying twice Axiom 1.2, there exist $x^\ddagger, y^\ddagger \in \Delta(X)$ such that $x^* \succ x^\ddagger \succ y^\ddagger$ and $u'(y^\ddagger) > u'(x^\ddagger)$.

We consider cases. First, assume that there exists $x \in \Delta(X)$ such that $x \succ' x^*$. By Axiom 1.2, there exists $\hat{x} \succ' x^*$ and $u(x^*) > u(\hat{x})$. By (A.37), $x^* \succ \hat{x}$ which contradicts (A.35).

By Axiom 1.1, there exists $x^\S \in \Delta(X)$ such that $x^* \succ' x^\S$. By Axiom 1.2, there exists $y^\S \in \Delta(X)$ such that $x^* \succ' y^\S \succ' x^\S$. By Lemma A.34, $\alpha, \beta \in [u'(x^*), u'(y^\S)], \alpha > \beta$, we have

$$H'(\alpha) > H'(\beta). \tag{A.38}$$

Let $z^\S \in \Delta(X)$ be such that $u'(x^*) > u'(z^\S) > u'(y^\S)$. Now there exists $\alpha^0 \in (0,1)$ such that

$$u'(x^*) > u'(\alpha^0 y^\ddagger + (1-\alpha^0) z^\S) > u'(\alpha^0 x^\ddagger + (1-\alpha^0) z^\S) > u'(y^\S).$$

By (A.38),

$$\alpha^0 y^\ddagger + (1-\alpha^0) z^\S \succ' \alpha^0 x^\ddagger + (1-\alpha^0) z^\S. \tag{A.39}$$

Now $u(x^*) > u(x^\ddagger) > u(y^\ddagger)$ and so

$$u(x^*) > u(\alpha^0 x^\ddagger + (1-\alpha^0) z^\S) > u(\alpha^0 y^\ddagger + (1-\alpha^0) z^\S).$$

There exists $\alpha^1 \in (0,1)$ such that

$$u\big(\alpha^1[\alpha^0 y^\ddagger + (1-\alpha^0) z^\S] + (1-\alpha^1) x^*\big) > u(x^\dagger).$$

By (A.37), we have

$$\alpha^1[\alpha^0 x^\ddagger + (1-\alpha^0) z^\S] + (1-\alpha^1) x^* \succ \alpha^1[\alpha^0 y^\ddagger + (1-\alpha^0) z^\S] + (1-\alpha^1) x^*.$$

By (A.36,A.39),

$$\alpha^1[\alpha^0 y^\ddagger + (1-\alpha^0) z^\S] + (1-\alpha^1) x^* \succ' \alpha^1[\alpha^0 x^\ddagger + (1-\alpha^0) z^\S] + (1-\alpha^1) x^*$$

which contradicts (A.35).

□

The next lemma shows that if the crisp consequence is in the interior of statewise preferences in some state, then all other states' utilities give a weak affine representation for the statewise preferences for this state.

Lemma A.36 Let $\succsim, \succsim'$ be preferences on $\Delta(X)$ that satisfy Axioms 1.1, 1.2, $4'$ and have censored EU representations with (H, u) and (H', u') respectively. If for all $x, y \in \Delta(X)$,

$$x \succ y \Rightarrow x \succsim' y \text{ and } x \succ' y \Rightarrow x \succsim y \tag{A.40}$$

and there exists $x^* \in \Delta(X)$ such that for all $x', y' \in \Delta(X), \alpha \in (0,1)$,

$$x' \succ y' \Rightarrow \alpha x' + (1-\alpha)x^* \succ \alpha y' + (1-\alpha)x^* \text{ and } x' \succ' y' \Rightarrow \alpha x' + (1-\alpha)x^* \succ' \alpha y' + (1-\alpha)x^* \tag{A.41}$$

and there exist $\tilde{x}, \tilde{y} \in \Delta(X)$ such that

$$\tilde{x} \succ x^* \succ \tilde{y}, \tag{A.42}$$

then u' is a weak affine representation for $\succsim$.

Proof. By Axiom 1.2, there exist $\hat{x}, \hat{y} \in \Delta(X)$ such that

$$\tilde{x} \succ \hat{x} \succ x^* \succ \hat{y} \succ \tilde{y}.$$

By Lemma A.34, for all $\alpha, \beta \in [u(\hat{y}), u(\hat{x})], \alpha > \beta$, we have

$$H(\alpha) > H(\beta). \tag{A.43}$$

Assume, per contra, that u' is not a weak affine representation for $\succsim$. Thus there exist $\tilde{x}^\dagger \succ \tilde{y}^\dagger$ and $u'(\tilde{y}^\dagger) \geq u'(\tilde{x}^\dagger)$. By Axiom 1.1, $\succsim'$ is non-trivial and so for some $a, b \in \Delta(X)$, $u'(a) > u'(b)$. By Axiom 1.2, there exist $x^\dagger, y^\dagger \in \Delta(X)$ such that $x^\dagger \succ y^\dagger$ and $u'(y^\dagger) > u'(x^\dagger)$.

By Axiom 1.1, there exists $x^\ddagger \in \Delta(X)$ such that $x^* \succ' x^\ddagger$ or $x^\ddagger \succ' x^*$. W.l.o.g. assume that $x^* \succ' x^\ddagger$. By Axiom 1.2, there exists $y^\ddagger \in \Delta(X)$ such that $x^* \succ' y^\ddagger \succ' x^\ddagger$. By Lemma A.34, for all $\alpha, \beta \in [u'(x^*), u'(y^\ddagger)], \alpha > \beta$, we have

$$H'(\alpha) > H'(\beta). \tag{A.44}$$

Let $z^\ddagger \in \Delta(X)$ be such that $u'(x^*) > u'(z^\ddagger) > u'(y^\ddagger)$.

Now there exists $\alpha^0 \in (0,1)$ such that

$$u'(x^*) > u'(\alpha^0 y^\dagger + (1-\alpha^0)z^\ddagger) > u'(\alpha^0 x^\dagger + (1-\alpha^0)z^\ddagger) > u'(y^\ddagger).$$

By (A.44),

$$\alpha^0 y^\dagger + (1-\alpha^0)z^\ddagger \succ' \alpha^0 x^\dagger + (1-\alpha^0)z^\ddagger. \tag{A.45}$$

Next, there exists $\alpha^1 \in (0,1)$ such that

$$u(\hat{x}) > u\big(\alpha^1[\alpha^0 y^\dagger + (1-\alpha^0)z^\ddagger] + (1-\alpha^1)x^*\big) > u(\hat{y}).$$

By (A.43), we have

$$\alpha^1[\alpha^0 x^\dagger + (1-\alpha^0)z^\ddagger] + (1-\alpha^1)x^* \succ \alpha^1[\alpha^0 y^\dagger + (1-\alpha^0)z^\ddagger] + (1-\alpha^1)x^*.$$

By (A.41,A.45),

$$\alpha^1[\alpha^0 y^\dagger + (1-\alpha^0)z^\ddagger] + (1-\alpha^1)x^* \succ' \alpha^1[\alpha^0 x^\dagger + (1-\alpha^0)z^\ddagger] + (1-\alpha^1)x^*$$

which contradicts (A.40).

□

The following axiom summarizes the additional properties for the state independent Boolean expected utility from Remark 1.2.

Axiom A.1 1. $\succsim$ satisfies monotonicity with between states comparisons: For all $s, s' \in S$, $f, g \in H, x, y \in \Delta(X)$,

$$(x_s, f_{-s}) \succ (y_s, f_{-s}) \Longrightarrow (x_{s'}, g_{-s'}) \succsim (y_{s'}, g_{-s'}),$$

2. There exists a constant act $z \in \Delta(X)$ such that z is a crisp act.
3. There exist constant acts $x, y \in \Delta(X)$ such that $x \succ y$ and x and y are equally crisp.

The next lemma is a simple observation that connects crisp and equally crisp acts.

Lemma A.37 Assume that $\succsim$ has a state dependent Boolean expected utility representation with $(\mathbb{P}, u)$. If $f^* \in H$ is crisp and g^*, f^* are equally crisp, then g^* is crisp.

Proof. Let $f, g \in H$ and $\alpha \in (0,1)$. Now we have by the definitions of crisp and equally crisp acts,

$$f \succsim g \iff \alpha f + (1-\alpha)f^* \succsim \alpha g + (1-\alpha)f^* \iff \alpha f + (1-\alpha)g^* \succsim \alpha g + (1-\alpha)g^*.$$

This shows the claim. □

Next, by using the above lemmas, we can show that all the state dependent utilities are positive affine transformations of each other.

Lemma A.38 Assume that $\succsim$ satisfies Axioms 1-1.6 and A.1 and has a state dependent Boolean expected utility representation with $(\mathbb{P}, u)$. Then there exist an affine $u^* : \Delta(X) \to \mathbb{R}$ and $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that for all $s \in S^{\mathcal{P}}$,

$$u_s = a_s u^* + b_s.$$

Proof. First, we observe that by Axioms 1.3 and A.1, if $c \in \Delta(X)$ is a crisp act, then for all $s \in S, x, y \in \Delta(X), \alpha \in (0,1)$,

$$x \succ_s y \Rightarrow \alpha x + (1-\alpha)c \succ_s \alpha y + (1-\alpha)c. \tag{A.46}$$

Additionally, by Axiom A.1, for all $s, s' \in S, x, y \in \Delta(X)$,

$$x \succ_s y \Rightarrow x \succsim_{s'} y. \tag{A.47}$$

Let $z^* \in \Delta(X)$ be the crisp act from Axiom A.1. We consider three cases. 1) $\exists\, s \in S^{\mathcal{P}}, u_s(z^*) = \sup u_s(\Delta(X))$, 2) $\exists\, s \in S^{\mathcal{P}}, u_s(z^*) = \inf u_s(\Delta(X))$, 3) $\forall\, s \in S^{\mathcal{P}}, \sup u_s(\Delta(X)) \neq u_s(z^*) \neq \inf u_s(\Delta(X))$. The case 2) follows symmetrically to 1) so we only show 1).

Assume that there exists $s \in S^{\mathcal{P}}$ such that $u_s(z^*) = \sup u_s(\Delta(X))$. Let $s' \in S^{\mathcal{P}}$. By (A.46,A.47) and Lemma A.35, $u_{s'}$ is a weak affine representation for $\succsim_s$. By Lemmas A.14 and A.15 and Axiom 1.4, $\succsim_s$ satisfies Axioms 1.1, 1.2, 4′. By Proposition A.12, there exist $a_{s'} > 0, b_{s'} \in \mathbb{R}$ such that $u_{s'} = a_{s'} u_s + b_{s'}$. This shows the claim since $s' \in S^{\mathcal{P}}$ was arbitrary.

Next, assume that for all $s \in S^{\mathcal{P}}$, $\sup u_s(\Delta(X)) \neq u_s(z^*) \neq \inf u_s(\Delta(X))$. By Axiom A.1, let $x^*, y^* \in \Delta(X)$ be such that $x^* \succ y^*$ and x^* and y^* are equally crisp. By the case considered, $\mathrm{pr}_{S^{\mathcal{P}}} u(z^*) \in \operatorname{int} \mathrm{pr}_{S^{\mathcal{P}}} u(H)$. Thus there exist $\beta \in (0,1)$ and $h \in H$ such that for all $s \in S^{\mathcal{P}}$,

$$u_s\Big(\beta h_s + (1-\beta)\Big(\frac{1}{2}x^* + \frac{1}{2}y^*\Big)\Big) = u_s(z^*).$$

By Lemma A.48, $\beta h + (1-\beta)x^* \succ \beta h + (1-\beta)y^*$. By Lemmas A.13, A.17, and A.37, $\beta h + (1-\beta)x^*$ and $\beta h + (1-\beta)y^*$ are crisp. By Lemma A.13, there exists $s^* \in S^{\mathcal{P}}$ such that $\beta h_{s^*} + (1-\beta)x^*_{s^*} \succ_{s^*} \beta h_{s^*} + (1-\beta)y^*_{s^*}$. By the crispness,

$$\beta h_{s^*} + (1-\beta)x^*_{s^*} \succ_{s^*} \beta h_{s^*} + (1-\beta)\Big(\frac{1}{2}x^* + \frac{1}{2}y^*\Big) \sim_{s^*} z^* \succ_{s^*} \beta h_{s^*} + (1-\beta)y^*_{s^*}. \tag{A.48}$$

Let $s' \in S^{\mathcal{P}}$. By (A.46,A.47,A.48) and Lemma A.36, $u_{s'}$ is a weak affine representation for $\succsim_{s^*}$. Thus as above, by Proposition A.12, there exist $a_{s'} > 0, b_{s'} \in \mathbb{R}$ such that $u_{s'} = a_{s'} u_{s^*} + b_{s'}$. This shows the claim since $s' \in S^{\mathcal{P}}$ was arbitrary. □

Finally, by putting above together, we have a state independent dual-self expected utility.

Theorem A.39 $\succsim$ satisfies Axioms 1-1.6 and A.1 iff $\succsim$ has a state independent Boolean expected utility representation.

Proof. First, we show the sufficiency of the axioms. Let $z^* \in \Delta(X)$ be a crisp act and $x^*, y^* \in \Delta(X), x^* \succ y^*$ be equally crisp acts from A.1. By Theorem 1.1, there exists a state dependent dual-

self expected utility representation $(u, \mathbb{P})$ for $\succsim$ such that for all $s \in S$, $u_s(z^*) = 0$. By Lemma A.38, there exist an affine $u^* : \Delta(X) \to \mathbb{R}$ and $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that for all $s \in S^{\mathcal{P}}$,

$$u_s = a_s u^* + b_s.$$

Since $x^* \succ y^*$, $u^*(x^*) > u^*(y^*)$. By Corollary A.69 and the proofs of Propositions A.26 and A.27, there exists a state dependent dual-self expected utility $(\tilde{u}, \tilde{\mathbb{P}})$ for $\succsim$ such that for all $s \in S$,

$$\tilde{u}_s(z^*) = 0 \text{ and } \tilde{u}(x^*) - \tilde{u}(y^*) = 1. \tag{A.49}$$

By Lemma A.38, there exist an affine $\tilde{u}^* : \Delta(X) \to \mathbb{R}$ and $\tilde{a} \in \mathbb{R}^S_{++}, \tilde{b} \in \mathbb{R}^S$ such that for all $s \in S^{\mathcal{P}}$,

$$\tilde{u}_s = \tilde{a}_s \tilde{u}^* + \tilde{b}_s.$$

By an affine transformation for u^*, assume w.l.o.g. that $\tilde{u}^*(z^*) = 0$ and $\tilde{u}^*(x^*) - \tilde{u}^*(y^*) = 1$. By (A.49), we have for all $s \in S^{\mathcal{P}}$, $\tilde{a}_s = 1$ and $\tilde{b}_s = 0$. Now we can choose for all $s \in S \setminus S^{\mathcal{P}}$, $u_s = \tilde{u}^*$ without affecting representation by the definition of state dependent dual-self expected utility. This gives a state independent dual-self expected utility.

Next, we show the necessity of the axioms. By Theorem 1.1, we only show Axiom A.1. The properties (2) and (3) follow trivially since all constant acts are crisp and so equally crisp with each other and the preferences are non-trivial. Finally, (1) follows from the state independent utility and statewise monotonicity. □

A.4 Dual-Self Identification

For this section, we assume that $S = S^{\mathcal{P}}$ since by the definition of dual-self and dual-self variational representations probabilities for null states are always identified and the utilities are not. Thus this is without loss of generality.

With an abuse of definitions we define following.

Definition Let $\succsim \subseteq H \times H$. (u, I) is a *state dependent dual-self representation for* $\succsim$ if $u = (u_s)_{s \in S}$ and for all $s \in S$, $u_s : \Delta(X_s) \to \mathbb{R}$ is affine, $I : u(H) \to \mathbb{R}$ is C-additive, positive homogeneous, and monotonic and for all $f, g \in H$,

$$f \succsim g \iff I\big(u(f)\big) \geq I\big(u(g)\big).$$

Definition Let $\succsim \subseteq H \times H$. (u, I) is a *state dependent variational dual-self representation for* $\succsim$ if $u = (u_s)_{s\in S}$ and for all $s \in S$, $u_s : \Delta(X_s) \to \mathbb{R}$ is affine, $I : u(H) \to \mathbb{R}$ is C-additive, and monotonic and for all $f, g \in H$,

$$f \succsim g \iff I\big(u(f)\big) \geq I\big(u(g)\big).$$

Since the state space is assumed to be finite $\mathbb{R}^S$ is metrizable space and hence convergence is characterized by the convergence of sequences.

We start with preliminary results.

A.4.1 Dual-Self Variational Representation Additive Between Equally Crisp Acts

We first show that any dual-self variational representation is additive between any equally crisp acts. The next lemmas show that the set of equally crisp acts is convex.

Lemma A.40 Let f^*, g^* and $\widetilde{f}^*, \widetilde{g}^*$ be equally crisp acts and $\alpha \in [0, 1]$ Then $\alpha f^* + (1-\alpha)\widetilde{f}^*, \alpha g^* + (1-\alpha)\widetilde{g}^*$ are equally crisp acts.

Proof. If $\alpha \in \{0, 1\}$ then the claim follows directly. So assume $\alpha \in (0, 1)$. Let $f, g \in H$, $\beta \in (0, 1)$. Now for $h \in \{f, g\}, h^* \in \{f^*, g^*\}, \widetilde{h}^* \in \{\widetilde{f}^*, \widetilde{g}^*\}$,

$$\alpha h + (1-\alpha)(h^*\beta\widetilde{h}^*) \equiv (\alpha+\beta-\alpha\beta)\Big(\frac{\alpha}{\alpha+\beta-\alpha\beta}h + \frac{\beta-\alpha\beta}{\alpha+\beta-\alpha\beta}h^*\Big) + (1-\alpha-\beta+\alpha\beta)\widetilde{h}^* \quad \text{(A.50)}$$

and

$$\alpha h+(1-\alpha)(h^*\beta\widetilde{h}^*)\equiv(1-\beta+\alpha\beta)\Big(\frac{\alpha}{1-\beta+\alpha\beta}h+\frac{1-\alpha-\beta+\alpha\beta}{1-\beta+\alpha\beta}\widetilde{h}^*\Big)+\big(1-(1-\beta+\alpha\beta)\big)h^*. \quad \text{(A.51)}$$

Thus where equivalencies $(\star)$ follow from the assumptions that f^*, g^* and $\widetilde{f}^*, \widetilde{g}^*$ are equally crisp.

$$\begin{aligned}
&\alpha f + (1-\alpha)(f^*\beta\widetilde{f}^*) \succsim \alpha g + (1-\alpha)(f^*\beta\widetilde{f}^*)\\
&\overset{\text{(A.50)}}{\iff} (\alpha+\beta-\alpha\beta)\Big(\frac{\alpha}{\alpha+\beta-\alpha\beta}f + \frac{\beta-\alpha\beta}{\alpha+\beta-\alpha\beta}f^*\Big) + (1-\alpha-\beta+\alpha\beta)\widetilde{f}^*\\
&\qquad\qquad \succsim (\alpha+\beta-\alpha\beta)\Big(\frac{\alpha}{\alpha+\beta-\alpha\beta}g + \frac{\beta-\alpha\beta}{\alpha+\beta-\alpha\beta}f^*\Big) + (1-\alpha-\beta+\alpha\beta)\widetilde{f}^*\\
&\overset{(\star)}{\iff} (\alpha+\beta-\alpha\beta)\Big(\frac{\alpha}{\alpha+\beta-\alpha\beta}f + \frac{\beta-\alpha\beta}{\alpha+\beta-\alpha\beta}f^*\Big) + (1-\alpha-\beta+\alpha\beta)\widetilde{g}^*\\
&\qquad\qquad \succsim (\alpha+\beta-\alpha\beta)\Big(\frac{\alpha}{\alpha+\beta-\alpha\beta}g + \frac{\beta-\alpha\beta}{\alpha+\beta-\alpha\beta}f^*\Big) + (1-\alpha-\beta+\alpha\beta)\widetilde{g}^*\\
&\overset{\text{(A.50)}}{\iff} \alpha f + (1-\alpha)(f^*\beta\widetilde{g}^*) \succsim \alpha g + (1-\alpha)(f^*\beta\widetilde{g}^*)
\end{aligned}$$

$$\overset{(A.51)}{\iff} (1-\beta+\alpha\beta)\Big(\frac{\alpha}{1-\beta+\alpha\beta}f+\frac{1-\alpha-\beta+\alpha\beta}{1-\beta+\alpha\beta}\widetilde{g}^*\Big)+\big(1-(1-\beta+\alpha\beta)\big)f^*$$
$$\succsim (1-\beta+\alpha\beta)\Big(\frac{\alpha}{1-\beta+\alpha\beta}g+\frac{1-\alpha-\beta+\alpha\beta}{1-\beta+\alpha\beta}\widetilde{g}^*\Big)+\big(1-(1-\beta+\alpha\beta)\big)f^*$$
$$\overset{(\star)}{\iff} (1-\beta+\alpha\beta)\Big(\frac{\alpha}{1-\beta+\alpha\beta}f+\frac{1-\alpha-\beta+\alpha\beta}{1-\beta+\alpha\beta}\widetilde{g}^*\Big)+\big(1-(1-\beta+\alpha\beta)\big)g^*$$
$$\succsim (1-\beta+\alpha\beta)\Big(\frac{\alpha}{1-\beta+\alpha\beta}g+\frac{1-\alpha-\beta+\alpha\beta}{1-\beta+\alpha\beta}\widetilde{g}^*\Big)+\big(1-(1-\beta+\alpha\beta)\big)g^*$$
$$\overset{(A.51)}{\iff} \alpha f+(1-\alpha)(g^*\beta\widetilde{g}^*)\succsim \alpha g+(1-\alpha)(g^*\beta\widetilde{g}^*).$$

□

Corollary A.41 Let f^*, g^* be such that for all $f, g \in H$ and $\alpha \in (0,1)$,

$$\alpha f+(1-\alpha)f^*\succsim \alpha g+(1-\alpha)f^* \iff \alpha f+(1-\alpha)g^*\succsim \alpha g+(1-\alpha)g^*.$$

Then for all $\beta \in (0,1)$,

$$\alpha f+(1-\alpha)f^*\succsim \alpha g+(1-\alpha)f^* \iff \alpha f+(1-\alpha)(f^*\beta g^*)\succsim \alpha g+(1-\alpha)(f^*\beta g^*).$$

Proof. Now (f^*, g^*) and (f^*, f^*) are equally crisp. Thus for $\beta \in (0,1)$, by Lemma A.40, f^* and $\beta f^*+(1-\beta)g^*$ are equally crisp which shows the claim. □

Corollary A.42 Let f^*, g^* be equally crisp acts and $\alpha, \beta \in [0,1]$ Then $\alpha f^*+(1-\alpha)g^*, \beta f^*+(1-\beta)g^*$ are equally crisp acts.

Proof. Assume w.l.o.g. $\alpha \geq \beta$. Now (f^*, g^*) and (g^*, g^*) are equally crisp by Lemma A.40, $(g^*, \frac{\alpha-\beta}{1-\beta}f^*+\big(1-\frac{\alpha-\beta}{1-\beta}\big)g^*)$ are equally crisp. Now since (f^*, f^*) and $(g^*, \frac{\alpha-\beta}{1-\beta}f^*+\big(1-\frac{\alpha-\beta}{1-\beta}\big)g^*)$ are equally crisp, by Lemma A.40, $(\beta f^*+(1-\beta)g^*, \alpha f^*+(1-\alpha)g^*)$ are equally crisp since

$$\beta f^*+(1-\beta)\Big(\frac{\alpha-\beta}{1-\beta}f^*+\big(1-\frac{\alpha-\beta}{1-\beta}\big)g^*\Big)\equiv \alpha f^*+(1-\alpha)g^*.$$

□

Corollary A.43 Let f^*, g^* be equally crisp acts and $f \in H, \beta \in (0,1)$ Then $\beta f+(1-\beta)f^*, \beta f+(1-\beta)g^*$ are equally crisp acts.

Proof. Now (f^*, g^*) and (f, f) are equally crisp. Thus for $\beta \in (0,1)$, by Lemma A.40, $\beta f+(1-\beta)f^*, \beta f+(1-\beta)g^*$ are equally crisp which shows the claim. □

The next two lemmas show that the dual-self variational representation is affine between equally crisp acts.

Lemma A.44 Let (u, I) be a dual-self variational representation for $\succsim$ and f^*, g^* equally crisp acts. Let $f \in H, \beta \in (0,1)$ be such that $f \in \operatorname{int} u(H)$. Then the mapping $(0,1) \ni \alpha \overset{J}{\mapsto} I \circ u\big(\beta f + (1-\beta)(f^* \alpha g^*)\big)$ has locally constant derivative at the differentiability points.

Proof. Let Ω be the differentiability points of the mapping J. Assume, per contra, that there exists a point $x \in (0,1)$ such that x does not have a neighborhood U such that for all $y, z \in U \cap \Omega$, $\nabla J(y) = \nabla J(z)$. First, since $u(f) \in \operatorname{int} u(H)$, there exists $\varepsilon > 0$ such that $B_\infty(u(f), \varepsilon) \subseteq u(H)$. Second, since J is Lipschitz, by the counter assumption, there exist $y, z \in (0,1)$ such that $|y - z| < \frac{\beta\varepsilon}{\|u(f^*) - u(g^*)\|_\infty}$ and $|x - y|, |x - z| < \frac{\min\{x, 1-x\}}{3}$, $\nabla J(y) > \nabla J(z)$. Since I is C-additive and hence especially 1-Lipschitz, we have $1/\beta |J(y) - J(z)| \leq 1/\beta \|u(f^*) - u(g^*)\|_\infty |y - z| < \varepsilon$. Thus there exists $g \in H$ such that $u(g) = u(f) + 1/\beta \overline{1}(J(z) - J(y))$. By C-additivity, we have since $1/2(1+\beta) = \beta + (1-\beta)/2$ and $1 - 1/2(1+\beta) = 1/2(1-\beta)$,

$$
\begin{aligned}
&I\Big(1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)} u(g) + \frac{(1-\beta)/2}{1/2(1+\beta)} u(f^* y g^*)\Big) + \big(1 - 1/2(1+\beta)\big) u(f^* y g^*)\Big) \\
&= I\Big(\beta u(f) + 1/2(1-\beta) u(f^* y g^*) + 1/2(1-\beta) u(f^* y g^*) + \beta 1/\beta \overline{1} \times (J(z) - J(y)))\Big) \\
&\overset{\text{C-add.}}{=} J(y) + J(z) - J(y) = J(z)
\end{aligned}
$$

Additionally since $1 > 2z - y > 0$, we have

$$
\begin{aligned}
&I\Big(1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)} u(f) + \frac{(1-\beta)/2}{1/2(1+\beta)} u\big(f^*(2z-y) g^*\big)\Big) + \big(1 - 1/2(1+\beta)\big) u(f^* y g^*)\Big) \\
&= I\Big(\beta u(f) + 1/2(1-\beta) u\big(f^*(2z-y) g^*\big) + 1/2(1-\beta) u(f^* y g^*)\Big) = J(z).
\end{aligned}
$$

Thus by the representation,

$$
\begin{aligned}
&1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)} g + \frac{(1-\beta)/2}{1/2(1+\beta)} (f^* y g^*)\Big) + \big(1 - 1/2(1+\beta)\big)(f^* y g^*) \qquad \text{(A.52)} \\
&\sim 1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)} f + \frac{(1-\beta)/2}{1/2(1+\beta)} \big(f^*(2z-y) g^*\big)\Big) + \big(1 - 1/2(1+\beta)\big)(f^* y g^*).
\end{aligned}
$$

But since $\nabla J(y) > \nabla J(z)$, there exists $0 < \delta < \frac{1-x}{3}$ such that

$$
\frac{J(y+\delta) - J(y)}{\delta} > \frac{J(z+\delta) - J(z)}{\delta} \Rightarrow J(y+\delta) + J(z) - J(y) > J(z+\delta). \qquad \text{(A.53)}
$$

Thus since $y + 2\delta < 1$

$$
\begin{aligned}
&I\Big(1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)} u(g) + \frac{(1-\beta)/2}{1/2(1+\beta)} u(f^* y g^*)\Big) + \big(1 - 1/2(1+\beta)\big) u\big(f^*(y+2\delta) g^*\big)\Big) \\
&= I\Big(\beta u(f) + 1/2(1-\beta) u(f^* y g^*) + 1/2(1-\beta) u\big(f^*(y+2\delta) g^*\big) + \beta 1/\beta \overline{1} \times (J(z) - J(y))\Big) \\
&\overset{\text{C-add.}}{=} J(y+\delta) + J(z) - J(y) \overset{\text{(A.53)}}{>} J(z+\delta) \\
&= I\Big(\beta u(f) + 1/2(1-\beta) u\big(f^*(2z-y) g^*\big) + 1/2(1-\beta) u\big(f^*(y+2\delta) g^*\big)\Big)
\end{aligned}
$$

$$= I\Big(1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)}u(f)+\frac{(1-\beta)/2}{1/2(1+\beta)}u\big(f^*(2z-y)g^*\big)\Big)+\big(1-1/2(1+\beta)\big)u\big(f^*(y+2\delta)g^*\big)\Big).$$

Thus by the representation,

$$1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)}g+\frac{(1-\beta)/2}{1/2(1+\beta)}(f^*yg^*)\Big)+\big(1-1/2(1+\beta)\big)\big(f^*(y+2\delta)g^*\big)$$
$$\succ 1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)}f+\frac{(1-\beta)/2}{1/2(1+\beta)}\big(f^*(2z-y)g^*\big)\Big)+\big(1-1/2(1+\beta)\big)\big(f^*(y+2\delta)g^*\big).$$

but this contradicts Corollary A.41 and (A.52). □

Lemma A.45 Let (u, I) be a variational representation for $\succsim$ and f^*, g^* equally crisp acts. Let $f \in H, \beta \in (0,1)$ be such that $f \in \operatorname{int} u(H)$. Then the mapping $[0,1] \ni \alpha \overset{J}{\mapsto} I \circ u\big(\beta f + (1-\beta)(f^*\alpha g^*)\big)$ is affine.

Proof. Since I is C-additive and hence 1-Lipschitz, especially J is Lipschitz and it is differentiable almost everywhere. Let the set of differentiability points be Ω. Define

$$\widetilde{J}(\alpha) = \begin{cases} \nabla J(\alpha), & \text{if } J \text{ is differentiable at } \alpha \\ \liminf\limits_{\widetilde{\alpha}\to\alpha, \widetilde{\alpha}\in\Omega} \nabla J(\widetilde{\alpha}), & \text{if } J \text{ is not differentiable at } \alpha \end{cases}$$

Since J is differentiable almost everywhere $\widetilde{J}$ is well-defined. Let $0 < \varepsilon < 1/2$. We show that $\widetilde{J}$ is constant in the set $(0,1)$. By Lemma A.44, for all $\alpha \in (0,1)$, there exists $\varepsilon_\alpha > 0, c \in \mathbb{R}$ such that for all $x \in B_\infty(\alpha, \varepsilon_\alpha) \cap \Omega$, $\nabla J(x) = c$. Thus by the definition of $\widetilde{J}$, for all $x \in B_\infty(\alpha, \varepsilon_\alpha)$, $\widetilde{J}(x) = c$. Thus $\widetilde{J}$ is locally constant function in the set $(0,1)$. Since the interval $(0,1)$ is connected, by (Viro et al., 2008, Problem 12.2x) J' is constant on the set $(0,1)$. Thus there exists $c \in \mathbb{R}$ such that for all $x \in (0,1) \cap \Omega$, $\nabla J(x) = c$. Since J is Lipschitz, we have for all $x \in [0,1]$,

$$J(x) - J(0) = \int_{[0,x]\cap\Omega} \nabla J(y)\, dy = \int_{(0,x)\cap\Omega} \nabla J(y)\, dy = \int_{(0,x)\cap\Omega} c\, dy = cx.$$

Thus especially, the mapping J is affine. □

The next two lemmas show that the directional derivative of the representation to the direction $u(g^*) - u(f^*)$ is constant always. First, we show that it is constant on each indifference level.

Lemma A.46 Let (u, I) be a dual-self variational representation for $\succsim$ and f^*, g^* equally crisp acts and denote $h^* \coloneqq f^* 1/2 g^*$. Let $f, g \in H, \beta, \alpha \in (0,1)$ be such that $f, g \in \operatorname{int} u(H)$ and $\alpha f + (1-\alpha)h^* \sim \beta g + (1-\beta)h^*$. Then the affine mappings $[0,1] \ni \gamma \overset{J^f}{\longmapsto} I \circ u\big(\alpha f + (1-\alpha)(f^*\gamma g^*)\big)$ and $[0,1] \ni \gamma \overset{J^g}{\longmapsto} I \circ u\big(\beta g + (1-\beta)(f^*\gamma g^*)\big)$ with derivatives ∇J^f and ∇J^g respectively has $\nabla J^f = \frac{1-\alpha}{1-\beta}\nabla J^g$.

Proof. Assume w.l.o.g. $\alpha \geq \beta$. Then

$$\alpha\Big(\frac{\beta}{\alpha}g + \frac{\alpha-\beta}{\alpha}h^*\Big) + (1-\alpha)h^* \equiv \beta g + (1-\beta)h^*.$$

Define the mapping

$$[0,1] \ni \gamma \overset{J^{gh^*}}{\longmapsto} I \circ u\Big(\alpha\Big(\frac{\beta}{\alpha}g + \frac{\alpha-\beta}{\alpha}h^*\Big) + (1-\alpha)(f^*\gamma g^*)\Big)$$

Thus since f^* and g^* are equally crisp acts and by Corollary A.41, we have by the representation, for all $\gamma \in [0,1]$,

$$\begin{aligned} J^{gh^*}(\gamma) &= I \circ u\Big(\alpha\Big(\frac{\beta}{\alpha}g + \frac{\alpha-\beta}{\alpha}h^*\Big) + (1-\alpha)(f^*\gamma g^*)\Big) \\ &= I \circ u\big(\alpha f + (1-\alpha)(f^*\gamma g^*)\big) = J^f(\gamma). \end{aligned}$$

Thus especially J^{gh^*} is affine with derivative $\nabla J^{gh^*} = \nabla J^f$. But on the other hand for all $\gamma \in [0,1]$, we have

$$\alpha\Big(\frac{\beta}{\alpha}g + \frac{\alpha-\beta}{\alpha}h^*\Big) + (1-\alpha)(f^*\gamma g^*) \equiv \beta g + (1-\beta)\Big(f^*\frac{(\alpha-\beta)/2 + \gamma(1-\alpha)}{1-\beta}g^*\Big)$$

and so

$$\begin{aligned} &J^{gh^*}(\gamma) = I \circ u\Big(\alpha\Big(\frac{\beta}{\alpha}g + \frac{\alpha-\beta}{\alpha}h^*\Big) + (1-\alpha)(f^*\gamma g^*)\Big) \\ &= I \circ u\Big(\beta g + (1-\beta)\Big(f^*\frac{(\alpha-\beta)/2 + \gamma(1-\alpha)}{1-\beta}g^*\Big)\Big) = J^g\Big(\frac{(\alpha-\beta)/2 + \gamma(1-\alpha)}{1-\beta}\Big). \end{aligned}$$

Thus especially by the chain rule of derivatives, $\nabla J^{gh^*} = \nabla J^g \frac{1-\alpha}{1-\beta}$ which shows the claim. □

The next lemma shows that the directional derivative is constant at different indifference levels.

Lemma A.47 Let (u, I) be a state dependent dual-self variational representation for $\succsim$ and f^*, g^* equally crisp acts and denote $h^* := f^*1/2g^*$. For all $f, \beta \in H$, define the mapping $[0,1] \ni \gamma \overset{J^f_\beta}{\longmapsto} I \circ u\big(\beta f + (1-\beta)(f^*\gamma g^*)\big)$ with derivative ∇J^f_β. Define the mappings

$$K^* : (\inf I(u(H)), \sup I(u(H))) \to \mathbb{R} \cup \{\infty\} \text{ and } K_* : (\inf I(u(H)), \sup I(u(H))) \to \mathbb{R} \cup \{\infty\}$$

by for all $v \in (\inf I(u(H)), \sup I(u(H)))$,

$$K^*(v) = \sup\Big\{\frac{1}{1-\beta}\nabla J^f_\beta \Big| f \in H, \beta \in (0,1), u(f) \in \operatorname{int} u(H), I \circ u\big(\beta f + (1-\beta)h^*\big) = v\Big\}$$

and

$$K_*(v) = \inf\Big\{\frac{1}{1-\beta}\nabla J^f_\beta \Big| f \in H, \beta \in (0,1), u(f) \in \operatorname{int} u(H), I \circ u\big(\beta f + (1-\beta)h^*\big) = v\Big\}$$

where $\sup \varnothing = \inf \varnothing = \infty$. Then K^* is constant and finite and $K^* = K_*$.

Proof. We show first that locally K^* is constant and finite and $K^* = K_*$. Let $v \in (\inf I(u(H)), \sup I(u(H)))$. First, there exist $f^1, f^2 \in H$ such that $I(u(f^1)) > v > I(u(f^2))$. Since S is finite, there exists $f^3 \in H$ such that $u(f^3) \in \operatorname{int} u(H)$. By mixture continuity, there exist $\alpha^1, \alpha^2, \alpha^3 \in (0,1)$ such that $\alpha^1 + \alpha^2 + \alpha^3 = 1$ and $I \circ u(\alpha^1 f^1 + \alpha^2 f^2 + \alpha^3 f^3) = v$. Denote

$f^4 := \alpha^1 f^1 + \alpha^2 f^2 + \alpha^3 f^3$. Since $\alpha^3 > 0$ and $u(f^3) \in \operatorname{int} u(H)$ and $u(H)$ is a convex set $u(f^4) \in \operatorname{int} u(H)$. Thus there exists $\varepsilon > 0$ such that $B_\infty(u(f^4), \varepsilon) \subseteq u(H)$. Let $\beta \in (1/2, 1)$ be such that

$$(\frac{1}{\beta} - 1)\left\|u(f^4)\right\|_\infty + \frac{1-\beta}{\beta}\|u(h^*)\|_\infty < \varepsilon.$$

Thus by the choice of β, there exists $f \in H$ such that

$$u(f) = u(f^4) + (\frac{1}{\beta} - 1)u(f^4) - \frac{1-\beta}{\beta}u(h^*).$$

Thus especially $u(f) \in \operatorname{int} u(H)$ and

$$\beta u(f) + (1-\beta)u(h^*) = u(f^4)$$

and so $I\big(\beta u(f) + (1-\beta)u(h^*)\big) = v$. First, for all $\widetilde{f} \in H, \widetilde{\beta} \in (0,1)$ such that $u(\widetilde{f}) \in \operatorname{int} u(H), I \circ u\big(\widetilde{\beta}\widetilde{f} + (1-\widetilde{\beta})h^*\big) = v$, we have $\widetilde{\beta}\widetilde{f} + (1-\widetilde{\beta})h^* \sim (1-\beta)f + (1-\beta)h^*$ and by Lemma A.46,

$$\frac{1}{1-\widetilde{\beta}}\nabla J_{\widetilde{\beta}}^{\widetilde{f}} = \frac{1}{(1-\beta)}\nabla J_\beta^f \tag{A.54}$$

Thus since $\widetilde{f} \in H, \widetilde{\beta} \in (0,1)$ were arbitrary such that $u(\widetilde{f}) \in \operatorname{int} u(H), I \circ u\big(\widetilde{\beta}\widetilde{f} + (1-\widetilde{\beta})h^*\big) = \widetilde{v}$,

$$K^*(v) \overset{(A.54)}{=} \frac{1}{(1-\beta)}\nabla J_\beta^f \overset{(A.54)}{=} K_*(v). \tag{A.55}$$

We consider two cases: 1) there exists $\alpha^0 \in (0,1)\backslash\{1/2\}$ such that $\beta f + (1-\beta)h^* \sim \beta f + (1-\beta)f^*\alpha^0 g^*$. Now $J_\beta^f(1/2) = J_\beta^f(\alpha^0)$ and by Lemma A.45, J_β^f is affine and hence for all $\alpha \in (0,1)$,

$$J_\beta^f(\alpha) = J_\beta^f(1/2). \tag{A.56}$$

Thus $K^*(v) = 0$. We show that there exists a neighborhood A of v such that for all $\widetilde{v} \in A$, $K^*(\widetilde{v}) = 0 = K_*(\widetilde{v})$. Assume, per contra, that such a neighborhood A of v does not exist. Since $u(f) \in \operatorname{int} u(H)$, there exists $\varepsilon > 0$ such that $B_\infty(u(f), \varepsilon) \subseteq u(H)$. Now there exists $\widetilde{v} \in B(v, 1/2\beta\varepsilon)$ such that $K^*(\widetilde{v}) \neq 0$ or $K_*(\widetilde{v}) \neq 0$. Thus there exists $\widetilde{f} \in H$, $\widetilde{\beta} \in (0,1)$ such that $u(\widetilde{f}) \in \operatorname{int} u(H)$, $I \circ u(\beta\widetilde{f} + (1-\beta)h^*) = \widetilde{v}$, and $\nabla J_{\widetilde{\beta}}^{\widetilde{f}} \neq 0$. Now $u(f) + (\widetilde{v} - v) \in u(H)$ and hence there exists $f^5 \in H$ such that $u(f^5) = u(f) + 1/\beta(\widetilde{v} - v)\bar{1}$. Thus $u(\beta f^5 + (1-\beta)h^*) = u(\beta f + (1-\beta)h^*) + (\widetilde{v} - v)\bar{1}$ and so by C-additivity, $I \circ u(\beta f^5 + (1-\beta)h^*) = \widetilde{v}$. By Lemma A.46, $0 \neq \nabla J_\beta^{\widetilde{f}} = \nabla J_\beta^{f^5}$ and thus especially there exists $\alpha^1 \in (0,1)$ such that

$$I \circ u(\beta f^5 + (1-\beta)(f^*\alpha^1 g^*)) \neq \widetilde{v} = I \circ u(\beta f^5 + (1-\beta)h^*).$$

However, then by C-additivity,

$$\begin{aligned}
&I\big(u(\beta f + (1-\beta)(f^*\alpha^1 g^*))\big) = I\big(u(\beta f + (1-\beta)(f^*\alpha^1 g^*))\big) + (\widetilde{v} - v) - (\widetilde{v} - v) \\
&\overset{\text{C-add.}}{=} I\big(u(\beta f + (1-\beta)(f^*\alpha^1 g^*)) + \bar{1}\beta 1/\beta(\widetilde{v} - v)\big) - (\widetilde{v} - v) \\
&= I\big(u(\beta f^5 + (1-\beta)(f^*\alpha^1 g^*))\big) - (\widetilde{v} - v) \neq I\big(u(\beta f^5 + (1-\beta)h^*)\big) - (\widetilde{v} - v)
\end{aligned}$$

$$= I\big(u(\beta f + (1-\beta)h^*) + \bar{1}\beta 1/\beta(\widetilde{v} - v)\big) - (\widetilde{v} - v) \overset{\text{C-add.}}{=} I\big(u(\beta f + (1-\beta)h^*)\big).$$

This contradicts (A.56). Thus there exists a neighborhood A of v such that for all $\widetilde{v} \in A$, $K^*(\widetilde{v}) = K^*(v) = K_*(\widetilde{v})$.

2) There exists $\alpha^0 \in (0,1)$ such that $\beta f + (1-\beta)h^* \not\sim \beta f + (1-\beta)f^*\alpha^0 g^*$. Thus $J^f_\beta(\alpha^0) \neq J^f_\beta(1/2)$. Since J^f_β is affine, we have

$$v^* \coloneqq \max\{J^f_\beta(1/4), J^f_\beta(3/4)\} > J^f_\beta(1/2) > \min\{J^f_\beta(1/4), J^f_\beta(3/4)\} =: v_*.$$

Assume w.l.o.g. $v^* = I\big(\beta u(f) + (1-\beta)u(f^* 3/4 g^*)\big)$ and $v_* = I\big(\beta u(f) + (1-\beta)u(f^* 1/4 g^*)\big)$ since the other case follows symmetrically. Let $\widetilde{v} \in [v_*, v^*]$. Then by Lemma A.45,

$$I\Big(\beta u(f) + (1-\beta)u\Big[f^*\Big(1/4 + 1/2\frac{\widetilde{v} - v_*}{v^* - v_*}\Big)g^*\Big]\Big) = \widetilde{v}$$

Now since $1/2(1+\beta) = \beta + (1-\beta)/2$ and $1 - 1/2(1+\beta) = 1/2(1-\beta)$

$$\begin{aligned}
&1/2(1+\beta)\Big(\frac{\beta}{1/2(1+\beta)}f + \frac{1/2(1-\beta)}{1/2(1+\beta)}\frac{\widetilde{v} - v_*}{v^* - v_*}f^* + \frac{1/2(1-\beta)}{1/2(1+\beta)}\frac{v^* - \widetilde{v}}{v^* - v_*}g^*\Big) + \big(1 - 1/2(1+\beta)\big)h^* \\
&\equiv \beta f + (1-\beta)f^*\Big(1/4 + 1/2\frac{\widetilde{v} - v_*}{v^* - v_*}\Big)g^*.
\end{aligned}$$

Denote

$$f^6 \coloneqq \frac{\beta}{1/2(1+\beta)}f + \frac{1/2(1-\beta)}{1/2(1+\beta)}\frac{\widetilde{v} - v_*}{v^* - v_*}f^* + \frac{1/2(1-\beta)}{1/2(1+\beta)}\frac{v^* - \widetilde{v}}{v^* - v_*}g^*.$$

Then we have for all $\gamma \in (0,1)$, as above

$$1/2(1+\beta)f^6 + \big(1 - 1/2(1+\beta)\big)(f^*\gamma g^*) \equiv \beta f + (1-\beta)f^*\Big(\frac{\widetilde{v} - v_*}{2(v^* - v_*)} + \frac{\gamma}{2}\Big)g^*.$$

Thus for all $\gamma \in (0,1)$,

$$\begin{aligned}
J^{f^6}_{1/2(1+\beta)}(\gamma) &= I \circ u\Big(1/2(1+\beta)f^6 + \big(1 - 1/2(1+\beta)\big)(f^*\gamma g^*)\Big) \\
&= I \circ u\Big(\beta f + (1-\beta)f^*\Big(\frac{\widetilde{v} - v_*}{2(v^* - v_*)} + \frac{\gamma}{2}\Big)g^*\Big) = J^f_\beta\Big(\frac{\widetilde{v} - v_*}{2(v^* - v_*)} + \frac{\gamma}{2}\Big).
\end{aligned}$$

Thus by the chain rule of derivatives and by the affinity of J functions from Lemma A.45, we have

$$\nabla J^{f^6}_{1/2(1+\beta)} = 1/2 \nabla J^f_\beta$$

and thus

$$\frac{1}{1 - 1/2(1+\beta)}\nabla J^{f^6}_{1/2(1+\beta)} = \frac{2}{(1-\beta)}\nabla J^{f^6}_{1/2(1+\beta)} = \frac{1}{(1-\beta)}\nabla J^f_\beta. \tag{A.57}$$

And finally for all $\widetilde{f} \in H, \widetilde{\beta} \in (0,1)$ such that $u(\widetilde{f}) \in \operatorname{int} u(H), I \circ u\big(\widetilde{\beta}\widetilde{f} + (1-\widetilde{\beta})h^*\big) = \widetilde{v}$, we have $\widetilde{\beta}\widetilde{f} + (1-\widetilde{\beta})h^* \sim 1/2(1+\beta)f^6 + \big(1 - 1/2(1+\beta)\big)h^*$ and by Lemma A.46,

$$\frac{1}{1-\widetilde{\beta}}\nabla J^{\widetilde{f}}_{\widetilde{\beta}} = \frac{1}{1 - 1/2(1+\beta)}\nabla J^{f^6}_{1/2(1+\beta)} \overset{\text{(A.57)}}{=} \frac{1}{(1-\beta)}\nabla J^f_\beta \overset{\text{(A.55)}}{=} K^*(v) = K_*(v) \tag{A.58}$$

Thus since $\widetilde{f} \in H, \widetilde{\beta} \in (0,1)$ were arbitrary such that $u(\widetilde{f}) \in \operatorname{int} u(H), I \circ u\big(\widetilde{\beta}\widetilde{f} + (1-\widetilde{\beta})h^*\big) = \widetilde{v}$,

$$K^*(\widetilde{v}) \overset{\text{(A.58)}}{=} K^*(v) = K_*(v) \overset{\text{(A.58)}}{=} K_*(\widetilde{v}).$$

Thus there exists a neighborhood of v such that $K^* = K_*$ and K^* is constant and finite. Since $v \in (\inf I(u(H)), \sup I(u(H)))$ was arbitrary, locally $K^* = K_*$ and K^* is constant and finite. Since the interval $(\inf I(u(H)), \sup I(u(H)))$ is connected, by (Viro et al., 2008, Problem 12.2x) K^* and K_* are constants and finite on the set $(\inf I(u(H)), \sup I(u(H)))$ and $K^* = K_*$. □

Finally, the next lemma shows that the constant directional derivative gives additivity of the representation between equally crisp acts.

Lemma A.48 Let (u, I) be a state dependent dual-self variational representation for $\succsim$. f^*, g^* are equally crisp acts iff for all $\varphi \in u(H), \alpha \in (0,1]$ such that $\varphi + \alpha(u(f^*) - u(g^*)) \in u(H)$,

$$I\big(\varphi + \alpha(u(f^*) - u(g^*))\big) = I(\varphi) + \alpha(I(u(f^*)) - I(u(g^*))).$$

Proof. Let first f^*, g^* be equally crisp acts. Let $\varphi \in \operatorname{int} u(H), \alpha \in \mathbb{R}_{++}$ such that $\varphi + \alpha(u(f^*) - u(g^*)) \in \operatorname{int} u(H)$. Define the mapping $[0,\alpha] \ni \beta \overset{J}{\mapsto} I\big(\varphi + \beta(u(f^*) - u(g^*))\big)$. Now let $\beta \in (0,\alpha)$. Then $\varphi + \beta(u(f^*) - u(g^*)) \in \operatorname{int} u(H)$. Thus there exists $\varepsilon > 0$ such that $B_\infty(\varphi + \beta(u(f^*) - u(g^*)), \varepsilon) \subseteq u(H)$. Let $\gamma \in (1/2, 1)$ be such that

$$\big(\frac{1}{\gamma} - 1\big)\|\varphi + \beta(u(f^*) - u(g^*))\|_\infty + \frac{1-\gamma}{\gamma}\|u(h^*)\|_\infty < \varepsilon.$$

Thus by the choice of γ, there exists $f \in H$ such that

$$u(f) = \varphi + \beta(u(f^*) - u(g^*)) + \big(\frac{1}{\gamma} - 1\big)\varphi + \beta(u(f^*) - u(g^*)) - \frac{1-\gamma}{\gamma}u(h^*).$$

Thus $u(\gamma f + (1-\gamma)h^*) = \varphi + \beta(u(f^*) - u(g^*))$. Now for all $\zeta \in (0,1)$ such that $\beta + (1-\gamma)(\zeta - 1/2) \geq 0$,

$$I\big(u(\gamma f + (1-\gamma)(f^* \zeta g^*))\big) = J(\beta + (1-\gamma)(\zeta - 1/2)). \tag{A.59}$$

Denote $\zeta_* := \max\{0, 1/2 - \frac{\beta}{1-\gamma}\}$ and $\zeta^* := \min\{1, 1/2 + \frac{\alpha-\beta}{1-\gamma}\}$. Define $[\zeta_*, \zeta^*] \ni \zeta \overset{J^\beta}{\longmapsto} J(\beta + (1-\gamma)(\zeta - 1/2))$ where $\zeta^* > 1/2 > \zeta_*$. Thus by Lemma A.47, J^β is affine and $\frac{\nabla J^\beta}{1-\gamma} = I\big(u(f^*)\big) - I\big(u(g^*)\big)$. Thus by the chain rule and (A.59),

$$\nabla J(\beta) = \frac{\nabla J^\beta(1/2)}{1-\gamma} = I\big(u(f^*)\big) - I\big(u(g^*)\big). \tag{A.60}$$

Thus finally since I is 1-Lipschitz, we have

$$I\big(\varphi + \alpha(u(f^*) - u(g^*))\big) - I(\varphi) = \int_{(0,\alpha)} \nabla J(\beta)\, d\beta \overset{\text{(A.60)}}{=} \alpha\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big).$$

Thus we have for all $\varphi \in \operatorname{int} u(H), \alpha \in (0,1]$ such that $\varphi + \alpha(u(f^*) - u(g^*)) \in \operatorname{int} u(H)$,

$$I\big(\varphi + \alpha(u(f^*) - u(g^*))\big) = I(\varphi) + \alpha(I(u(f^*)) - I(u(g^*))).$$

Finally, since the mapping I is continuous, we have for all $\varphi \in u(H), \alpha \in (0,1]$ such that $\varphi + \alpha(u(f^*) - u(g^*)) \in u(H)$,

$$I\big(\varphi + \alpha(u(f^*) - u(g^*))\big) = I(\varphi) + \alpha(I(u(f^*)) - I(u(g^*))).$$

Next, assume that $f^*, g^* \in H$ are such that for all $\varphi \in u(H), \alpha \in (0,1]$ such that $\varphi + \alpha(u(f^*) - u(g^*)) \in u(H)$,

$$I\big(\varphi + \alpha(u(f^*) - u(g^*))\big) = I(\varphi) + \alpha(I(u(f^*)) - I(u(g^*))). \tag{A.61}$$

We show that f^* and g^* are equally crisp. Let $f, g \in H, \alpha \in (0,1)$. Then we have for $h \in \{f, g\}$,

$$I\big(\alpha u(h) + (1-\alpha)g^*\big) \overset{(A.61)}{=} I\big(\alpha u(h) + (1-\alpha)f^*\big) - I\big(u(f^*)\big) + I\big(u(g^*)\big).$$

Thus

$$I\big(\alpha u(f) + (1-\alpha)g^*\big) \geq I\big(\alpha u(g) + (1-\alpha)g^*\big) \Leftrightarrow I\big(\alpha u(f) + (1-\alpha)f^*\big) \geq I\big(\alpha u(g) + (1-\alpha)f^*\big),$$

which proves the claim since $I \circ u$ represents $\succsim$. □

A.4.2 Behavioral Directional Derivatives

By the previous part, dual-self variational representation is additive between equally crisp acts. Next we show that this allows us to capture the derivatives of the representation at differentiability points. These derivatives will give us the probabilities of the decision maker. First, we define how to capture behaviorally directional derivatives. The next definition gives the level of change of preferences at f to the direction of g at mixture level α using the β mixture of equally crisp acts as the measuring tool. Due to non-convex preferences, we focus on the smallest amount of change with C. We keep the crisp acts f^*, g^* such that for all $s \in S^{\mathcal{P}}$, $f_s^* \succ_s g_s^*$ as fixed and denote $h^* \coloneqq 1/2 f^* + 1/2 g^*$.

Definition For $f, g \in H, \beta \in (0,1)$, define for all $\alpha \in (0,1)$,

$$A_{f,\beta}^{g,\alpha} \coloneqq \Big\{\widehat{\alpha} \in [-1/2\alpha, 1/2\alpha] \,|\, \beta f + (1-\beta)\big((\alpha\widehat{\alpha} + 1/2)f^* + (1/2 - \alpha\widehat{\alpha})g^*\big) \sim \alpha g + (1-\alpha)\big(\beta f + (1-\beta)h^*\big)\Big\}$$

and

$$C_{f,\beta}^{g,\alpha} = \begin{cases} \arg\min\{|\hat{\alpha}| \,|\, \hat{\alpha} \in A_{f,\beta}^{g,\alpha}\}, & \text{if } A_{f,\beta}^{g,\alpha} \neq \varnothing \\ \infty, & \text{if } A_{f,\beta}^{g,\alpha} = \varnothing \end{cases}$$

where $\sup \varnothing = \infty = \inf \varnothing$.

The next lemmas show that $\lim_{\alpha\to 0} C^{g,\alpha}_{f,\beta}$ captures the directional derivative of the representation at f to the direction g.

The first lemma shows that for a small enough α, $C^{g,\alpha}_{f,\beta}$ exists and the argmin in it is well-defined. First, we observe that the argmin is always at most a singleton by the affinity of the representation between equally crisp acts, Lemma A.48.

Lemma A.49 Let $\succsim$ have a state dependent dual-self variational representation. Let $f, g \in H$ and $\beta \in (0,1)$. Let $(\alpha_i)_{i=1}^{\infty}$ and $\lim_{i\to\infty} \alpha_i = 0$. Then there exists i^1 such that for all $i > i^1$, $C^{g,\alpha_i}_{f,\beta} \in \mathbb{R}$.

Proof. Assume that $\alpha^0 \in (0, 1/2]$ since the other case follows symmetrically. First, there exists $i^1 \in \mathbb{N}$ such that for all $i > i^1$, $C^{g,\alpha_i}_{f,\beta} \in \mathbb{R}$. Since the sets

$$\Big\{\alpha \in [0,1]\Big|\alpha g + (1-\alpha)\big(\beta f + (1-\beta)h^*\big) \succ \beta f + (1-\beta)g^*\Big\} \text{ and}$$
$$\Big\{\alpha \in [0,1]\Big|\beta f + (1-\beta)f^* \succ (1-\alpha)\big(\beta f + (1-\beta)h^*\big)\Big\}$$

are open by Axiom 1.2 and include 0, there exists $\varepsilon > 0$ such that for all $\alpha \in (0,\varepsilon)$,

$$\beta f + (1-\beta)f^* \succ \alpha g + (1-\alpha)\big(\beta f + (1-\beta)h^*\big) \succ \beta f + (1-\beta)g^*.$$

Thus especially for all $\alpha \in (0,\varepsilon)$, $C^{g,\alpha_i}_{f,\beta} \in \mathbb{R}$. Since $\lim_{i\to\infty} \alpha_i = 0$, there exists $i^1 \in \mathbb{N}$ such that for all $i > i^1$, $\alpha_i \in [0,\varepsilon)$. □

Lemma A.50 Assume that $\succsim$ has a state dependent dual-self variational representation. Let $f, g \in H$ and $\beta, \alpha \in (0,1)$. If $C^{g,\alpha}_{f,\beta} \in \mathbb{R}$, then

$$\beta f + (1-\beta)\big((\alpha C^{g,\alpha}_{f,\beta} + 1/2)f^* + (1/2 - \alpha C^{g,\alpha}_{f,\beta})g^*\big) \sim \alpha g + (1-\alpha)\big(\beta f + (1-\beta)h^*\big)$$

Proof. Assume, per contra,

$$\beta f + (1-\beta)\big((\alpha C^{g,\alpha}_{f,\beta} + 1/2)f^* + (1/2 - \alpha C^{g,\alpha}_{f,\beta})g^*\big) \not\sim \alpha g + (1-\alpha)\big(\beta f + (1-\beta)h^*\big).$$

Assume w.l.o.g.

$$\beta f + (1-\beta)\big((\alpha C^{g,\alpha}_{f,\beta} + 1/2)f^* + (1/2 - \alpha C^{g,\alpha}_{f,\beta})g^*\big) \succ \alpha g + (1-\alpha)\big(\beta f + (1-\beta)h^*\big)$$

since the other case is symmetric. Now the set

$$A \coloneqq \Big\{\alpha \in [0,1]\Big|\alpha\big(\beta f + (1-\beta)f^*\big) + (1-\alpha)\big(\beta f + (1-\beta)g^*\big) \succ \alpha g + (1-\alpha)\big(\beta f + (1-\beta)h^*\Big\}$$

is open by Axiom 1.2 and includes $\alpha C^{g,\alpha}_{f,\beta} + 1/2$. Thus there exists $\varepsilon > 0$ such that for all $\gamma \in (-\varepsilon,\varepsilon)$, $\alpha C^{g,\alpha}_{f,\beta} + 1/2 + \gamma \in A$. Thus for all $\gamma \in (-\varepsilon/\alpha, \varepsilon/\alpha)$,

$$\beta f + (1-\beta)\Big(\big(\alpha(C^{g,\alpha}_{f,\beta} + \gamma) + 1/2\big)f^* + \big(1/2 - \alpha(C^{g,\alpha}_{f,\beta} + \gamma)\big)g^*\Big)$$

$$\equiv (\alpha C^{g,\alpha}_{f,\beta} + 1/2 + \alpha\gamma)(\beta f + (1-\beta)f^*) + (1 - \alpha C^{g,\alpha}_{f,\beta} + 1/2 + \alpha\gamma)(\beta f + (1-\beta)g^*)$$

$$\succ \alpha g + (1-\alpha)\big(\beta f + (1-\beta)h^*\big).$$

However, this contradicts the definition of $C^{g,\alpha}_{f,\beta}$ since there exists $\widehat{\alpha} \in [-1/2\alpha, 1/2\alpha]$ such that $\left|C^{g,\alpha}_{f,\beta} - \widehat{\alpha}\right| < \frac{\varepsilon}{\alpha}$ and

$$\beta f + (1-\beta)\Big((\alpha\widehat{\alpha} + 1/2)f^* + (1/2 - \alpha\widehat{\alpha}))g^*\Big) \sim \alpha g + (1-\alpha)(\beta f + (1-\beta)h^*)$$

□

The next two lemmas show how $\lim_{\alpha\to 0} C^{g,\alpha}_{f,\beta}$ captures the directional derivative to direction g at differentiability points.

Lemma A.51 Let (u, I) be a state dependent dual-self variational representation and $f, g \in H, \beta \in (0,1)$. Let $(\alpha_i)_{i=1}^{\infty} \subset (0,1)$ be such that $\lim_{i\to\infty} \alpha_i = 0$. Then there exists i^1 such that for all $i > i^1$,

$$\frac{I\Big(\alpha_i u(g) + (1-\alpha_i)u\big(\beta f + (1-\beta)h^*\big)\Big) - I\Big(u\big(\beta f + (1-\beta)h^*\big)\Big)}{\alpha_i}$$
$$= (1-\beta)C^{g,\alpha_i}_{f,\beta}\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big).$$

Proof. By Lemmas A.49 and A.50, there exists $i^1 \in \mathbb{N}$ such that for all $i > i^1$,

$$I\Big(\alpha_i u(g) + (1-\alpha_i)u\big(\beta f + (1-\beta)h^*\big)\Big)$$
$$= I\Big(\beta u(f) + (1-\beta)\big((\alpha_i C^{g,\alpha_i}_{f,\beta} + 1/2)u(f^*) + (1/2 - \alpha_i C^{g,\alpha_i}_{f,\beta})u(g^*)\big)\Big).$$

Let $i > i^1$. We have by Lemmas A.48 and A.50,

$$I\Big(\alpha_i u(g) + (1-\alpha_i)u\big(\beta f + (1-\beta)h^*\big)\Big)$$
$$= I\Big(\beta u(f) + (1-\beta)\big((\alpha_i C^{g,\alpha_i}_{f,\beta} + 1/2)u(f^*) + (1/2 - \alpha_i C^{g,\alpha_i}_{f,\beta})u(g^*)\big)\Big)$$
$$= I\Big(\beta f + (1-\beta)g^*\Big) + (1-\beta)(1/2 + \alpha_i C^{g,\alpha_i}_{f,\beta})\big(I(f^*) - I(g^*)\big)$$
$$= I\Big(\beta f + (1-\beta)h^*\Big) + (1-\beta)(\alpha_i C^{g,\alpha_i}_{f,\beta})\big(I(f^*) - I(g^*)\big).$$

The claim follows by subtracting $I\Big(\beta f + (1-\beta)h^*\Big)$ from both sides and dividing both sides by α_i.

□

Lemma A.52 Let S be finite and (u, I) be a state dependent dual-self representation $f, g \in H$, $\beta \in (0,1)$, and $(\alpha_i)_{i=1}^{\infty} \subset (0,1)$ be such that $\lim_{i\to\infty} \alpha_i = 0$. If

$$\lim_{i\to\infty} C^{g,\alpha_i}_{f,\beta} \in \mathbb{R} \text{ or } \lim_{i\to\infty} \frac{I\Big(\alpha_i u(g) + (1-\alpha_i)u\big(\beta f + (1-\beta)h^*\big)\Big) - I\Big(u\big(\beta f + (1-\beta)h^*\big)\Big)}{\alpha_i} \in \mathbb{R},$$

then both limits exist and

$$\lim_{i\to\infty} \frac{I\Big(\alpha_i u(g) + (1-\alpha_i)u\big(\beta f + (1-\beta)h^*\big)\Big) - I\Big(u\big(\beta f + (1-\beta)h^*\big)\Big)}{\alpha_i}$$
$$= (1-\beta)\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big) \lim_{i\to\infty} C^{g,\alpha_i}_{f,\beta}.$$

Proof. By Lemma A.51, there exists i^1 such that for all $i > i^1$,

$$\frac{I\Big(\alpha_i u(g) + (1-\alpha_i)u\big(\beta f + (1-\beta)h^*\big)\Big) - I\Big(u\big(\beta f + (1-\beta)h^*\big)\Big)}{\alpha_i}$$
$$= (1-\beta)C^{g,\alpha_i}_{f,\beta}\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big).$$

Thus taking the limit of $i \to \infty$ shows the claim since either the limit on the left or the right hand side exists. □

The next lemma shows how interior points are the same for all representations.

Lemma A.53 Let (u, I) and $(\widetilde{u}, \widetilde{I})$ be state dependent dual-self variational representations for $\succsim$. Let $f \in H$. Then $u(f) \in \operatorname{int} u(H)$ iff $\widetilde{u}(f) \in \operatorname{int} \widetilde{u}(H)$

Proof. We show w.l.o.g. that if $u(f) \in \operatorname{int} u(H)$, then $\widetilde{u}(f) \in \operatorname{int} \widetilde{u}(H)$. By Lemma A.31 and Proposition A.12, for all $s \in S^{\mathcal{P}}$, there exist $A_s > 0, B_s \in \mathbb{R}$ such that $\widetilde{u}_s = A_s u_s + B_s$. Thus if $u(f) \in \operatorname{int} u(H)$, there exists $\varepsilon > 0$ such that $B_\infty(u(f), \varepsilon) \subseteq \operatorname{int} u(H)$ and so $B_\infty(u(f), \varepsilon) \subseteq \operatorname{int} u(H)$. Thus by above, $B_\infty(\widetilde{u}(f), \min_{A_s | s\in S^{\mathcal{P}}} \varepsilon) \subseteq \operatorname{int} \widetilde{u}(H)$, which shows the claim. □

The next lemma shows how we can relate directional derivatives of different representations to each other using the behavioral directional derivatives.

Lemma A.54 Let (u, I) and $(\widetilde{u}, \widetilde{I})$ be state dependent dual-self variational representations for $\succsim$. Let $A \in \mathbb{R}^S_{++}, B \in \mathbb{R}^S$ be such that for all $s \in S$, $\widetilde{u}_s = A_s u_s + B_s$. Denote $A^{-1} \coloneqq (A_s^{-1})_{s\in S}$. Let $f \in H$ be such that $u(f) \in \operatorname{int} u(H)$. If I is differentiable at $u(f)$ with derivative $\nabla I\big(u(f)\big)$, then $\widetilde{I}$ is differentiable at $\widetilde{u}(f)$ with derivative

$$\frac{\widetilde{I}(\widetilde{u}(f^*)) - \widetilde{I}(\widetilde{u}(g^*))}{I\big(u(f^*)\big) - I\big(u(g^*)\big)} A^{-1} \nabla I\big(u(f)\big).$$

Proof. Let the derivative of I at $u(f)$ be $\nabla I\big(u(f)\big)$. By Lemma A.31 and Proposition A.12, for all $s \in S^{\mathcal{P}}$, there exist $A_s > 0, B_s \in \mathbb{R}$ such that

$$\widetilde{u}_s = A_s u_s + B_s. \tag{A.62}$$

By Lemma A.53, $\widetilde{u}(f) \in \operatorname{int} \widetilde{u}(H)$. Let $v \in \mathbb{R}^S$. Then there exists $\alpha^* > 0$ such that $\widetilde{u}(f) + \alpha^* v, \widetilde{u}(f) - \alpha^* v \in \widetilde{u}(H)$. Let $g, \widetilde{g} \in H$ be such that $\widetilde{u}(g) = \widetilde{u}(f) + \alpha^* v$ and $\widetilde{u}(\widetilde{g}) = \widetilde{u}(f) - \alpha^* v$. Let $\beta \in (0,1)$ be such that $\frac{1}{\beta}\widetilde{u}(f) - \frac{1-\beta}{\beta}u(h^*) \in \widetilde{u}(H)$. Let $\widetilde{f} \in H$ be such that $u(\widetilde{f}) = \frac{1}{\beta}\widetilde{u}(f) - \frac{1-\beta}{\beta}u(h^*)$. Then

$$\widetilde{u}(\beta\widetilde{f} + (1-\beta)h^*) = \widetilde{u}(f). \tag{A.63}$$

Thus by (A.62) $u\big(\beta\widetilde{f} + (1-\beta)h^*\big) = u(f)$. Let $(\alpha_i)_{i=1}^{\infty} \subseteq (0,1)$ be such that $\alpha_i \to 0$. Since I is differentiable at $u\big(\beta\widetilde{f} + (1-\beta)h^*\big) = u(f)$ and so

$$\lim_{i\to\infty} \frac{I\big(\alpha_i u(g) + (1-\alpha_i)u(f)\big) - I\big(u(f)\big)}{\alpha_i} = \nabla I\big(u(f)\big) \cdot \big(u(g) - u(f)\big) \in \mathbb{R}.$$

Thus by Lemma A.52, $\lim_{i\to\infty} C_{\widetilde{f},\beta}^{g,\alpha_i}$ exists and

$$(1-\beta)\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big) \lim_{i\to\infty} C_{\widetilde{f},\beta}^{g,\alpha_i} = \nabla I\big(u(f)\big) \cdot \big(u(g) - u(f)\big).$$

Thus by Lemma A.52 and (A.63),

$$\begin{aligned}\lim_{i\to\infty} \frac{\widetilde{I}\big(\alpha_i\widetilde{u}(g) + (1-\alpha_i)\widetilde{u}(f)\big) - I\big(\widetilde{u}(f)\big)}{\alpha_i} &= (1-\beta)\Big(\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)\Big) \lim_{i\to\infty} C_{\widetilde{f},\beta}^{g,\alpha_i} \qquad \text{(A.64)}\\ &= \frac{\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)}{I\big(u(f^*)\big) - I\big(u(g^*)\big)} \nabla I\big(u(f)\big) \cdot \big(u(g) - u(f)\big).\end{aligned}$$

Symmetrically for $\widetilde{g}$, we have

$$\begin{aligned}\lim_{i\to\infty} \frac{\widetilde{I}\big(\alpha_i\widetilde{u}(\widetilde{g}) + (1-\alpha_i)\widetilde{u}(f)\big) - I\big(\widetilde{u}(f)\big)}{\alpha_i} &= (1-\beta)\Big(\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)\Big) \lim_{i\to\infty} C_{\widetilde{f},\beta}^{\widetilde{g},\alpha_i} \qquad \text{(A.65)}\\ &= \frac{\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)}{I\big(u(f^*)\big) - I\big(u(g^*)\big)} \nabla I\big(u(f)\big) \cdot \big(u(\widetilde{g}) - u(f)\big).\end{aligned}$$

Next by (A.62), we have

$$\begin{aligned}\nabla I\big(u(f)\big) \cdot \big(u(g) - u(f)\big) &= \sum_{s\in S} \nabla I\big(u(f)\big)_s A_s^{-1}\big(\widetilde{u}(g) - \widetilde{u}(f)\big)\\ &= \sum_{s\in S} A_s^{-1} \nabla I\big(u(f)\big)_s \alpha^* v_s = A^{-1}\nabla I\big(u(f)\big) \cdot (\alpha^* v).\end{aligned}$$

Since $\widetilde{u}(g) - \widetilde{u}(f) = -\big(\widetilde{u}(\widetilde{g}) - \widetilde{u}(f)\big)$, by (A.62), we have $u(g) - u(f) = -\big(u(\widetilde{g}) - u(f)\big)$. Thus by (A.64,A.65) and the choice of g, we have

$$\begin{aligned}\lim_{i\to\infty} \frac{\widetilde{I}\big(\widetilde{u}(f) + \alpha_i\alpha^* v)\big) - I\big(\widetilde{u}(f)\big)}{\alpha^*\alpha_i} &= \frac{1}{\alpha^*} \frac{\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)}{I\big(u(f^*)\big) - I\big(u(g^*)\big)} A^{-1}\nabla I\big(u(f)\big) \cdot (\alpha^* v)\\ = \lim_{i\to\infty} \frac{\widetilde{I}\big(\widetilde{u}(f) - \alpha_i\alpha^* v)\big) - I\big(\widetilde{u}(f)\big)}{-\alpha^*\alpha_i}&\end{aligned}$$

Since the sequence $(\alpha_i)_{i=1}^{\infty} \subseteq (0,1)$ such that $\alpha_i \to 0$ was arbitrary, we have that

$$\lim_{\alpha\to 0} \frac{\widetilde{I}\big(\widetilde{u}(f) + \alpha v)\big) - I\big(\widetilde{u}(f)\big)}{\alpha} = \frac{\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)}{I\big(u(f^*)\big) - I\big(u(g^*)\big)} A^{-1}\nabla I\big(u(f)\big) \cdot v.$$

Since $v \in \mathbb{R}^S$ was arbitrary, $\widetilde{I}$ is differentiable at $\widetilde{u}(f)$ with derivative

$$\frac{\widetilde{I}(\widetilde{u}(f^*)) - \widetilde{I}(\widetilde{u}(g^*))}{I(u(f^*)) - I(u(g^*))} A^{-1} \nabla I(u(f)).$$

□

Finally, we have some useful simple lemmas. The first lemma relates Clarke derivatives to standard derivatives.

Lemma A.55 Let S be finite and (u, I) be a state dependent dual-self variational representation for $\succsim$. Then for all $f \in H$ such that $u(f) \in \operatorname{int} u(H)$,

$$\partial I(u(f)) = \operatorname{convex}\{\lim_{i\to\infty} \nabla I(\varphi_i) | (\varphi_i)_{i=1}^{\infty} \subseteq u(H), \lim_{i\to\infty} \varphi_i = u(f), \forall i \in \mathbb{N}, I \text{ differentiable at } \varphi_i\}.$$

Proof. Since I is Lipschitz on $u(H)$, this follows directly from Clarke's (1983) Theorem 2.5.1. □

The next lemma shows how for dual-self representation the derivative gives the value of the representation.

Lemma A.56 Let S be finite and (u, I) be a state dependent dual-self representation for $\succsim$. Let $f \in H$ be such that $u(f) \in \operatorname{int}_\infty u(H)$. If I is differentiable at $u(f)$ with derivative p, then $I(u(f)) = p \cdot u(f)$.

Proof. Since $u(f) \in \operatorname{int}_\infty u(H)$, there exists $\alpha^0 > 0$ such that $(1 + \alpha^0)u(f), (1 - \alpha^0)u(f) \in u(H)$. Since $u(H)$ is convex and I is positive homogeneous, we have for all $\alpha \in (-\alpha^0, \alpha^0) \setminus \{0\}$,

$$I\big((1+\alpha)u(f)\big) = (1+\alpha)I\big(u(f)\big) \Longrightarrow \frac{I\big(\alpha u(f) + u(f)\big) - I\big(u(f)\big)}{\alpha} = I\big(u(f)\big).$$

Thus by taking the limit $\alpha \to 0$, we have by the differentiability of I at $u(f)$,

$$I\big(u(f)\big) = \lim_{\alpha\to 0} \frac{I\big(\alpha u(f) + u(f)\big) - I\big(u(f)\big)}{\alpha} = p \cdot u(f).$$

□

The last lemma is an observation of separating any vector that has a constant value for all probabilities to a constant number and a vector perpendicular to the probabilities y.

Lemma A.57 Let $C \subseteq \Delta(S)$. If there exist $x \in \mathbb{R}^S_{++}$ and $\beta \in \mathbb{R}_+$ such that for all $p \in C$,

$$\sum_{s\in S} x_s p_s = \beta$$

then there exists $y \in C^\perp$ such that

$$x = \bar{1}_S \beta + y$$

Proof. Let $x \in \mathbb{R}^S_{++}$ be as in the antecedent. Since $\sum_{s\in S} p_s = 1$, we have for all $p \in \widetilde{C}$,

$$\sum_{s\in S}(x - \beta\overline{1}_S)_s p_s = 0.$$

Thus

$$(x - \beta\overline{1}_S) \in \widetilde{C}^\perp \coloneqq \{y \in \mathbb{R}^S | \, \forall\, p \in \widetilde{C}, \sum_{s\in S} y_s p_s = 0\}.$$

So there exists $y \in \widetilde{C}^\perp$ such that

$$y = x - \beta\overline{1}_S \;\Rightarrow\; \beta\overline{1}_S + y = x.$$

□

A.4.3 Partial Identification

The next lemma shows the general identification for dual-self variational representation's utilities and probabilities.

Lemma A.58 Let (u, I) and $(\widetilde{u}, \widetilde{I})$ be state dependent dual-self variational representations for $\succsim$. Then there exist $x \in \left(\bigcup_{\varphi\in u(H)} \partial I(\varphi)\right)^\perp, \alpha > 0, B \in \mathbb{R}^S \beta \in \mathbb{R}$ such that when we denote the mapping for all $\mu \in \Delta(S)$,

$$(1 + x)(\mu) \coloneqq \left((1 + x_s)\mu_s\right)_{s\in S}$$

we have

$$\bigcup_{\varphi\in \operatorname{int} \widetilde{u}(H)} \partial\widetilde{I}(\varphi) = (1 + x) \bigcup_{\varphi\in \operatorname{int} u(H)} \partial I(\varphi).$$

and for all $S^{\mathcal{P}}$,

$$(\widetilde{u}_s)_{s\in S^{\mathcal{P}}} = \left(\frac{\alpha}{1 + x_s} u_s + B_s\right)_{s\in S^{\mathcal{P}}}.$$

Proof. By Lemma A.31 and Proposition A.12, there exist $A \in \mathbb{R}^S_{++}, B \in \mathbb{R}^S_{++}$ such that for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = A_s u_s + B_s.$$

Denote

$$\alpha \coloneqq \frac{\widetilde{I}(\widetilde{u}(f^*)) - \widetilde{I}(\widetilde{u}(g^*))}{I(u(f^*)) - I(u(g^*))} > 0.$$

By Lemmas A.53 and A.54, I is differentiable at $\varphi \in \operatorname{int} u(H)$ with derivative $\nabla I(\varphi)$ if and only if $\widetilde{I}$ is differentiable at $A\varphi + B \in \operatorname{int} \widetilde{u}(H)$ with derivative

$$\alpha A^{-1} \nabla I(u(f)).$$

Thus for all $\varphi \in \operatorname{int} u(H)$, by Lemma A.55,

$$\partial \widetilde{I}(A\varphi + B) = \alpha A^{-1} \partial I(\varphi).$$

Thus by Lemma A.53,

$$\bigcup_{\varphi \in \operatorname{int} \widetilde{u}(H)} \partial \widetilde{I}(\varphi) = \bigcup_{\varphi \in \operatorname{int} u(H)} \alpha A^{-1} \partial I(\varphi).$$

By continuity of A^{-1} operator, we have

$$\bigcup_{\varphi \in \operatorname{int} \widetilde{u}(H)} \partial \widetilde{I}(\varphi) = \alpha A^{-1} \bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi).$$

Thus for all $p \in \bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)$, we have

$$\alpha A^{-1} p \in \Delta(S).$$

Thus by Lemma A.57, there exists $x \in \left(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\right)^{\perp}$ such that

$$1 + x = \alpha A^{-1}.$$

Thus especially for all $s \in S$,

$$A_s = \frac{\alpha}{1 + x_s}.$$

This proves the claim. □

The next lemma shows that the values of the dual-self variational representation are unique up to a positive affine transformation.

Lemma A.59 Let (u, I) and $(\widetilde{u}, \widetilde{I})$ be state dependent dual-self variational representations for $\succsim$. Then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that for all $f \in H$,

$$\widetilde{I}(\widetilde{u}(f)) = \alpha I(u(f)) + \beta$$

where

$$\alpha = \frac{\widetilde{I}(\widetilde{u}(f^*)) - \widetilde{I}(\widetilde{u}(g^*))}{I(u(f^*)) - I(u(g^*))} \text{ and } \beta = \widetilde{I}(\widetilde{u}(g^*)) - \alpha I(u(g^*)).$$

Proof. Let α and β be as above. Define the mapping $K : (\inf I(u(H)), \sup I(u(H))) \to \mathbb{R} \cup \{\infty\}$

$$K(v) = \sup \left\{ \left| \widetilde{I}(\widetilde{u}(f)) - \alpha I(u(f)) + \beta \right| | f \in H, I \circ u(f) = v \right\}$$

where $\sup \varnothing = \infty$. We show that K is locally constant. Let $v \in (\inf I(u(H)), \sup I(u(H)))$. First, there exist $f^1, f^2 \in H$ such that $I(u(f^1)) > v > I(u(f^2))$. Since S is finite there exists $f^3 \in H$ such that $u(f^3) \in \operatorname{int} u(H)$. By mixture continuity, there exist $\alpha^1, \alpha^2, \alpha^3 \in (0, 1)$ such that $\alpha^1 + \alpha^2 + \alpha^3 = 1$ and $I \circ u(\alpha^1 f^1 + \alpha^2 f^2 + \alpha^3 f^3) = v$. Denote $f^4 := \alpha^1 f^1 + \alpha^2 f^2 + \alpha^3 f^3$. Since $\alpha^3 > 0$ and $u(f^3) \in \operatorname{int} u(H)$

and $u(H)$ is a convex set $u(f^4) \in \operatorname{int} u(H)$. Thus there exists $\varepsilon > 0$ such that $B_\infty(u(f^4), \varepsilon) \subseteq u(H)$. Let $\beta \in (1/2, 1)$ be such that

$$(\frac{1}{\beta} - 1)\Big\|u(f^4)\Big\|_\infty + \frac{1-\beta}{\beta}\|u(h^*)\|_\infty < \varepsilon.$$

Thus by the choice of β, there exists $f \in H$ such that

$$u(f) = u(f^4) + (\frac{1}{\beta} - 1)u(f^4) - \frac{1-\beta}{\beta}u(h^*).$$

Thus especially $u(f) \in \operatorname{int} u(H)$ and

$$\beta u(f) + (1-\beta)u(h^*) = u(f^4)$$

and so $I\big(\beta u(f) + (1-\beta)u(h^*)\big) = v$. By Lemma A.47, for all $\alpha \in (0,1)$, we have

$$\begin{aligned} &I\big(\beta u(f) + (1-\beta)u(f^*\alpha g^*)\big) \\ &= I\big(\beta u(f) + (1-\beta)u(h^*)\big) + (1-\beta)(\alpha - 1/2)\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big) \end{aligned} \tag{A.66}$$

and

$$\begin{aligned} &\widetilde{I}\big(\beta\widetilde{u}(f) + (1-\beta)\widetilde{u}(f^*\alpha g^*)\big) \\ &= \widetilde{I}\big(\beta\widetilde{u}(f) + (1-\beta)\widetilde{u}(h^*)\big) + (1-\beta)(\alpha - 1/2)\Big(\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)\Big). \end{aligned} \tag{A.67}$$

Thus let

$$\widetilde{v} \in \Big(v - 1/2(1-\beta)\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big), v + 1/2(1-\beta)\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big)\Big).$$

and let $\widetilde{\alpha} \in (0,1)$ be such

$$\widetilde{v} = I\big(\beta u(f) + (1-\beta)u(h^*)\big) + (1-\beta)(\widetilde{\alpha} - 1/2)\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big).$$

Let $\widetilde{f} \in H$ be such that $I \circ u(\widetilde{f}) = \widetilde{v}$. Then $\widetilde{f} \sim \beta\widetilde{u}(f) + (1-\beta)\widetilde{u}(f^*\widetilde{\alpha}g^*)\big)$. Thus by (A.66,A.67),

$$\begin{aligned} I\big(u(\widetilde{f})\big) &= \widetilde{I}\big(\beta\widetilde{u}(f) + (1-\beta)u(f^*\widetilde{\alpha}g^*)\big) \\ &= I\big(\beta u(f) + (1-\beta)u(h^*)\big) + (1-\beta)(\widetilde{\alpha} - 1/2)\Big(I\big(u(f^*)\big) - I\big(u(g^*)\big)\Big) \end{aligned}$$

and

$$\begin{aligned} \widetilde{I}\big(\widetilde{u}(\widetilde{f})\big) &= \widetilde{I}\big(\beta\widetilde{u}(f) + (1-\beta)\widetilde{u}(f^*\widetilde{\alpha}g^*)\big) \\ &= \widetilde{I}\big(\beta\widetilde{u}(f) + (1-\beta)\widetilde{u}(h^*)\big) + (1-\beta)(\widetilde{\alpha} - 1/2)\Big(\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)\Big). \end{aligned}$$

Hence

$$\begin{aligned} &\widetilde{I}\big(\widetilde{u}(\widetilde{f})\big) - \alpha I\big(u(\widetilde{f})\big) - \beta \\ &= \widetilde{I}\big(\beta\widetilde{u}(f) + (1-\beta)\widetilde{u}(h^*)\big) + (1-\beta)(\widetilde{\alpha} - 1/2)\Big(\widetilde{I}\big(\widetilde{u}(f^*)\big) - \widetilde{I}\big(\widetilde{u}(g^*)\big)\Big) \end{aligned}$$

$$-\frac{\widetilde{I}(\widetilde{u}(f^*))-\widetilde{I}(\widetilde{u}(g^*))}{I(u(f^*))-I(u(g^*))}\Big(I\big(\beta u(f)+(1-\beta)u(h^*)\big)+(1-\beta)(\widetilde{\alpha}-1/2)\Big(I\big(u(f^*)\big)-I\big(u(g^*)\big)\Big)\Big)-\beta$$

$$=\widetilde{I}\big(\beta\widetilde{u}(f)+(1-\beta)\widetilde{u}(h^*)\big)-\alpha I\big(\beta u(f)+(1-\beta)u(h^*)\big)-\beta.$$

Since the last line does not depend on $\widetilde{f},\widetilde{\alpha},\widetilde{v}$, we have for all

$$\widetilde{v}\in\Big(v-1/2(1-\beta)\Big(I\big(u(f^*)\big)-I\big(u(g^*)\big)\Big),v+1/2(1-\beta)\Big(I\big(u(f^*)\big)-I\big(u(g^*)\big)\Big)\Big),$$

$$K(\widetilde{v})=\Big|\widetilde{I}\big(\beta\widetilde{u}(f)+(1-\beta)\widetilde{u}(h^*)\big)-\alpha I\big(\beta u(f)+(1-\beta)u(h^*)\big)-\beta\Big|.$$

Thus K is constant in a neighborhood of v. Since $v\in(\inf I(u(H)),\sup I(u(H)))$ was arbitrary, K is locally constant function. Since the interval $(\inf I(u(H)),\sup I(u(H)))$ is connected, by (Viro et al., 2008, Problem 12.2x) K is constant function. Finally, let $f\in H$ be such that $I\circ u(f)=I\circ u(g^*)$. Then we have $f\sim g^*$ and so

$$\widetilde{I}\big(\widetilde{u}(f)\big)-\alpha I\big(u(f)\big)-\beta=\widetilde{I}\big(\widetilde{u}(g^*)\big)-\alpha I\big(u(g^*)\big)-\widetilde{I}\big(\widetilde{u}(g^*)\big)+\alpha I\big(u(g^*)\big)=0.$$

Since f such that $I\circ u(f)=I\circ u(g^*)$ was arbitrary, we have $K\big(I\circ u(g^*)\big)=0$. Thus K is a constant function at 0. Finally, by the continuity of I and $\widetilde{I}$, we have for all $f\in H$,

$$\widetilde{I}\big(\widetilde{u}(f)\big)=\alpha I\big(u(f)\big)+\beta.$$

□

These results characterize the identification for dual-self variational representation.

Corollary A.60 Let $(u,\mathbb{C})$ be a state dependent tight dual-self variational representation for $\succsim$, and $(\widetilde{u},\widetilde{\mathbb{C}})$ be a utility-cost set pair. Then $(\widetilde{u},\widetilde{\mathbb{C}})$ is a state dependent tight dual-self variational representation for $\succsim$ if and only if there exist $x\in\mathbb{R}^S_{++},B\in\mathbb{R}^S,\beta\in\mathbb{R}$ such that for all $p,q\in\cup_{c\in\mathbb{C}}\operatorname{dom}c$,

$$\sum_{s\in S}x_sp_s=\sum_{s\in S}x_sq_s=:\alpha,$$

for all $s\in S^{\mathcal{P}}$,

$$\widetilde{u}_s=\frac{1}{x_s}u_s+B_s,$$

and for all $f\in H$,

$$\max_{\widetilde{c}\in\widetilde{\mathbb{C}}}\min_{p\in\Delta(S)}\sum_{s\in S}p_s\widetilde{u}_s(f_s)+\widetilde{c}(p)=\max_{c\in\mathbb{C}}\min_{p\in\Delta(S)}\sum_{s\in S}p_s\widetilde{u}_s(f_s)+\alpha c(x/\alpha p)-\sum_{s\in S}B_sp_s+\beta.$$

Especially $(\widetilde{u},\alpha(\mathbb{C}\circ x/\alpha)-dB+\beta)$, where multiplication and subtractions are done elementwise, is another state dependent tight dual-self variational representation for $\succsim$ and

$$\overline{\operatorname{co}}\bigcup_{\widetilde{c}\in\widetilde{\mathbb{C}}}\operatorname{dom}\widetilde{c}=\overline{\operatorname{co}}\bigcup_{c\in\mathbb{C}}\frac{x}{\alpha}\operatorname{dom}c.$$

Proof. This follows symmetrically to Corollary A.62 since for $I : u(H) \to \mathbb{R}$ by defining

$$I\big(u(f)\big) = \max_{c\in\mathbb{C}} \min_{p\in\Delta(S)} p \cdot u(f) + c(p),$$

we have by Chandrasekher et al.'s (2020) Section S.3,

$$\overline{\text{co}} \bigcup_{\varphi\in u(H)} \partial I(\varphi) = \overline{\text{co}} \bigcup_{c\in\mathbb{C}} \operatorname{dom} c.$$

□

We move onto dual-self representation. The next lemma shows the identification for probabilities and utilities.

Lemma A.61 Let (u, I) and $(\widetilde{u}, \widetilde{I})$ be state dependent dual-self representations for $\succsim$. Then there exist $x, y \in \big(\bigcup_{\varphi\in u(H)} \partial I(\varphi)\big)^{\perp}, \alpha > 0, \beta \in \mathbb{R}$ such that when we denote the mapping for all $\mu \in \Delta(S)$,

$$(1 + x)(\mu) \coloneqq \big((1 + x_s)\mu_s\big)_{s\in S}$$

we have

$$\bigcup_{\varphi\in\operatorname{int}\widetilde{u}(H)} \partial\widetilde{I}(\varphi) = (1 + x) \bigcup_{\varphi\in\operatorname{int} u(H)} \partial I(\varphi).$$

and for all $S^{\mathcal{P}}$,

$$(\widetilde{u}_s)_{s\in S^{\mathcal{P}}} = \left(\frac{\alpha}{1 + x_s}(u_s + y_s) + \beta\right)_{s\in S^{\mathcal{P}}}.$$

Proof. By the proof of Lemma A.58, there exist $x \in \big(\bigcup_{\varphi\in u(H)} \partial I(\varphi)\big)^{\perp}, \alpha > 0, B \in \mathbb{R}^S \beta \in \mathbb{R}$ such that

$$(\widetilde{u}_s)_{s\in S^{\mathcal{P}}} = \left(\frac{\alpha}{1 + x_s}u_s + B_s\right)_{s\in S^{\mathcal{P}}} \tag{A.68}$$

where

$$\alpha = \frac{\widetilde{I}(\widetilde{u}(f^*)) - \widetilde{I}(\widetilde{u}(g^*))}{I\big(u(f^*)\big) - I\big(u(g^*)\big)}.$$

Let $\beta \coloneqq \widetilde{I}(\widetilde{u}(g^*)) - \alpha I\big(u(g^*)\big)$.

By Lemmas A.53 and A.54, I is differentiable at $\varphi \in \operatorname{int} u(H)$ with derivative $\nabla I(\varphi)$ if and only if $\widetilde{I}$ is differentiable at $A\varphi + B \in \operatorname{int}\widetilde{u}(H)$ with derivative

$$(1 + x)\nabla I\big(u(f)\big).$$

Let $f \in H$ be such that $u(f) \in \operatorname{int} u(H)$ and $u(f)$ is a differentiability point of I. Then by Lemma A.56, $I(u(f)) = \nabla I(u(f)) \cdot u(f)$ and $\widetilde{I}(\widetilde{u}(f)) = (1 + x)\nabla I(u(f)) \cdot \widetilde{u}(f)$. Thus we have by (A.68)

$$\widetilde{I}(\widetilde{u}(f)) = (1 + x)\nabla I(u(f)) \cdot \frac{\alpha}{1 + x}u(f) + B$$

$$= \alpha \nabla I(u(f)) \cdot u(f) + (1+x)\nabla I(u(f)) \cdot B = \alpha I(u(f)) + (1+x)\nabla I(u(f)) \cdot B.$$

Thus by Lemma A.59,

$$(1+x)\nabla I(u(f)) \cdot B = \widetilde{I}(\widetilde{u}(f)) - \alpha I(u(f)) = \beta$$

and so

$$(1+x)\nabla I(u(f)) \cdot B = \widetilde{I}(\widetilde{u}(f)) - \alpha I(u(f)) = \beta$$

Since f was arbitrary such that $u(f) \in \operatorname{int} u(H)$ and $u(f)$ is a differentiability point of I, we have by Lemma A.55, for all $p \in \bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)$,

$$\beta = (1+x)p \cdot B = \sum_{s \in S} \big((1+x_s)B_s\big)p_s$$

and since $(1+x)p \in \Delta(S)$

$$0 = \sum_{s \in S} \big((1+x_s)B_s - \beta(1+x_s)\big)p_s.$$

Thus by Lemma A.57, there exists $\alpha y \in \big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\big)^{\perp}$ such that

$$\alpha y = (1+x)B - (1+x)\beta \iff B = \frac{\alpha y}{1+x} + \beta.$$

Thus since $\big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\big)^{\perp}$ is a linear space $y \in \big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\big)^{\perp}$ and for all $s \in S^{\mathcal{P}}$,

$$(\widetilde{u}_s)_{s \in S^{\mathcal{P}}} = \left(\frac{\alpha}{1+x_s}(u_s + y_s) + \beta\right)_{s \in S^{\mathcal{P}}}.$$

Now for all differentiability points, $p \cdot u(f) = I(u(f))$ and $(1+x)p \cdot \widetilde{u}(f) = \widetilde{I}(\widetilde{u}(f))$. Thus

$$(1+x)p \cdot (\frac{\alpha}{1+x}(1+x)u(f) + B) = \widetilde{I}(\widetilde{u}(f))$$

and so

$$(1+x)p \cdot B = \widetilde{I}(\widetilde{u}(f)) - \alpha I\big(u(f)\big).$$

Since this holds for all differentiability points, this shows the claim. □

The next lemma shows the identification for dual-self representation.

Corollary A.62 Let $(u, \mathbb{P})$ be a state dependent tight dual-self representation for $\succsim$, and $(\widetilde{u}, \widetilde{\mathbb{P}})$ be a utility-probability set collection pair. Then $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self representation for $\succsim$ if and only if there exist $x \in \mathbb{R}^S_{++}, y \in \mathbb{R}^S$ such that for all $p, q \in \bigcup_{P \in \mathbb{P}} P$,

$$\sum_{s \in S} x_s p_s = \sum_{s \in S} x_s q_s =: \alpha \text{ and } \sum_{s \in S} y_s p_s = \sum_{s \in S} y_s q_s,$$

for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = \frac{1}{x_s}(u_s + y_s),$$

and for all $f \in H$,

$$\max_{\widetilde{P}\in\widetilde{\mathbb{P}}} \min_{\widetilde{p}\in\widetilde{P}} \sum_{s\in S} \widetilde{p}_s \widetilde{u}_s(f_s) = \max_{P\in x/\alpha\mathbb{P}} \min_{p\in P} \sum_{s\in S} p_s \widetilde{u}_s(f_s).$$

Especially $(\widetilde{u}, x/\alpha\mathbb{P})$ is another state dependent tight dual-self representation for $\succsim$ and

$$\overline{\text{co}} \bigcup_{\widetilde{P}\in\widetilde{\mathbb{P}}} \widetilde{P} = \overline{\text{co}} \bigcup_{P\in\mathbb{P}} \frac{x}{\alpha} P.$$

Proof. Let $(\widetilde{u}, \widetilde{\mathbb{P}})$ be a state dependent tight dual-self representation for $\succsim$. Define $I : u(H) \to \mathbb{R}$ and $\widetilde{I} : \widetilde{u}(H) \to \mathbb{R}$ by for all $f \in H$,

$$I\big(u(f)\big) = \max_{P\in\mathbb{P}} \min_{p\in P} p \cdot u(f) \text{ and } \widetilde{I}\big(\widetilde{u}(f)\big) = \max_{\widetilde{P}\in\widetilde{\mathbb{P}}} \min_{\widetilde{p}\in\widetilde{P}} \widetilde{p} \cdot \widetilde{u}(f).$$

Now (u, I) and $(\widetilde{u}, \widetilde{I})$ are state dependent dual-self representations for $\succsim$. By Chandrasekher et al.'s (2020) Lemma B.3

$$\overline{\text{co}} \bigcup_{\varphi\in u(H)} \partial I(\varphi) = \overline{\text{co}} \bigcup_{P\in\mathbb{P}} P \text{ and } \overline{\text{co}} \bigcup_{\widetilde{\varphi}\in\widetilde{u}(H)} \partial \widetilde{I}(\widetilde{\varphi}) = \overline{\text{co}} \bigcup_{\widetilde{P}\in\widetilde{\mathbb{P}}} \widetilde{P}.$$

Thus by Lemma A.61, there exist $\widetilde{x}, \widetilde{y} \in \big(\bigcup_{P\in\mathbb{P}} P\big)^{\perp}, \widetilde{\alpha} > 0, \beta \in \mathbb{R}$ such that

$$\bigcup_{\varphi\in\text{int}\,\widetilde{u}(H)} \partial \widetilde{I}(\varphi) = (1 + \widetilde{x}) \bigcup_{\varphi\in\text{int}\,u(H)} \partial I(\varphi).$$

and for all $S^{\mathcal{P}}$,

$$(\widetilde{u}_s)_{s\in S^{\mathcal{P}}} = \left(\frac{\alpha}{1 + \widetilde{x}_s} (u_s + \widetilde{y}_s) + \beta \right)_{s\in S^{\mathcal{P}}}.$$

Thus define $y = \widetilde{y} + \widetilde{\alpha}^{-1}\beta$, $x = \widetilde{\alpha}^{-1}(1 + \widetilde{x})$, and $\alpha = \widetilde{\alpha}^{-1}$ which gives the utility identification. Additionally, we have for all $f \in H$,

$$\max_{P\in x/\alpha\mathbb{P}} \min_{\widetilde{p}\in\widetilde{P}} \widetilde{p} \cdot \widetilde{u}(f) = \alpha \max_{P\in\mathbb{P}} \min_{p\in P} p \cdot u(f) + \beta.$$

And $x/\alpha\mathbb{P} \in \mathcal{K}\big(\Delta(S)\big)$ since $\|x/\alpha\|_\infty < \infty$. Thus $(\widetilde{u}, x/\alpha\mathbb{P})$ is a state dependent dual-self representation for $\succsim$ and thus by Lemma A.59, for all $f \in H$,

$$\max_{\widetilde{P}\in\widetilde{\mathbb{P}}} \min_{\widetilde{p}\in\widetilde{P}} \sum_{s\in S} \widetilde{p}_s \widetilde{u}_s(f_s) = \max_{P\in x/\alpha\mathbb{P}} \min_{p\in P} \sum_{s\in S} p_s \widetilde{u}_s(f_s).$$

□

A.4.4 Full Identification

Next, we move on to full-identification and show how Axiom 1.7 characterizes the full identification. The next lemma shows that instead of utility trade-off for acts in Axiom 1.7, we can use statewise preference trade-offs for them.

Lemma A.63 Let $\succsim$ have a state dependent dual-self variational representation with (u, I). Let $f, g \in H$ be equally crisp acts such that there exist $s, s' \in S^{\mathcal{P}}$ such that $u_s(f_s) > u_s(g_s)$ and $u_{s'}(g_{s'}) > u_{s'}(f_{s'})$. Then there exist equally crisp acts f^1, g^1 and $s^1, s^2 \in S^{\mathcal{P}}$ such that $f^1 \sim g^1$, $f^1_{s^1} \succ_{s^1} g^1_{s^1}$, and $g^1_{s^2} \succ_{s^2} f^1_{s^2}$.

Proof. Assume that there exist $\widetilde{f}^*, \widetilde{g}^*, s, s' \in S$ such that $u_s(\widetilde{f}^*)_s > u_s(\widetilde{g}^*)_s$ and $u_{s'}(\widetilde{g}^*)_{s'} > u_{s'}(\widetilde{f}^*_{s'})$ and $\widetilde{f}^*, \widetilde{g}^*$ are equally crisp. Assume without loss of generality $\widetilde{f}^* \succsim \widetilde{g}^*$. By Lemma A.48, for all $\varphi \in u(H), \alpha \in (0, 1]$ such that $\varphi + \alpha(u(\widetilde{f}^*) - u(\widetilde{g}^*)) \in u(H)$,

$$I\big(\varphi + \alpha(u(\widetilde{f}^*) - u(\widetilde{g}^*))\big) = I(\varphi) + \alpha\big(I(u(\widetilde{f}^*)) - I(u(\widetilde{g}^*))\big). \tag{A.69}$$

Let $g^0 \in H$ be such that $u(f^0) \in \operatorname{int} u(H)$. Now there exists $\alpha^* > 0$ such that $u(f^0) + \alpha^*(u(\widetilde{f}^*) - u(\widetilde{g}^*)), u(f^0) - \alpha^* \bar{1}\big(I(u(\widetilde{f}^*)) - I(u(\widetilde{g}^*))\big) \in u(H)$. Let $f^0, f^1 \in H$ be such that

$$u(f^0) = u(f^0) + \alpha^*(u(\widetilde{f}^*) - u(\widetilde{g}^*))$$

and

$$u(f^1) = u(f^0) - \alpha^* \bar{1}\big(I(u(\widetilde{f}^*)) - I(u(\widetilde{g}^*))\big).$$

By Equation (A.69) and C-additivity of I, (g^0, f^0) and (g^0, f^1) are equally crisp acts. Then by Lemma A.40, $(g^0, 1/2 f^0 + 1/2 f^1)$ are equally crisp acts. Denote $f^2 \coloneqq 1/2 f^0 + 1/2 f^1$. Now we have since $u(H)$ is a convex set

$$\begin{aligned} I(u(f^2)) &= I\Big(u(f^0) + 1/2\alpha^*(u(\widetilde{f}^*) - u(\widetilde{g}^*)) - 1/2\alpha^* \bar{1}\big(I(u(\widetilde{f}^*)) - I(u(\widetilde{g}^*))\big)\Big) \\ &\overset{\text{C-add.}}{=} I\Big(u(f^0) + 1/2\alpha^*\big(u(\widetilde{f}^*) - u(\widetilde{g}^*)\big)\Big) - 1/2\alpha^*\big(I(u(\widetilde{f}^*)) - I(u(\widetilde{g}^*))\big) \\ &\overset{(A.69)}{=} I\big(u(f^0)\big) + 1/2\alpha^*\big(I(u(\widetilde{f}^*)) - I(u(\widetilde{g}^*))\big) - 1/2\alpha^*\big(I(u(\widetilde{f}^*)) - I(u(\widetilde{g}^*))\big) = I(u(f^0)). \end{aligned} \tag{A.70}$$

Thus $f^2 \sim f^0$. First, since $u_{s'}(\widetilde{g}^*_{s'}) > u_{s'}(\widetilde{f}^*_{s'})$, we have

$$u_{s'}(f^2_{s'}) = u_{s'}(f^0_{s'}) + 1/2\alpha^*\big(u_{s'}(\widetilde{f}^*_{s'}) - u_{s'}(\widetilde{g}^*_{s'})\big) + 1/2\big(I(u(\widetilde{g}^*)) - I(u(\widetilde{f}^*))\big) < u_{s'}(f^0_{s'}).$$

Since $s' \in S^{\mathcal{P}}$, by Lemma A.55, there exists $\varphi \in \operatorname{int} u(H)$ such that I is differentiable at φ and $\nabla I(\varphi)_{s'} > 0$. By Lemma A.48, we have

$$\nabla I(\varphi) \cdot \big(u(f^2) - u(f^0)\big) = 0.$$

Thus there exists s'' such that $u_{s''}(f^2_{s''}) > u_{s''}(g^0_{s''})$ and $\nabla I(\varphi)_{s''} > 0$. Let g_3, f^3, α^3 be such that $u(g^3) = \varphi$, $u(f^3) = \varphi + \alpha^3\big(u(f^2) - u(g^0)\big)$. By Lemma A.48, (g^3, f^3) are equally crisp acts and by (A.70) $g^3 \sim f^3$ and

$$u_{s''}(f^3_{s''}) - u_{s''}(g^3_{s''}) > 0 > u_{s'}(f^3_{s'}) - u_{s'}(g^3_{s'}).$$

Additionally, by the differentiability of I at $u(g^3)$, there exists

$$0 < \varepsilon < \min\{u_{s''}(f^3_{s''}) - u_{s''}(g^3_{s''}), u_{s'}(g^3_{s'}) - u_{s'}(f^3_{s'})\}$$

such that

$$I(u(g^3) + \mathrm{pr}_{s'}\,\varepsilon) > I(u(g^3)) + \nabla(I(\varphi))_{s'}\frac{\varepsilon}{2} \text{ and } I(u(g^3) - \mathrm{pr}_{s'\prime}\,\varepsilon) < I(u(g^3)) - \nabla(I(\varphi))_{s'}\frac{\varepsilon}{2}.$$

Thus

$$I(u(f^3_{s''}, g^3_{-s''})) \leq I(u(g^3) - \mathrm{pr}_{s'\prime}\,\varepsilon) < I(u(g^3)) \text{ and } I(u(f^3_{s'}, g^3_{-s'})) \geq I(u(g^3) + \mathrm{pr}_{s'\prime}\,\varepsilon) > I(u(g^3)).$$

Thus $f^3_{s'} \succ_{s'} g^3_{s'}$ and $g^3_{s''} \succ_{s''} f^3_{s''}$ and (g^3, f^3) are equally crisp. □

The next lemma shows the equivalency between Axiom 1.7 and equally crisp acts being statewise ranked.

Lemma A.64 Assume that $\succsim$ has a state dependent dual-self variational representation with (u, I). Then the following are equivalent:

(I) If f, g are equally crisp acts, then for all $s \in S$, $f_s \succsim_s g_s$ or for all $s \in S$, $g_s \succsim_s f_s$.

(II) For all f, g and $s, s' \in S$ such that $f \sim g$, $f_s \succ_s g_s$, and $f_{s'} \succ_{s'} g_{s'}$, there exist $h \in H, \alpha \in (0,1)$ such that $\alpha h + (1-\alpha)f \not\sim \alpha h + (1-\alpha)g$.

Proof. Assume that (II) does not hold and we show that (I) does not hold. By the counter assumption, there exist $f, g \in H, s, s' \in S$ such that $f \sim g$, $f_s \succ_s g_s$, and $f_{s'} \succ_{s'} g_{s'}$ and for all $h \in H, \alpha \in (0,1)$, $\alpha h + (1-\alpha)f \not\sim \alpha h + (1-\alpha)g$. Now we have for all $h, h', \alpha \in (0,1)$,

$$\alpha h + (1-\alpha)f \sim \alpha h + (1-\alpha)g \text{ and } \alpha h' + (1-\alpha)f \sim \alpha h' + (1-\alpha)g$$

Thus

$$\alpha h + (1-\alpha)f \succsim \alpha h' + (1-\alpha)f \iff \alpha h + (1-\alpha)g \succsim \alpha h' + (1-\alpha)g.$$

Hence f, g are equally crisp and I) does not hold.

Assume that (I) does not hold and we show that (II) does not hold. Assume that there exist $\widetilde{f}^*, \widetilde{g}^*, s, s' \in S$ such that $\widetilde{f}^*_s \succ_s \widetilde{g}^*_s$ and $\widetilde{g}^*_{s'} \succ_{s'} \widetilde{f}^*_{s'}$ and $\widetilde{f}^*, \widetilde{g}^*$ are equally crisp. First, by Lemma A.31, $u_s(\widetilde{f}^*)_s > u_s(\widetilde{g}^*)_s$ and $u_{s'}(\widetilde{g}^*)_{s'} > u_{s'}(\widetilde{f}^*_{s'})$. Thus by Lemma A.63, there exist equally crisp acts f^1, g^1 and $s^1, s^2 \in S^{\mathcal{P}}$ such that $f^1 \sim g^1$, $f^1_{s^1} \succ_{s^1} g^1_{s^1}$, and $g^1_{s^2} \succ_{s^2} f^1_{s^2}$. However, for all $h \in H, \alpha \in (0,1)$, by Lemma A.48,

$$I\big(\alpha u(h) + (1-\alpha)u(f^1)\big) = I\big(\alpha u(h) + (1-\alpha)u(g^1)\big) + (1-\alpha)\big(I(u(f^1)) - I(u(g^1))\big)$$

$$= I\big(\alpha u(h) + (1-\alpha)u(g^1)\big)$$

and hence $\alpha h + (1-\alpha)f^1 \sim \alpha h + (1-\alpha)g^1$. This contradicts (II). □

The next lemma shows that Axiom 1.7 characterizes the full-dimensional uncertainty that gives the identification by previous partial identification.

Lemma A.65 Assume that $\succsim$ has a state dependent dual-self variational representation with (u, I). Then the following are equivalent:

(I) $\succsim$ satisfies Axiom 1.7.

(II) $\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi)$ has a non-empty interior in $\Delta(S^{\mathcal{P}})$.

Proof. Enumerate $S^{\mathcal{P}} = \{\tilde{s}_1, \ldots, \tilde{s}_n\}$ where $n = |S^{\mathcal{P}}|$. Assume that (I) holds. Then if f, g are equally crisp acts, by Lemmas A.63 and A.64, for all $s, s' \in S^{\mathcal{P}}$, $u_s(f_s) - u_s(g_s) = u_{s'}(f_{s'}) - u_{s'}(g_{s'})$. Thus $\Big(\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi)\Big)^{\perp} = \{\bar{0}\}$. Hence $\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi)$ contains $|S^{\mathcal{P}}| = n$ linearly independent probabilities $\{p^1, \ldots, p^n\}$. Consider the set $\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi) - p^n$ where the subtraction is vector subtraction in $\mathbb{R}^{S^{\mathcal{P}}}$. Now the set $\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi) - p^n \supseteq \{p^1 - p^n, \ldots, p^{n-1} - p^n\}$ is linearly independent. Consider the matrix $P = [p^i - p^n]_{1 \le i \le n-1}$. First, the rank of P is $n-1$ by above. Second, we have for all $1 \le i \le n-1$, $\sum_{j=1}^n p^i_j - p^n_j = 0$. Third, since the row rank of P is $n-1$ and the last row $\tilde{s}_n$ is linearly dependent on rows $\{\tilde{s}_1, \ldots, \tilde{s}_{n-1}\}$, the row vectors of P for rows $\{\tilde{s}_1, \ldots, \tilde{s}_{n-1}\}$ are linearly independent. Next, consider the matrix $\tilde{P} = [\mathrm{pr}_{S^{\mathcal{P}} \setminus \{\tilde{s}_n\}} p^i - p^n]_{1 \le i \le n-1}$. This is collecting row vectors of P for rows $\{\tilde{s}_1, \ldots, \tilde{s}_{n-1}\}$ and thus by above, the rank of $\tilde{P}$ is $n-1$. Thus especially $\{\mathrm{pr}_{S^{\mathcal{P}} \setminus \{\tilde{s}_n\}} p^1 - p^n, \ldots, \mathrm{pr}_{S^{\mathcal{P}} \setminus \{\tilde{s}_n\}} p^{n-1} - p^n\}$ are linearly independent. Thus the set $\mathrm{pr}_{S^{\mathcal{P}} \setminus \{\tilde{s}_n\}} \mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi) - p^n$ contains the zero vector and $n-1$ linearly independent vectors and is a convex set. Hence the set $\mathrm{pr}_{S^{\mathcal{P}} \setminus \{\tilde{s}_n\}} \mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi) - p^n$ cannot be contained in any $n-2$-dimensional set and thus it has a non-empty interior in $\mathbb{R}^{S^{\mathcal{P}} \setminus \{\tilde{s}_n\}}$ (Boyd and Vandenberghe, 2004, Section 2.5.2). Thus especially the set $\mathrm{pr}_{S^{\mathcal{P}} \setminus \{\tilde{s}_n\}} \mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi)$ has a non-empty interior in $\mathbb{R}^{S^{\mathcal{P}} \setminus \{\tilde{s}_n\}}$ as a translation. Thus by the choice of topology for $\Delta(S^{\mathcal{P}})$, $\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi)$ has a non-empty interior in $\Delta(S^{\mathcal{P}})$.

Assume that (II) holds. Then $\mathrm{pr}_{S^{\mathcal{P}}} \overline{\mathrm{co}} \bigcup_{\varphi \in \mathrm{int}\, u(H)} \partial I(\varphi)$ contains $|S^{\mathcal{P}}|$ linearly independent probabilities. Let f, g be equally crisp acts. Assume w.l.o.g. that there exists $s \in S^{\mathcal{P}}$ such that

$u_s(f_s) > u_s(g_s)$ or for all $s \in S^{\mathcal{P}}$, $u_s(f_s) = u_s(g_s)$. Then by Lemmas A.48 and A.55, for all $p, q \in \overline{\text{co}} \bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)$,

$$p \cdot \big(u(f) - u(g)\big) = q \cdot \big(u(f) - u(g)\big) \coloneqq \alpha.$$

By Lemma A.57, there exists $y \in \Big(\overline{\text{co}} \bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)\Big)^{\perp}$ such that for all $s \in S$,

$$\alpha + y = u(f) - u(g).$$

Since $\Big(\overline{\text{co}} \bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)\Big)^{\perp} = \{\bar{0}\}$, we have for all $s \in S^{\mathcal{P}}$, $u_s(f_s) - u_s(g_s) = \alpha$. Now $\alpha \geq 0$ by the above assumption. Thus for all $s \in S^{\mathcal{P}}$, $u_s(f_s) \geq u_s(g_s)$. Thus by Lemma A.31, for all $s \in S^{\mathcal{P}}$, $f_s \succsim_s g_s$. This shows (I) by Lemma A.64.

□

A.4.5 Identification Characterization with Crisp Acts and Equally Crisp Acts

Next, we show Corollaries 1.6 and 1.11 and characterize the identification with crisp and equally crisp acts.

Preliminary Results

We first show some simple preliminary results. The first result shows for a dual-self expected utility taking a convex combination with a crisp act does not affect the derivative of the representation.

Lemma A.66 Let (u, I) be a state dependent dual-self representation for $\succsim$. If $c \in H$ is crisp, $\alpha \in (0,1)$ and $h \in H$ is such that $u(h) \in \text{int}\, u(H)$ and I is differentiable at $u(h)$, then I is differentiable at $\alpha u(c) + (1-\alpha)u(h)$ with

$$\nabla I\big(\alpha u(c) + (1-\alpha)u(h)\big) = \nabla I\big(u(h)\big).$$

Proof. Let $\alpha^* \in (0,1)$. As in the proof of Lemma A.32, there exist equally crisp $f^*, g^* \in H, a > 0$ such that for all $s \in S^{\mathcal{P}}$, $u_s(f_s^*) - u_s(g_s^*) = a$ and $f_s^* \succ_s g_s^*$. Denote $h^* = \frac{1}{2}f^* + \frac{1}{2}g^*$.

We use notation from Appendix A.4 from using equally crisp acts f^*, g^*. Since $u(h) \in \text{int}\, u(H)$, there exist $\hat{h} \in H, \beta \in (0,1)$ such that for all $s \in S^{\mathcal{P}}$, $u_s(\beta\hat{h} + (1-\beta)h^*) = u_s(h)$.

Let $g \in H$. Denote $\tilde{h} = \frac{\alpha^*}{\alpha^* + (1-\alpha^*)\beta} c + \frac{(1-\alpha^*)\beta}{\alpha^* + (1-\alpha^*)\beta}\hat{h}$, $\tilde{\beta} = \alpha^* + (1-\alpha^*)\beta$, $\tilde{g} = \alpha^* c + (1-\alpha^*)g$. We will show that for all $\alpha \in (0,1)$,

$$C^{g,\alpha}_{\hat{h},\beta} = C^{\tilde{g},\alpha}_{\tilde{h},\tilde{\beta}}. \tag{A.71}$$

This follows from the definition of $C^{g,\alpha}_{\hat{h},\beta}$ and the observation that for all $\widehat{\alpha} \in [-1/2\alpha, 1/2\alpha]$, we have

$$\begin{aligned}
&\beta\hat{h} + (1-\beta)\big((\alpha\widehat{\alpha} + 1/2)f^* + (1/2 - \alpha\widehat{\alpha})g^*\big) \sim \alpha g + (1-\alpha)\big(\beta\hat{h} + (1-\beta)h^*\big) \\
\iff &\alpha^* c + (1-\alpha^*)\Big(\beta\hat{h} + (1-\beta)\big((\alpha\widehat{\alpha} + 1/2)f^* + (1/2 - \alpha\widehat{\alpha})g^*\big)\Big) \\
&\qquad\qquad \sim \alpha^* c + (1-\alpha^*)\Big(\alpha g + (1-\alpha)\big(\beta\hat{h} + (1-\beta)h^*\big)\Big) \\
\iff &\tilde{\beta}\tilde{h} + (1-\tilde{\beta})\big((\alpha\widehat{\alpha} + 1/2)f^* + (1/2 - \alpha\widehat{\alpha})g^*\big)\Big) \sim \Big(\alpha\tilde{g} + (1-\alpha)\big(\tilde{\beta}\tilde{h} + (1-\tilde{\beta})h^*\big)\Big).
\end{aligned}$$

where the second equivalence follows as an identity.

Since I is differentiable at $u(h)$, the differentiability follows from Lemma A.52 as in the proof of Lemma A.54 since for all $v \in \mathbb{R}^{S^{\mathcal{P}}}$, there exist $g \in H, a^* > 0$ such that for all $s \in S^{\mathcal{P}}$, $u_s(g) = u_s(h) + a^* v$ and so

$$u_s(\alpha^* c + (1-\alpha^*)g) = u_s\big(\tilde{\beta}\tilde{h} + (1-\tilde{\beta})h^*\big) + (1-\alpha^*)a^* v.$$

Additionally, the derivative being the same follows from Lemma A.52.

□

Next, we show that the expected utility of crisp acts is the same for all probabilities.

Lemma A.67 Let (u, I) be a state dependent dual-self representation for $\succsim$. If $c \in H$ is crisp, then for all $p \in \bigcup_{\varphi \in u(H)} \partial I(\varphi)$,

$$p \cdot u(c) = I(u(c)).$$

Proof. Let $h \in H$ be such that $u(h) \in \operatorname{int} u(H)$ and I is differentiable at $u(h)$. Then by Lemmas A.56 and A.66, for all $\alpha \in [0, 1)$,

$$I\big(u(\alpha c + (1-\alpha)h)\big) = \nabla I\big(u(h)\big) \cdot u(\alpha c + (1-\alpha)h).$$

Thus by the continuity of I and by taking limit $\alpha \to 1$,

$$I\big(u(c)\big) = \nabla I\big(u(h)\big) \cdot u(c).$$

Finally, the claim follows from Lemma A.55. □

Identification with Crisp and Equally Crisp Acts

Corollary A.68 Let $(u, \mathbb{P})$ be a state dependent dual-self representation for $\succsim$ and $x \in \mathbb{R}^S_{++}, y \in \mathbb{R}^S$. The following two conditions are equivalent:

1. $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent dual-self representation for $\succsim$ such that for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = \frac{1}{x_s}(u_s + y_s)$$

and

$$0, 1 \in \bigcap_{s \in S^{\mathcal{P}}} \widetilde{u}_s(\Delta(X_s)).$$

2. There exist crisp acts f^*, g^* such that $u_s(g_s^*) = -y_s$ and $u_s(f_s^*) = -y_s + x_s$.

Proof. Follows from Lemma A.67 and Proposition 1.3 since $x \in \mathbb{R}^S_{++}$. □

Corollary A.69 Let $(u, \mathbb{C})$ be a state dependent dual-self variational representation for $\succsim$ and $x \in \mathbb{R}^S_{++}, B \in \mathbb{R}^S$. The following two conditions are equivalent:

1. $(\widetilde{u}, \widetilde{\mathbb{C}})$ is a state dependent dual-self variational representation for $\succsim$ such that for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = \frac{1}{x_s} u_s + \frac{B_s}{x_s}$$

and

$$0, 1 \in \bigcap_{s \in S^{\mathcal{P}}} \widetilde{u}_s(\Delta(X_s)).$$

2. There exist equally crisp acts f^*, g^* such that $u_s(g_s^*) = -B_s$ and $u_s(f_s^*) = -B_s + x_s$.

Proof. Follows from Lemmas A.48 and A.72 and Theorem 1.9 since $x \in \mathbb{R}^S_{++}$. □

A.4.6 Relative Likelihood and Probability Characterizations

Next, we move on to proving Propositions 1.3 and 1.4. We show these results by characterizing the possible transformations for probabilities behaviorally and in terms of the set of probabilities: First, for the relative likelihood identification, we show when all the possible probability transformations change the probabilities for states s and s' similarly that is when for all $x \in \left(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\right)^{\perp}$, $x_s = x_{s'}$. Second, for the probability identification, we show when all the possible probability transformations keep the probability for state s the same that is when for all $x \in \left(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\right)^{\perp}$, $x_s = 0$. We show these results for state dependent dual-self expected utility.

Preliminary Results

We start with some simple lemmas. The next lemma is a trivial one showing that non-null states are characterized by zero probability.

Lemma A.70 Let (u, I) be a state dependent dual-self variational representation for $\succsim$. If $s \in S^{\mathcal{P}}$, then there exists $p \in \bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)$ such that $p_s > 0$.

Proof. Follows from the definition of $S^{\mathcal{P}}$ and the definition of Clarke derivatives. □

The next lemma shows that any utility trade-off between two states can be represented with acts that have statewise preference trade-offs in those two states.

Lemma A.71 Let (u, I) be a state dependent dual-self variational representation for $\succsim$. Let $x \in \mathbb{R}^S$ and $s, s' \in S^{\mathcal{P}}$ be such that $x_s > 0 > x_{s'}$ and for all $\tilde{s} \notin S^{\mathcal{P}}$, $x_{\tilde{s}} = 0$. Then there exist $f, g \in H, a > 0$ such that $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$, $u(f), u(g) \in \operatorname{int} u(H)$, and $u(f) - u(g) = \alpha x$.

Proof. Since $s, s' \in S^{\mathcal{P}}$ by Lemmas A.55 and A.70, there exist $f^s, f^{s'} \in H$ such that $u(f^s), u(f^{s'}) \in \operatorname{int} u(H)$, I is differentiable at $u(f^s), u(f^{s'})$ and

$$\nabla I(u(f^s))_s, \nabla I(u(f^{s'}))_{s'} > 0. \tag{A.72}$$

Let $f \in H$ be such that $u(f) \in \operatorname{int} u(H)$, $u_s(f_s) = \varphi^s_s$, and $u_{s'}(f_{s'}) = \varphi^{s'}_{s'}$. Since $u(f) \in \operatorname{int} u(H)$ and for all $\tilde{s} \notin S^{\mathcal{P}}$, $x_{\tilde{s}} = 0$, there exist $a > 0, g \in H$ such that $u(f) - u(g) = ax$. Additionally, since I is monotonic by (A.72), we have

$$(f_s, f^s_{-s}) \succ (g_s, f^s_{-s}) \text{ and } (g_{s'}, f^{s'}_{-s'}) \succ (f_{s'}, f^{s'}_{-s'}).$$

□

The last lemma shows that if x has a constant expected utility for all the probabilities, then the representation is additive in adding utility x to any utility φ.

Lemma A.72 Let (u, I) be a state dependent dual-self variational representation for $\succsim$. Let $x \in \mathbb{R}^S$ and $c \in \mathbb{R}$. The following two are equivalent:

1. $(x - \bar{1}c) \in \left(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi) \right)^{\perp}$.
2. For all $\varphi \in \operatorname{int} u(H)$ such that $\varphi + x \in \operatorname{int} u(H)$, we have

$$I(\varphi + x) = I(\varphi) + c.$$

Proof. We first show that (1)⇒(2). By Chandrasekher et al.'s (2020) Section S.3,

$$\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi) \subseteq \Delta(S).$$

Since $\operatorname{int} u(H)$ is convex, for all $\alpha \in [0, 1]$, $\varphi + \alpha x \in \operatorname{int} u(H)$ Define a mapping $[0, 1] \ni \alpha \overset{J}{\mapsto} I(\varphi + \alpha x)$. Since I is C-additive, it is 1-Lipschitz. So especially J is Lipschitz and so differentiable almost everywhere. Now if J is differentiable at $\alpha \in (0, 1)$, we have by Lemma A.55,

$$J'(\alpha) \leq \sup_{\psi \in \partial I(\psi + \alpha x)} \psi \cdot x = \sup_{\psi \in \partial I(\psi + \alpha x)} \psi \cdot (x - \bar{1}c) + \psi \cdot \bar{1}c = \sup_{\psi \in \partial I(\psi + \alpha x)} \psi \cdot \bar{1}c = c$$

and symmetrically

$$J'(\alpha) \geq \inf_{\psi \in \partial I(\psi + \alpha x)} \psi \cdot x = c.$$

Let $\Omega \subseteq (0,1)$ be the set of differentiability points for J. By the Fundamental theorem of calculus, we have

$$I(\varphi + x) = J(0) + J(1) - J(0) = I(\varphi) + \int_\Omega J'(x)\, dx = I(\varphi) + c.$$

Next, we show that (2)⇒(1). Let $\varphi \in \operatorname{int} u(H)$ be such that I is differentiable at φ. Now we have for all $\alpha \in (0,1)$ such that $\varphi + \alpha(x - \bar{1}c), \varphi + \alpha x \in \operatorname{int} u(H)$, by C-additivity and (2),

$$I(\varphi + \alpha(x - \bar{1}c)) = I(\varphi + \alpha(x - \bar{1}c)) = I(\varphi + \alpha x) - \alpha c = I(\varphi) + \alpha c - \alpha c = I(\varphi).$$

Now there exists $\alpha^* \in (0,1)$ such that for all $\alpha \in (0,1)$. $\varphi + \alpha(x - \bar{1}c), \varphi + \alpha x \in \operatorname{int} u(H)$ and thus by the differentiability of I,

$$\nabla I(\varphi) \cdot (x - \bar{1}c) = \lim_{\alpha \to 0} \frac{I(\varphi + \alpha(x - \bar{1}c)) - I(\varphi)}{\alpha} = 0.$$

Thus since φ was arbitrary, we have by Lemma A.55, for all $p \in \big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\big)$, $p \cdot (x - \bar{1}c) = 0$. This shows (1). □

Relative Likelihood Identification

Our first lemma for relative likelihood identification shows that our behavioral assumption for relative likelihood identification characterizes their identification. This follows as a corollary of this lemma and our previous partial identification result.

Lemma A.73 Let (u, I) be a state dependent dual-self variational representation for $\succsim$ and $s, s' \in S^{\mathcal{P}}, s \neq s'$. The following two are equivalent:

1. For all $f, g \in H$ such that $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$, there exist $h, h' \in H$ and $\alpha \in (0,1)$ such that
$$\alpha h + (1-\alpha) f \succ \alpha h' + (1-\alpha) f \text{ and } \alpha h + (1-\alpha) g \prec \alpha h' + (1-\alpha) g.$$
2. For all $x \in \big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\big)^\perp$, $x_s = x_{s'}$.

Proof. We will first show that (1)⇒(2). Assume, per contra, there exists $x \in \big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\big)^\perp$, $x_s \neq x_{s'}$. Assume w.l.o.g. $x_s > x_{s'}$. Let $c = -\frac{1}{2}x_s - \frac{1}{2}x_{s'}$. By (A.71), there exist $f, g \in H, a > 0$ such that $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$ and $u(f) - u(g) = a(x + c)$. Let $h \in H$, $\alpha \in (0,1)$ with $u(h) \in \operatorname{int} u(H)$. By Lemma A.72, we have

$$I\big(\alpha u(h) + (1-\alpha) u(f)\big) = I\Big(\alpha u(h) + (1-\alpha) u(g) + (1-\alpha)\big(u(f) - u(g)\big)\Big)$$

$$= I\big(\alpha u(h) + (1-\alpha)u(g) + (1-\alpha)a(x+c)\big) = I\big(\alpha u(h) + (1-\alpha)u(g)\big) + (1-\alpha)ac.$$

Since I is continuous, we have for all $h \in H$, $\alpha \in (0,1)$

$$I\big(\alpha u(h) + (1-\alpha)u(f)\big) = I\big(\alpha u(h) + (1-\alpha)u(g)\big) + (1-\alpha)ac.$$

Thus we have for all $h, h' \in H$,

$$\begin{aligned}
&\alpha h + (1-\alpha)f \succsim \alpha h' + (1-\alpha)f \iff I\big(\alpha u(h) + (1-\alpha)u(f)\big) \geq I\big(\alpha u(h') + (1-\alpha)u(f)\big) \\
&\iff I\big(\alpha u(h) + (1-\alpha)u(f)\big) + (1-\alpha)ac \geq I\big(\alpha u(h') + (1-\alpha)u(f)\big) + (1-\alpha)ac \\
&\iff I\big(\alpha u(h)+(1-\alpha)u(g)\big) \geq I\big(\alpha u(h')+(1-\alpha)u(g)\big) \iff \alpha h+(1-\alpha)f \succsim \alpha h'+(1-\alpha)f.
\end{aligned}$$

This contradicts (1).

Second, we show that (2)⇒(1). Assume, per contra, that there exist $f, g \in H$ such that $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$ and for all $h, h', \alpha \in (0,1)$,

$$\alpha h + (1-\alpha)f \succ \alpha h' + (1-\alpha)f \Rightarrow \alpha h + (1-\alpha)g \succsim \alpha h' + (1-\alpha)g.$$

We show first that for all $h, h', \in H$ with $u(h), u(h') \in \operatorname{int} u(H), \alpha \in (0,1)$,

$$\alpha h + (1-\alpha)f \succ \alpha h' + (1-\alpha)f \Rightarrow \alpha h + (1-\alpha)g \succ \alpha h' + (1-\alpha)g.$$

By the continuity of I and since $h \in \operatorname{int} u(H)$, there exist $h^\dagger \in H, a > 0$ such that $u(h^\dagger) = u(h) - \bar{1}a$ and $\alpha h^\dagger + (1-\alpha)f \succ \alpha h' + (1-\alpha)f$. So by assumption and C-additivity of I, we have

$$\alpha h + (1-\alpha)g \succ \alpha h^\dagger + (1-\alpha)g \succsim \alpha h' + (1-\alpha)g.$$

Symmetrically by the negation of the assumption, we have for all $h, h', \alpha \in (0,1)$,

$$\alpha h + (1-\alpha)g \succ \alpha h' + (1-\alpha)g \Rightarrow \alpha h + (1-\alpha)f \succsim \alpha h' + (1-\alpha)f$$

and so for all $h, h', \in H$ with $u(h), u(h') \in \operatorname{int} u(H), \alpha \in (0,1)$,

$$\alpha h + (1-\alpha)g \succ \alpha h' + (1-\alpha)g \Rightarrow \alpha h + (1-\alpha)f \succ \alpha h' + (1-\alpha)f.$$

So for all $h, h', \in H$ with $u(h), u(h') \in \operatorname{int} u(H), \alpha \in (0,1)$,

$$\alpha h + (1-\alpha)f \succsim \alpha h' + (1-\alpha)f \iff \alpha h + (1-\alpha)g \succsim \alpha h' + (1-\alpha)g. \tag{A.73}$$

Thus f and g are equally crisp acts in interior of $u(H)$. By Lemma A.48, we have for all $\varphi \in \operatorname{int} u(H), \alpha \in (0,1)$ such that $\varphi + \alpha(u(f) - u(g)) \in \operatorname{int} u(H)$,

$$I(\varphi + \alpha(u(f) - u(g))) = I(\varphi) + \alpha\Big(I\big(u(f)\big) - I\big(u(g)\big)\Big).$$

Thus for all differentiability points φ of I in $\operatorname{int} u(H)$, we have

$$\nabla I(\varphi) \cdot \big(u(f) - u(g)\big) = I\big(u(f)\big) - I\big(u(g)\big).$$

By Chandrasekher et al.'s (2020) Section S.3

$$\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi) \subseteq \Delta(S).$$

Thus we have for all differentiability points φ of I in $\operatorname{int} u(H)$, we have

$$\nabla I(\varphi) \cdot \Big(u(f) - u(g) - \bar{1}\big(I(u(f)) - I(u(g))\big)\Big) = I\big(u(f)\big) - I\big(u(g)\big) - \Big(I\big(u(f)\big) - I\big(u(g)\big)\Big) = 0.$$

Additionally, by Lemma A.31,

$$u_s(f_s) > u_s(g_s) \text{ and } u_{s'}(g_{s'}) > u_{s'}(f_{s'})$$

and so

$$u_s(f_s) - u_s(g_s) - \big(I(u(f)) - I(u(g))\big) \neq u_{s'}(f_{s'}) - u_{s'}(g_{s'}) - \big(I(u(f)) - I(u(g))\big).$$

By Lemma A.55,

$$u(f) - u(g) - \bar{1}\big(I(u(f)) - I(u(g))\big) \in \Big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\Big)^{\perp},$$

which is a contradiction with (2). □

The next lemma characterizes in terms of probabilities when the relative likelihood for states is identified.

Lemma A.74 Let $P \subseteq \Delta(S), P \neq \varnothing, s, s' \in S, s \neq s'$. The following two are equivalent:

1. For all $x \in P^{\perp}$, $x_s = x_{s'}$.
2. There exist $p, q \in P$ such that for all $\tilde{s} \in S \setminus \{s, s'\}$, $p_{\tilde{s}} = q_{\tilde{s}}$ and $\frac{p_s}{p_{s'}} \neq \frac{q_s}{q_{s'}}$.

Proof. We will first show that (2)⇒(1). Assume, per contra, that there exists $x \in P^{\perp}$ such that $x_s \neq x_{s'}$. Let p, q be as in (2). Then

$$\sum_{\tilde{s} \in S} x_{\tilde{s}} p_{\tilde{s}} = 0 = \sum_{\tilde{s} \in S} x_{\tilde{s}} q_{\tilde{s}}$$

Since for all $\tilde{s} \in S \setminus \{s, s'\}$, $p_{\tilde{s}} = q_{\tilde{s}}$ and so $p_s + p_{s'} = q_s + q_{s'}$,

$$(x_s - x_{s'})p_s + x_s(p_s + p_{s'}) = x_s p_s + x_{s'} p_{s'} = x_s q_s + x_{s'} q_{s'} = (x_s - x_{s'})q_s + x_s(q_s + q_{s'}).$$

Hence,

$$(x_s - x_{s'})(p_s - q_s) = 0.$$

Since $x_s \neq x_{s'}$ by the counter assumption, we have $p_s = q_s$ which is a contradiction.

Next, we show that (1)⇒(2). Denote $n^{\perp} = \dim P^{\perp}$ and $n = \dim P$. Now $n^{\perp} + n = |S|$. There exist $(c^i)_{i=1}^{n^{\perp}} \subseteq P^{\perp}$ linearly independent vectors and $(p^i)_{i=1}^{n} \subseteq P$ linearly independent vectors. First, we show that $n \geq 2$. Since P is non-empty, $n \geq 1$. Assume, per contra, $n = 1$. Thus P is a

singleton $P = \{p\}$. If $p_s = 0$, then $(1_s, 0_{-s}) \in P^\perp$ which is a contradiction. Similarly $p_{s'} \neq 0$. Now $((\frac{1}{p}_{s'})_s, (-\frac{1}{p}_s)_{s'}, 0_{-s,s'}) \in P^\perp$ that is a contradiction. So $n \geq 2$.

We consider two cases. First, assume that for all $i \in \{1, \ldots, n^\perp\}$,

$$c^i_s = 0. \tag{A.74}$$

Now we can consider the matrix C formed by row vectors $(c^i)_{i=1}^{n^\perp}$ and each column in $\tilde{s} \in S$. By (A.74), we can reduce C into Smith normal form $\tilde{C}$ formed by row vectors $(\tilde{c}^i)_{i=1}^{n^\perp}$ such that for each $i \in \{1, \ldots, n^\perp\}$, there exists $s^i \in S \setminus \{s, s'\}$ such that $\tilde{c}^i_{s^i} = -1$, for all $j \in \{1, \ldots, n^\perp\} \setminus \{i\}$, $\tilde{c}^i_{s^j} = 0$, and for all $k, l \in \{1, \ldots, n^\perp\}$, $k \neq l$, $s^k \neq s^l$. Now $(\tilde{c}^i)_{i=1}^{n^\perp} \subseteq P^\perp$ are linearly independent and for all $i \in \{1, \ldots, n^\perp\}$,

$$\tilde{c}^i_s = \tilde{c}^i_{s'} = 0. \tag{A.75}$$

Denote $S^\perp = \{s^1 \ldots, s^{n^\perp}\}$. Now we have for all $i \in \{1, \ldots, n^\perp\}$ and $p \in P$,

$$p_{s^i} = \sum_{\tilde{s} \in S \setminus S^\perp} \tilde{c}^i_{\tilde{s}} p^j_{\tilde{s}}. \tag{A.76}$$

Let $\pi : \{1, \ldots, n-2\} \to S \setminus (S^\perp \cup \{s, s'\})$ be a one-to-one function. We show by induction that for each $i \in \{0, \ldots, n-2\}$, there exists a linearly independent collection of probabilities $(\tilde{p}^{j,i})_{j=1}^{n-i}$ such that for all $j, k \in \{1, \ldots, n-i\}$, $1 \leq l \leq i$,

$$p^{j,i}_{\pi(l)} = p^{k,i}_{\pi(l)}. \tag{A.77}$$

For the first step $i = 0$, define for all $j \in \{1, \ldots, n\}$, $\tilde{p}^{j,i} = p^j$. For the induction step, assume that for $0 \leq i \leq n-2$, there exists $(\tilde{p}^{j,i})_{j=1}^{n-i}$ that satisfy (A.77). First, we show that

$$\min \{p^{j,i}_{\pi(i+1)} | j \in \{1, \ldots, n-i\}\} \neq \max \{p^{j,i}_{\pi(i+1)} | j \in \{1, \ldots, n-i\}\}. \tag{A.78}$$

Assume, per contra, that for all $j, k \in \{1, \ldots, n-i\}$, $p^{j,i}_{\pi(i+1)} = p^{k,i}_{\pi(i+1)}$. We collect $(\tilde{p}^{j,i})_{j=1}^{n-i}$ into a matrix P^i where each $p^{j,i}$ is a column vector. The column rank of P^i is $n-i$ since each column is linearly independent. Thus the row rank of P^i is $n-i$. However, each row $\tilde{s} \in S^\perp$ is linearly dependent on rows $S \setminus S^\perp$ by (A.76) and each row $\tilde{s} \in \{\pi(1), \ldots, \pi(i), \pi(i+1)\}$ is constant and so since the rows sum to 1, linearly dependent on the rows $S \setminus \{\pi(1), \ldots, \pi(i), \pi(i+1)\}$. Thus the maximum row rank for P^i is $|S| - n^\perp - i - 1 = n - i - 1$ which is a contradiction. This shows (A.78).

Let

$$j^* \in \arg\max \{p^{j,i}_{\pi(i+1)} | j \in \{1, \ldots, n-i\}\} \text{ and } j_* \in \arg\min \{p^{j,i}_{\pi(i+1)} | j \in \{1, \ldots, n-i\}\}.$$

By (A.78), there exists $\beta^i \in (0,1)$ such that for all $j \in \{1,\dots,n-i\} \setminus j_*$,

$$\frac{1}{2}p^{j,i}_{\pi(i+1)} + \frac{1}{2}p^{j^*,i}_{\pi(i+1)} > \beta^i > p^{j_*,i}_{\pi(i+1)}.$$

Thus for each $j \in \{1,\dots,n-i\} \setminus j_*$, there exists $\alpha^j \in (0,1)$ such that

$$\alpha^j\Big(\frac{1}{2}p^{j,i}_{\pi(i+1)} + \frac{1}{2}p^{j^*,i}_{\pi(i+1)}\Big) + (1-\alpha^j)p^{j_*,i}_{\pi(i+1)} = \beta^i.$$

Denote for $j \in \{1,\dots,n-i-1\} \setminus j_*$,

$$\tilde{p}^{(j-\mathbb{1}(j>j_*)),i+1} = \alpha^j\Big(\frac{1}{2}p^{j,i} + \frac{1}{2}p^{j^*,i}\Big) + (1-\alpha^j)p^{j_*,i},$$

where $\mathbb{1}(j > j_*)$ is an indicator function for $j > j_*$. Now $(\tilde{p}^{j,i+1})_{j=1}^{n-i-1}$ are linearly independent since they have been created using elementary column operations using the above P^i. Additionally, it satisfies (A.77). This shows the induction step and concludes the induction.

Since $n \geq 2$, by the induction, there exist $p^*, p^\dagger \in P$ that are linearly independent and for all $\tilde{s} \in S \setminus (S^\perp \cup \{s, s'\})$, $p^*_{\tilde{s}} = p^\dagger_{\tilde{s}}$.

By (A.75,A.76), for all $\tilde{s} \in S \setminus \{s, s'\}$,

$$p^*_{\tilde{s}} = p^\dagger_{\tilde{s}}.$$

This shows the claim since $p^\dagger$ and p^* are linearly independent.

Second, we consider the case that there exists $i \in \{1,\dots,n^\perp\}$ such that $c^i_s \neq 0$. First, if a vector $\tilde{c} \in \mathbb{R}^S$ is such that $\tilde{c}_s = \tilde{c}_{s'} \neq 0$ and for all $\tilde{s} \in S \setminus \{s, s'\}$, $\tilde{c}_{\tilde{s}} = 0$, then

$$\tilde{c} \notin P^\perp \tag{A.79}$$

since each $p \in P$ is non-negative.

Now we can consider the matrix C formed by row vectors $(c^i)_{i=1}^{n^\perp}$ and each column in $\tilde{s} \in S$. Denote $s^{n^\perp} = s, s^{n^\perp+1} = s'$. By (A.79), we can reduce C into Smith normal form $\tilde{C}$ formed by row vectors $(\tilde{c}^i)_{i=1}^{n^\perp}$ such that for each $i \in \{1,\dots,n^\perp-1\}$, there exists $s^i \in S \setminus \{s, s'\}$ such that for each $\hat{i} \in \{1,\dots,n^\perp\}$, $\tilde{c}^{\hat{i}}_{s^i} = -1$, for all $j \in \{1,\dots,n^\perp\} \setminus \{i\}$, $\tilde{c}^{\hat{i}}_{s^j} = 0$, and for all $k, l \in \{1,\dots,n^\perp\}$, $k \neq l$, $s^k \neq s^l$. Now $(\tilde{c}^i)_{i=1}^{n^\perp} \subseteq P^\perp$ are linearly independent and for all $i \in \{1,\dots,n^\perp-1\}$,

$$\tilde{c}^i_s = \tilde{c}^i_{s'} = 0 \text{ and } \tilde{c}^{n^\perp}_s = \tilde{c}^{n^\perp}_{s'} = -1.$$

Denote $S^\perp = \{s^1 \dots, s^{n^\perp}, s^{n^\perp+1}\}$. Now we have for all $i \in \{1,\dots,n^\perp-1\}$ and $p \in P$,

$$p_{s^i} = \sum_{\tilde{s} \in S \setminus S^\perp} \tilde{c}^i_{\tilde{s}} p_{\tilde{s}} \text{ and } p_{s^{n^\perp}} + p_{s^{n^\perp+1}} = \sum_{\tilde{s} \in S \setminus S^\perp} \tilde{c}^{n^\perp}_{\tilde{s}} p_{\tilde{s}}. \tag{A.80}$$

Next, we show that there exists $s^\dagger \in S \setminus S^\perp$ such that

$$\sum_{i=1}^{n^\perp} \tilde{c}^i_{s^\dagger} \neq -1. \tag{A.81}$$

Assume, per contra, for all $\tilde{s} \in S \setminus S^\perp$,

$$\sum_{i=1}^{n^\perp} \tilde{c}^i_{\tilde{s}} = -1.$$

Then, we have for $p \in P$,

$$1 = \sum_{\tilde{s} \in S} p_{\tilde{s}} \overset{(A.80)}{=} \sum_{\tilde{s} \in S \setminus (S^\perp \cup \{s^\dagger\})} \Big(1 + \sum_{i=1}^{n^\perp} \tilde{c}^i_{\tilde{s}}\Big) p_{\tilde{s}} = 0$$

that is a contradiction which shows (A.81).

Let $\pi : \{1, \dots, n-2\} \to S \setminus (S^\perp \cup \{s^\dagger\})$ be a one-to-one function. By an induction as in the previous case for each $i \in \{0, \dots, n-2\}$, there exists a collection of linearly independent probabilities $(\tilde{p}^{j,i})_{j=1}^{n-i}$ such that for all $j, k \in \{1, \dots, n-i\}$, $1 \leq l \leq i$,

$$p^{j,i}_{\pi(l)} = p^{k,i}_{\pi(l)}. \tag{A.82}$$

Since $n \geq 2$, by the induction, there exist $p^*, p^\dagger \in P$ that are linearly independent and for all $\tilde{s} \in S \setminus (S^\perp \cup \{s^\dagger\})$, $p^*_{\tilde{s}} = p^\dagger_{\tilde{s}}$.

Now we have for $p \in \{p^*, p^\dagger\}$,

$$1 = \sum_{\tilde{s} \in S} p_{\tilde{s}} \overset{(A.80)}{=} \sum_{\tilde{s} \in S \setminus (S^\perp)} \Big(1 + \sum_{i=1}^{n^\perp} \tilde{c}^i_{\tilde{s}}\Big) p_{\tilde{s}} \overset{(A.81)}{\Longrightarrow} p_{s^\dagger} = \frac{1 - \sum_{\tilde{s} \in S \setminus (S^\perp \cup \{s^\dagger\})} \Big(1 + \sum_{i=1}^{n^\perp} \tilde{c}^i_{\tilde{s}}\Big) p_{\tilde{s}}}{1 + \sum_{i=1}^{n^\perp} \tilde{c}^i_{s^\dagger}}.$$

Thus $p^*_{s^\dagger} = p^\dagger_{s^\dagger}$.

Now for all $\tilde{s} \in S \setminus S^\perp$, $p^*_{\tilde{s}} = p^\dagger_{\tilde{s}}$. By (A.80), for all $\tilde{s} \in S \setminus \{s, s'\}$,

$$p^*_{\tilde{s}} = p^\dagger_{\tilde{s}} \text{ and } p^*_s + p^*_{s'} = p^\dagger_s + p^\dagger_{s'}.$$

This shows the claim since $p^\dagger$ and p^* are linearly independent.

□

The previous two lemmas give the relative likelihood identification characterization.

Proposition A.75 (Relative Likelihood Identification) Let $(u, \mathbb{P})$ be a state dependent tight dual-self representation for $\succsim$ and $s, s' \in S^{\mathcal{P}}$. The following four conditions are equivalent:

1. If $f, g \in H$ are such that and $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$, then there exist $h, h' \in H$ and $\alpha \in (0,1)$ such that

$$\alpha h + (1-\alpha) f \succ \alpha h + (1-\alpha) g \text{ and } \alpha h' + (1-\alpha) f \prec \alpha h' + (1-\alpha) g.$$

2. There exist $p, q \in \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P$ such that $\frac{p_s}{p_{s'}} \neq \frac{q_s}{q_{s'}}$ and for all $\tilde{s} \in S \setminus \{s, s'\}$, $p_{\tilde{s}} = q_{\tilde{s}}$.

3. If $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self representation for $\succsim$, then there exist $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}$ such that

$$\widetilde{u}_s = \alpha u_s + \beta \quad \text{and} \quad \widetilde{u}_{s'} = \alpha u_{s'} + \beta.$$

4. If $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self representation for $\succsim$, then

$$\Big\{ \frac{\widetilde{p}_s}{\widetilde{p}_{s'}} \Big| \widetilde{p} \in \overline{\text{co}} \bigcup_{\widetilde{P} \in \widetilde{\mathbb{P}}} \widetilde{P} \Big\} = \Big\{ \frac{p_s}{p_{s'}} \Big| p \in \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P \Big\}.$$

Proof. (1) $\iff$ (2) follows from Lemmas A.73 and A.74. (2) $\iff$ (3) and (2) $\iff$ (4) follows from Corollary A.62 and Lemma A.74 □

Probability Identification

The identification for a probability is similar to the identification for the relative likelihood of some states. First, we show that our behavioral assumption for probability identification characterizes its identification.

Lemma A.76 Let (u, I) be a state dependent dual-self variational representation for $\succsim$ and $s \in S^{\mathcal{P}}$. The following two are equivalent:

1. For all $f, g \in H$ such that $f \sim g$, $f_s \succ_s g_s$, and for some $s' \in S$, $g_{s'} \succ_{s'} f_{s'}$, there exist $h \in H$ and $\alpha \in (0, 1)$ such that

$$\alpha h + (1 - \alpha) f \not\sim \alpha h + (1 - \alpha) g.$$

2. For all $x \in \big(\bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)\big)^{\perp}$, $x_s = 0$.

Proof. We will first show that (1)⇒(2). Assume, per contra, there exists $\tilde{x} \in \big(\bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)\big)^{\perp}$, $\tilde{x}_s \neq 0$. Assume w.l.o.g. $\tilde{x}_s > 0$. Since for all $p \in \big(\bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)\big), s \notin S^{\mathcal{P}}$, we have $p_s = 0$ so $x = (\tilde{x}_{S^{\mathcal{P}}}, 0_{-S^{\mathcal{P}}}) \in \big(\bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)\big)^{\perp}$. We show that there exists $s' \in S^{\mathcal{P}}$ such that $x_{s'} < 0$. Assume, per contra, for all $s' \in S^{\mathcal{P}} x_{s'} \geq 0$. Since $s \in S^{\mathcal{P}}$, by Lemma A.70, there exists $p \in \bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)$ such that $p_s > 0$. So by the counterassumption and since p is non-negative,

$$\sum_{\tilde{s} \in S} x_{\tilde{s}} p_{\tilde{s}} \geq x_s p_s > 0$$

which is a contradiction since $x \in \big(\bigcup_{\varphi \in \text{int}\, u(H)} \partial I(\varphi)\big)^{\perp}$.

By Lemma A.71, there exist $f, g \in H$, $a > 0$ such that $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$, $u(f), u(g) \in \operatorname{int} u(H)$, and $u(f) - u(g) = ax$. Let $h \in H$ and $\alpha^* \in [0,1)$. Consider a mapping $J : [0,1] \to \mathbb{R}$, defined by for all $\alpha \in [0,1]$,

$$J(\alpha) = I \circ u\Big(\alpha^* h + (1-\alpha^*)\big(\alpha f + (1-\alpha)g\big)\Big) = I\big(\alpha^* u(h) + (1-\alpha^*)u(g) + (1-\alpha^*)ax\big).$$

Since $u(f), u(g) \in \operatorname{int} u(H)$, we have by the definition of Clarke derivative and since $x \in \big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\big)^{\perp}$, for all differentiability points α of J,

$$J'(\alpha) = 0.$$

Since J is differentiable almost everywhere, we have by the fundamental theorem of calculus,

$$I \circ u\big(\alpha^* h + (1-\alpha^*)g\big) = J(0) = J(1) = I \circ u\big(\alpha^* h + (1-\alpha^*)f\big)$$

and so $\alpha^* h + (1-\alpha^*)g \sim \alpha^* h + (1-\alpha^*)f$. Especially $f \sim g$ and this contradicts (1).

Second, we show that (2)⇒(1). Assume, per contra, that there exist $f, g \in H, s' \in S$ such that $f \sim g$, $f_s \succ_s g_s$ and $g_{s'} \succ_{s'} f_{s'}$ and for all $h\alpha \in (0,1)$,

$$\alpha h + (1-\alpha)f \sim \alpha h + (1-\alpha)g.$$

Let $\varphi \in \operatorname{int} u(H)$ be a differentiability point of I. Now there exist $h \in H$ with $u(h) \in \operatorname{int} u(H)$ and $\alpha^* \in (0,1)$ such that

$$u(\alpha^* h + (1-\alpha^*)g) = \varphi.$$

Now by the counter assumption, for all $\alpha \in [0,1]$,

$$\alpha^* h + (1-\alpha^*)g \sim \alpha^* h + (1-\alpha^*)\big(\alpha f + (1-\alpha)g\big).$$

Thus

$$\begin{aligned}
&(1-\alpha^*)\Delta I(\varphi) \cdot \big(u(f) - u(g)\big) \\
&= \lim_{\alpha \to 0} \frac{I[u(\alpha^* h + (1-\alpha^*)g) + \alpha(1-\alpha^*)(u(f) - u(g))] - I[u(\alpha^* h + (1-\alpha^*)g)]}{\alpha} \\
&= \lim_{\alpha \to 0} \frac{I[u(\alpha^* h + (1-\alpha^*)(\alpha f + (1-\alpha)g)] - I[u(\alpha^* h + (1-\alpha^*)g)]}{\alpha} = 0.
\end{aligned}$$

By Lemma A.55, $u(f) - u(g) \in \big(\bigcup_{\varphi \in \operatorname{int} u(H)} \partial I(\varphi)\big)^{\perp}$. But by Lemma A.31, $u_s(f_s) - u_s(g_s) \neq 0$ which is a contradiction with (2). □

The second lemma characterizes in terms of probabilities when the probabilities for a single state are identified.

Lemma A.77 Let $P \subseteq \Delta(S), P \neq \varnothing, s \in S$. The following two are equivalent:

1. For all $x \in P^\perp$, $x_s = 0$ and there exist $s' \in S, s \neq s', p \in P$ such that $p_{s'} > 0$.
2. There exist $p, q \in P$ such that $p_s \neq q_s$ and for all $\tilde{s} \in S \setminus \{s\}$,
$$\frac{p_{\tilde{s}}}{1-p_s} = \frac{q_{\tilde{s}}}{1-q_s}. \tag{A.83}$$

Proof. We show first that (2)⇒(1). First, by (2), $|P| \geq 2$ and so there exist $s' \in S, s \neq s', p \in P$ such that $p_{s'} > 0$. Assume, per contra, that there exists $x \in P^\perp$ such that $x_s \neq 0$. Let p, q be as in (2). Since $p_s \neq 1$ or $q_s \neq 1$, we have by (A.83) since $|S| \geq 2$,
$$p_s \neq 1 \text{ and } q_s \neq 1. \tag{A.84}$$
Since $p, q \in P^\perp$, we have
$$\sum_{\tilde{s} \in S} x_{\tilde{s}} p_{\tilde{s}} = 0 \text{ and } \sum_{\tilde{s} \in S} x_{\tilde{s}} q_{\tilde{s}} = 0.$$
By (A.84) and by multiplying the first equation by $(1-p_s)^{-1}$ and the second equation by $(1-q_s)^{-1}$, we have
$$\sum_{\tilde{s} \in S \setminus \{s\}} x_{\tilde{s}} \frac{p_{\tilde{s}}}{1-p_s} + x_s \frac{p_s}{1-p_s} = 0 = \sum_{\tilde{s} \in S \setminus \{s\}} x_{\tilde{s}} \frac{q_{\tilde{s}}}{1-q_s} + x_s \frac{q_s}{1-q_s}.$$
By (A.83), we have
$$x_s \frac{p_s}{1-p_s} = x_s \frac{q_s}{1-q_s} \xRightarrow{x_s \neq 0} \frac{p_s}{1-p_s} = \frac{q_s}{1-q_s}.$$
Now $x \mapsto \frac{x}{1-x}$ is strictly increasing function for $x \in [0, 1)$ and so by (A.84),
$$\frac{p_s}{1-p_s} = \frac{q_s}{1-q_s} \Rightarrow p_s = q_s.$$
This contradicts (2).

Next, we show that (1)⇒(2). Denote $n^\perp = \dim P^\perp$ and $n = \dim P$. Now $n^\perp + n = |S|$. There exist $(c^i)_{i=1}^{n^\perp} \subseteq P^\perp$ linearly independent vectors and $(p^i)_{i=1}^{n} \subseteq P$ linearly independent vectors.

First, we show that $n \geq 2$. Since P is non-empty, $n \geq 1$. Assume, per contra, $n = 1$. Thus P is a singleton $P = \{p\}$. If $p_s = 0$, then $(1_s, 0_{-s}) \in P^\perp$ which is a contradiction. By assumption, there exists $s' \in S, s' \neq s$ such that $p_{s'} > 0$. Now $((\frac{1}{p_{s'}})_s, (-\frac{1}{p_s})_{s'}, 0_{-s,s'}) \in P^\perp$ that is a contradiction. So $n \geq 2$.

Now we can consider the matrix C formed by row vectors $(c^i)_{i=1}^{n^\perp}$ and each column in $\tilde{s} \in S$. We can reduce C into Smith normal form $\tilde{C}$ formed by row vectors $(\tilde{c}^i)_{i=1}^{n^\perp}$ such that for each $i \in \{1, \ldots, n^\perp\}$, there exists $s^i \in S \setminus \{s\}$ such that $\tilde{c}^i_{s^i} = -1$, for all $j \in \{1, \ldots, n^\perp\} \setminus \{i\}$, $\tilde{c}^i_{s^j} = 0$, and for all $k, l \in \{1, \ldots, n^\perp\}$, $k \neq l$, $s^k \neq s^l$. Now $(\tilde{c}^i)_{i=1}^{n^\perp} \subseteq P^\perp$ are linearly independent and for all $i \in \{1, \ldots, n^\perp\}$, $\tilde{c}^i_s = 0$.

Denote $S^{\perp} = \{s^1 \ldots, s^{n^{\perp}}\}$. Now we have for all $i \in \{1, \ldots, n^{\perp}\}$ and $p \in P$,

$$p_{s^i} = \sum_{\tilde{s} \in S \setminus (S^{\perp} \cup \{s\})} \tilde{c}^i_{\tilde{s}} p^j_{\tilde{s}}. \tag{A.85}$$

Let $\pi : \{1, \ldots, n-1\} \to S \setminus (S^{\perp} \setminus \{s\})$ be a one-to-one function such that $\pi(n) = s$ and $\pi(n-1) = s^{\dagger}$. We show by induction that for each $i \in \{0, \ldots, n\}$, there exists a collection of linearly independent probabilities $(\tilde{p}^{j,i})_{j=1}^n$ such that for all $1 \leq m \leq i$, and $j, k \in \{1, \ldots, n\} \setminus \{m\}$,

$$p^{j,i}_{\pi(m)} = p^{k,i}_{\pi(m)} > p^{m,i}_{\pi(m)}. \tag{A.86}$$

This follows symmetrically to the proof in Lemma A.74 since at each step for $i < n-1$,

$$\min \{p^{j,i}_{\pi(i+1)} | j \in \{i+1, \ldots, n\}\} \neq \max \{p^{j,i}_{\pi(i+1)} | j \in \{i+1, \ldots, n\}\}$$

by linear independence and since for all $k \leq i$ and $j, l \in \{i+1, \ldots, n\}$, $p^{j,i}_{\pi(k)} = p^{l,i}_{\pi(k)}$.

Since $n \geq 2$, by the induction and by considering a convex combination of $(p^{j,n-1})_{j=1}^{n-1}$, there exist $a > 1$ and $p^*, p^{\dagger} \in P$ that are linearly independent and for all $\tilde{s} \in S \setminus (S^{\perp} \cup \{s\})$, $p^*_{\tilde{s}} = a p^{\dagger}_{\tilde{s}} > 0$. Let $s^{\dagger} \in S \setminus (S^{\perp} \cup \{s\})$. Now we have for all $\tilde{s} \in S \setminus (S^{\perp} \cup \{s\})$,

$$\frac{p^*_{\tilde{s}}}{p^*_{s^{\dagger}}} = \frac{a p^{\dagger}_{\tilde{s}}}{a p^{\dagger}_{s^{\dagger}}} = \frac{p^{\dagger}_{\tilde{s}}}{p^{\dagger}_{s^{\dagger}}}.$$

By (A.85), we have for all $i \in \{1, \ldots, n^{\perp}\}$ and $p \in P$,

$$\frac{p_{s^i}}{p_{s^{\dagger}}} = \sum_{\tilde{s} \in S \setminus (S^{\perp} \cup \{s\})} \tilde{c}^i_{\tilde{s}} \frac{p_{\tilde{s}}}{p_{s^{\dagger}}}.$$

Thus for all $\tilde{s} \in S \setminus \{s\}$,

$$\frac{p^*_{\tilde{s}}}{p^*_{s^{\dagger}}} = \frac{p^{\dagger}_{\tilde{s}}}{p^{\dagger}_{s^{\dagger}}}. \tag{A.87}$$

By taking a sum over $\tilde{s}$, we have

$$\frac{1 - p_s}{p^*_{s^{\dagger}}} = \frac{1 - p_s}{p^{\dagger}_{s^{\dagger}}}.$$

Thus by multiplying (A.87) by $\frac{p^*_{s^{\dagger}}}{1 - p_s}$, we have for all $\tilde{s} \in S \setminus \{s\}$,

$$\frac{p^*_{\tilde{s}}}{1 - p_s} = \frac{p^{\dagger}_{\tilde{s}}}{1 - p^{\dagger}_s}. \tag{A.88}$$

Finally, we show that $p^*_s \neq p^{\dagger}_s$. Assume, per contra, $p^*_s = p^{\dagger}_s$. Then by (A.88) for all $\tilde{s} \in S$, $p^*_{\tilde{s}} = p^{\dagger}_{\tilde{s}}$, which contradicts, that p^* and $p^{\dagger}$ are linearly independent. Thus $p^*_s \neq p^{\dagger}_s$ which shows the claim. □

The previous two lemmas give the probability dentification characterization.

Proposition A.78 (Probability Identification) Let $(u, \mathbb{P})$ be a state dependent tight dual-self representation for $\succsim$ and $s \in S^{\mathcal{P}}$. The following three conditions are equivalent:

1. If $f, g \in H$ are such that $f \sim g$, $f_s \succ_s g_s$, and there exists $s' \in S$ such that $g_{s'} \succ_{s'} f_{s'}$, then there exist $h \in H$ and $\alpha \in (0,1)$ such that
$$\alpha h + (1-\alpha) f \not\sim \alpha h + (1-\alpha) g.$$
2. $S^{\mathcal{P}} = \{s\}$ or there exist $p, q \in \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P$ such that $p_s \neq q_s$ and for all $\tilde{s} \in S \setminus \{s\}$,
$$\frac{p_{\tilde{s}}}{1-p_s} = \frac{q_{\tilde{s}}}{1-q_s}.$$
3. If $(\widetilde{u}, \widetilde{\mathbb{P}})$ is a state dependent tight dual-self representation for $\succsim$, then
$$\left\{ \widetilde{p}_s \middle| \widetilde{p} \in \overline{\text{co}} \bigcup_{\widetilde{P} \in \widetilde{\mathbb{P}}} \widetilde{P} \right\} = \left\{ p_s \middle| p \in \overline{\text{co}} \bigcup_{P \in \mathbb{P}} P \right\}.$$

Proof. (1) $\iff$ (2): We have two cases. If $S^{\mathcal{P}} = \{s\}$, then $\bigcup_{P \in \mathbb{P}} P = \{\delta_s\}$ and so the claim follows Lemma A.76 and since for all $x \in \mathbb{R}^S$, $\delta_s \cdot x = 0$ iff $x_s = 0$. If $S^{\mathcal{P}} \neq \{s\}$, then there exist $s' \in S, s \neq s'$ and $p \in \bigcup_{P \in \mathbb{P}} P$ such that $p_{s'} > 0$. Thus the equivalence follows from Lemmas A.76 and A.77. (2) $\iff$ (3) follows from Corollary A.62 and Lemma A.77 □

A.5 Social Choice Theory: Symmetric State Dependent MaxMin

We first show that non-negative symmetric set is either full dimensional or a singleton.

Lemma A.79 Let $P \subseteq \mathbb{R}^S_+$ be non-empty and $\bar{0} \notin P$ such that if $p, q \in P$ are such that $p \neq q$, then for all $a > 0$, $ap \neq q$ and for all permutations $\pi : S \to S$ and $p \in P$,
$$(p_{\pi(s)})_{s \in S} \in P. \tag{A.89}$$
Then there exists $c \in R_{++}$ such that $P = \{(c)_{s \in S}\}$ or $P^\perp = \{\bar{0}\}$.

Proof. We will show that there does not exist $p \in P, q \in P^\perp, s_1, s_2, s_3, s_4 \in S$ such that $p_{s_1} \neq p_{s_2}$ and $q_{s_3} \neq q_{s_4}$. Let $\pi^1 : S \to S$ and $\pi^2 : S \to S$ be two permutations such that $\pi^1(s_1) = s_3$, $\pi^1(s_2) = s_4$, $\pi^2(s_1) = s_4$, $\pi^2(s_2) = s_3$, and for all $s \notin \{s_1, s_2\}$, $\pi^1(s) = \pi^2(s)$. Now by (A.89) and $q \in P^\perp$, we have
$$\sum_{s \in S} p_{(\pi^1)^{-1}(s)} q_s = 0 = \sum_{s \in S} p_{(\pi^2)^{-1}(s)} q_s.$$
Since for all $s \notin \{s_3, s_4\}$, $(\pi^1)^{-1}(s) = (\pi^2)^{-1}(s)$, we have
$$p_{s_1} q_{s_3} + p_{s_2} q_{s_4} = p_{s_2} q_{s_3} + p_{s_1} q_{s_4} \Rightarrow (p_{s_1} - p_{s_2})(q_{s_4} - q_{s_3}) = 0.$$
Thus $p_{s_1} = p_{s_2}$ or $q_{s_4} = q_{s_3}$.

First, assume that P is a singleton with $p \in P$. Then by (A.89) and since P is a singleton, for all $s, s' \in S$, $p_s = p_{s'}$. This shows the claim since $\bar{0} \notin P$.

Assume next that P is not a singleton. Thus there exist $p, q \in P$ such that $p \neq q$. Now for all $a \in \mathbb{R}_{++}$, $ap \neq q$. Especially, there exist $p \in P, s_1, s_2 \in S$ such that $p_{s_1} \neq p_{s_2}$. Thus by above, for all $q \in P^{\perp}, s_3, s_4 \in S$, $q_{s_3} = q_{s_4}$. Since for all $s \in S$, $p_s \geq 0$ and for some $s' \in S$, $p_{s'} > 0$, this is possible only if $P^{\perp} = \{\bar{0}\}$. This shows the claim. □

The above lemma gives almost the characterization for symmetric state dependent maxmin. The next lemma shows that our behavioral symmetry assumption, Axiom 1.9, characterizes a symmetric set of probabilities. This shows the characterization.

Proposition A.80 (Symmetric State Dependent MaxMin) $\succsim$ satisfies Axioms 1-1.6, 1.8, 1.9 if and only if (1) or (2) of the following conditions holds:

1. For all $p \in \operatorname{int} \Delta(S)$, there exists $u = (u_s)_{s \in S}$ such that for all $s \in S$, $u_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\operatorname{int} u_s(\Delta(X_s)) \neq \varnothing$ and for all $f, g \in H$,
$$f \succsim g \iff \sum_{s \in S} p_s u_s(f_s) \geq \sum_{s \in S} p_s u_s(g_s).$$
2. There exists (u, P) that is a state dependent maxmin representation for $\succsim$ such that $\operatorname{int} P \neq \varnothing$, for all $p \in P$, permutations $\pi : S \to S$, $(p_{\pi(s)})_{s \in S} \in P$, and
$$\operatorname{int} \bigcap_{s \in S} u_s(\Delta(X_s)) \neq \varnothing.$$

Additionally, if (2) holds and $(\widetilde{u}, \widetilde{P})$ is another state dependent maxmin representation for $\succsim$, then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that
$$\widetilde{P} = P \text{ and for all } s \in S,\ \widetilde{u}_s = \alpha u_s + \beta.$$

Proof. We show first the if direction.

By an assumption, $\succsim$ satisfies Axioms 1-1.6,1.8, there exists a concave state dependent Boolean representation for $\succsim$. This corresponds to a state dependent maxmin representation (P, u) with probabilities $P \subseteq \Delta(S)$ and affine utility $u = (u_s)_{s \in S}$.

If $|S| = 1$, then by non-triviality, (1) holds. So assume that $|S| > 1$.

Since $f^* \succ g^*$, there exists $s^* \in S^{\mathcal{P}}$ such that $u_{s^*}(f^*_{s^*}) > u_{s^*}(g^*_{s^*})$. By the maxmin representation, $s^* \in S^{\mathcal{P}}$, and Lemma A.67, since the set of crisp acts is convex, we have
$$\frac{1}{|S|-1} f^* + \frac{|S|-2}{|S|-1} g^* \succ \frac{1}{|S|-1}(g^*_{s^*}, f^*_{-s^*}) + \frac{|S|-2}{|S|-1} g^* = \Big(g^*_{s^*}, \frac{1}{|S|-1} f^*_{-s^*} + \frac{|S|-2}{|S|-1} g^*_{-s^*}\Big).$$

By Axiom 1.9, for all $s \in S$,

$$\frac{1}{|S|-1}f^* + \frac{|S|-2}{|S|-1}g^* \succ \Big(g_s^*, \frac{1}{|S|-1}f_{-s}^* + \frac{|S|-2}{|S|-1}g_{-s}^*\Big).$$

Thus $S = S^{\mathcal{P}}$ and for all $s \in S$, $u_s(f_s^*) > u_s(g_s^*)$. Denote for each $s \in S$,

$$\tau_s = u_s(f_s^*) - u_s(g_s^*) \in \mathbb{R}_{++}. \tag{A.90}$$

Next assume that $h \in H, u(h) \in \operatorname{int} u(H)$ is such that I is differentiable at $u(h)$. Let $\pi : S \to S$ be a permutation. We will show that $\big(\nabla I(u(h))_{\pi(s)}\big)_{s\in S} \in P$.

First, since $|S| > 1$, there exists $\alpha^1 \in (0,1)$ such that for all $s \in S$,

$$u_s(f_s^*) > \alpha^1 u_s(\tfrac{3}{4}f_s^* + \tfrac{1}{4}g_s^*) + (1-\alpha^1)u_s(h_s) > \frac{1}{|S|}u_s(f_s^*) + \frac{|S|-1}{|S|}u_s(g_s^*).$$

Thus there exists $\theta \in (0,1)^S$ such that for each $s \in S$,

$$u_s(\theta_s f_s^* + (1-\theta_s)g_s^*) = \alpha^1 u_s(\tfrac{3}{4}f_s^* + \tfrac{1}{4}g_s^*) + (1-\alpha^1)u_s(h_s).$$

Now especially $\sum_{s\in S}\theta_s > 1$ and so $\alpha^2 = \big(\sum_{s\in S}\theta_s\big)^{-1} \in (0,1)$. Denote

$$\hat{h} = \alpha^2\alpha^1(\tfrac{3}{4}f^* + \tfrac{1}{4}g^*) + (1-\alpha^1)u_s(h_s) + (1-\alpha^2)g^*.$$

By Lemma A.66, I is differentiable at $u(\hat{h})$ and $\nabla I(\hat{h}) = \nabla I(h)$. Denote for all $s \in S$, $\hat{\gamma}_s = \theta_s\alpha^2$. Now $\sum_{s\in S}\gamma_s = 1$. Especially, I is differentiable at $\big(\gamma_s u_s(f_s^*) + (1-\gamma_s)u_s(g_s^*)\big)_{s\in S}$ with derivative $\nabla I(h)$.

We will show that I is differentiable at $\big(\gamma_{\pi(s)}u_s(f_s^*) + (1-\gamma_{\pi(s)})u_s(g_s^*)\big)_{s\in S}$ with a derivative $\big(\nabla I(u(h))_{\pi(s)}\big)_{s\in S}$.

Let $v \in \mathbb{R}^S$ and $a > 0$ be such that for all $s \in S$,

$$u_s(f_s^*) > u_s(\hat{h}_s) + av_s > u_s(g_s^*).$$

Thus there exists $\tilde{\theta} \in (0,1)^S$ such that for each $s \in S$,

$$u_s(\tilde{\theta}_s f_s^* + (1-\tilde{\theta}_s)g_s^*) = u_s(\hat{h}) + av_s.$$

Now for each $s \in S$,

$$\tilde{\theta}_s - \gamma_s = av_s(\tau_s^{-1}).$$

Assume w.l.o.g. $\sum_{s\in S} v_s \geq 0$ since the other case follows similarly. Now $\sum_{s\in S}\tilde{\theta}_s \geq 1$. Denote $\tilde{\alpha} = (\sum_{s\in S}\tilde{\theta}_s)^{-1}$. Now $\tilde{\alpha}\tilde{\theta} \in \Delta(S)$. Thus by Axiom 1.9,

$$\begin{aligned}&\tilde{\alpha}\big(\tilde{\theta}_s f_s^* + (1-\tilde{\theta}_s)g_s^*\big)_{s\in S} + (1-\tilde{\alpha})g^* = \big(\tilde{\alpha}\tilde{\theta}_s f_s^* + (1-\tilde{\alpha}\tilde{\theta}_s)g_s^*\big)_{s\in S}\\ &\sim \big(\tilde{\alpha}\tilde{\theta}_{\pi(s)} f_s^* + (1-\tilde{\alpha}\tilde{\theta}_{\pi(s)})g_s^*\big)_{s\in S} = \tilde{\alpha}\big(\tilde{\theta}_{\pi(s)} f_s^* + (1-\tilde{\theta}_{\pi(s)})g_s^*\big)_{s\in S} + (1-\tilde{\alpha})g^*.\end{aligned}$$

Since g^* is crisp, we have

$$\left(\tilde{\theta}_s f_s^* + (1-\tilde{\theta}_s)g_s^*\right)_{s\in S} \sim \left(\tilde{\theta}_{\pi(s)} f_s^* + (1-\tilde{\theta}_{\pi(s)})g_s^*\right)_{s\in S}.$$

Now

$$\left(\theta_{\pi(s)}u_s(f_s^*) + (1-\theta_{\pi(s)})u_s(g_s^*)\right)_{s\in S} = \left(\gamma_{\pi(s)}u_s(f_s^*) + (1-\gamma_{\pi(s)})u_s(g_s^*)\right)_{s\in S} + (av_{\pi(s)}\tau_{\pi(s)}^{-1}\tau_s)_{s\in S}.$$

By the representation,

$$I\big(u(\hat{h}) + av\big) = I\Big(\left(\gamma_{\pi(s)}u_s(f_s^*) + (1-\gamma_{\pi(s)})u_s(g_s^*)\right)_{s\in S} + (av_{\pi(s)}\tau_{\pi(s)}^{-1}\tau_s)_{s\in S}\Big)$$

Thus especially, by subtracting $I\big(u(\hat{h})\big) = I\big[\left(\gamma_{\pi(s)}u_s(f_s^*) + (1-\gamma_{\pi(s)})u_s(g_s^*)\right)_{s\in S}\big]$ from both hand sides, dividing by a and taking $a \to 0$, we have

$$\nabla I(u(h))\cdot v = \lim_{a\to 0} \frac{I\Big(\left(\gamma_{\pi(s)}u_s(f_s^*)+(1-\gamma_{\pi(s)})u_s(g_s^*)+av_{\pi(s)}\tau_{\pi(s)}^{-1}\tau_s\right)_s\Big) - I\Big(\left(\gamma_{\pi(s)}u_s(f_s^*)+(1-\gamma_{\pi(s)})u_s(g_s^*)\right)_s\Big)}{a}.$$

Since v was arbitrary, I is differentiable at $\left(\gamma_{\pi(s)}u_s(f_s^*) + (1-\gamma_{\pi(s)})u_s(g_s^*)\right)_{s\in S}$ and

$$\nabla I\Big(\left(\gamma_{\pi(s)}u_s(f_s^*) + (1-\gamma_{\pi(s)})u_s(g_s^*)\right)_{s\in S}\Big) = \Big(\frac{\tau_{\pi^{-1}(s)}}{\tau_s}\nabla I(u(h))_{\pi^{-1}(s)}\Big)_{s\in S}.$$

That is for all $p \in P$,

$$p = \Big(\frac{\tau_{\pi(s)}}{\tau_s}p_{\pi(s)}\Big)_{s\in S} \iff (\tau_s p_s)_{s\in S} = (\tau_{\pi(s)}p_{\pi(s)})_{s\in S}. \tag{A.91}$$

Define the set

$$\tau P = \{(\tau_s p_s)_{s\in S} | p \in P\}.$$

By (A.90,A.91), $\tau P \subseteq \mathbb{R}_{++}$ and for all permutations $\pi : S \to S$ and $p \in P$, $(\tau_{\pi(s)}p_{\pi(s)})_{s\in S} \in \tau P$. Additionally, let $p, q \in P, p \neq q$ and assume, per contra, that there exists $a > 0$ such that $(\tau_s p_s)_{s\in S} = (a\tau_s q_s)_{s\in S}$. Then $p = aq$, which is a contradiction since p and q are different probabilities that sum to 1.

By Lemma A.79, there exists $c \in \mathbb{R}_{++}$ such that $\tau P = \{(c)_{s\in S}\}$ or $(\tau P)^\perp = \{\bar{0}\}$.

We first consider the case that there exists $c \in \mathbb{R}_{++}$ such that $\tau P = \{(c)_{s\in S}\}$ and show that in this case (1) holds. This follows from the observation that since τP is a singleton, P is a singleton. Thus the claim follows from $\tau \in \mathbb{R}_{++}^S$ and Theorem 1.5.

Second, we consider the case that $(\tau P)^\perp = \{\bar{0}\}$ and show that (2) holds. First, $|S| > 1$ and so the first case does not hold. First, we show that $P^\perp = \{\bar{0}\}$. Assume, per contra there exists $x \in P^\perp, x \neq \bar{0}$. Now for all $p \in P$, since $\tau \in \mathbb{R}^S_{++}$,

$$0 = x \cdot p = \sum_{s \in S} \frac{x_s}{\tau_s} \tau_s p_s.$$

So $(\frac{x_s}{\tau_s})_{s \in S} \in (\tau P)^\perp$ and by (A.90) $(\frac{x_s}{\tau_s})_{s \in S} \neq \bar{0}$ which is a contradiction. Thus $P^\perp = \{\bar{0}\}$. By Lemma A.67, for all $s \in S$,

$$u_s(f^*_s) = I\big(u(f^*)\big) \text{ and } u_s(g^*_s) = I\big(u(g^*)\big).$$

Thus especially for all $s \in S$, $\tau_s = I\big(u(f^*)\big) - I\big(u(g^*)\big)$. By (A.91), for all $p \in P$ and permutations $\pi : S \to S$, $(p_{\pi(s)})_{s \in S} \in P$. By the proof of Lemma A.65, $\operatorname{int} P \neq \varnothing$. This shows (2) and the additional claim follows from Theorem 1.2 since if $\operatorname{int} \bigcap_{s \in S} u_s\big(\Delta(X_s)\big) \neq \varnothing$, then every state dependent maxmin representation is tight.

Next, we show the only if direction. By Theorem 1.1, we only need to show Axiom 1.9. First, assume that (1) holds. Then $\succsim$ satisfies independence axiom and every act is crisp. Let $p \in \operatorname{int} \Delta(S)$ and let $(u_s)_{s \in S}$ be the corresponding utility functions. Since for all $s \in S$, $\operatorname{int} u_s(\Delta(X_s)) \neq \varnothing$, there exist $g^*, f^* \in H, a > 0$ such that for all $s \in S$, $u_s(f^*_s) = u_s(g^*_s) + \frac{a}{p_s}$. Let $\gamma \in \Delta(S)$ and $\pi : S \to S$ be a permutation. Now we have

$$\sum_{s \in S} p_s u_s(\gamma_s f^*_s + (1-\gamma_s) g^*_s) = \sum_{s \in S} p_s \Big(u_s(g^*_s) + \frac{\gamma_s a}{p_s} \Big) = \sum_{s \in S} p_s u_s(g^*_s) + a \sum_{s \in S} \gamma_{s'}$$
$$= \sum_{s \in S} p_s u_s(g^*_s) + a \sum_{s' \in S} \gamma_{\pi(s')} = \sum_{s \in S} p_s \Big(u_s(g^*_s) + \frac{a \gamma_{\pi(s)}}{p_s} \Big) = \sum_{s \in S} p_s u_s(\gamma_{\pi(s)} f^*_s + (1-\gamma_{\pi(s)}) g^*_s).$$

This shows Axiom 1.9.

Second, assume that (2) holds. Since $\operatorname{int} \bigcap_{s \in S} u_s\big(\Delta(X_s)\big) \neq \varnothing$, there exist $f^*, g^* \in H, c^f, c^g \in \mathbb{R}, c^f > c^g$ such that for all $s \in S$, $u_s(f^*_s) = c^f$ and $u_s(g^*_s) = c^g$. As constant utility acts f^*, g^* are crisp acts. First, for all permutations $\pi : S \to S$ and $\gamma \in \Delta(S)$, we have

$$p \cdot u\Big((\gamma_s f^*_s + (1-\gamma_s) g^*_s)_{s \in S} \Big) = \sum_{s \in S} p_s \cdot c^g + \sum_{\tilde{s} \in S} p_s \gamma_s (c^f - c^g) \tag{A.92}$$
$$= \sum_{s \in S} p_{\pi(s)} \cdot c^g + \sum_{\tilde{s} \in S} p_{\pi(s)} \gamma_{\pi(s)} (c^f - c^g) = (p_{\pi(s)})_{s \in S} \cdot u\Big((\gamma_{\pi(s)} f^*_s + (1-\gamma_{\pi(s)}) g^*_s)_{s \in S} \Big).$$

Let $\gamma \in \Delta(S)$ and permutation $\pi : S \to S$. First, we show that

$$p \in \arg\min_{q \in P} q \cdot u\Big((\gamma_s f^*_s + (1-\gamma_s) g^*_s)_{s \in S} \Big) \Rightarrow (p_{\pi(s)})_s \in \arg\min_{q \in P} q \cdot u\Big((\gamma_{\pi(s)} f^*_s + (1-\gamma_{\pi(s)}) g^*_s)_s \Big). \tag{A.93}$$

Assume, per contra, that there exists $q \in P$ such that

$$q \cdot u\Big(\big(\gamma_{\pi(s)} f^*_s + (1-\gamma_{\pi(s)}) g^*_s\big)_{s\in S}\Big) < (p_{\pi(s)})_{s\in S} \cdot u\Big(\big(\gamma_{\pi(s)} f^*_s + (1-\gamma_{\pi(s)}) g^*_s\big)_{s\in S}\Big).$$

Then by Equation (A.92) and by using the permutation π^{-1},

$$(q_{\pi^{-1}(s)})_{s\in S} \cdot u\Big(\big(\gamma_s f^*_s + (1-\gamma_s) g^*_s\big)_{s\in S}\Big) < p \cdot u\Big(\big(\gamma_s f^*_s + (1-\gamma_s) g^*_s\big)_{s\in S}\Big)$$

which is a contradiction by the symmetry of P.

Now we have by (A.92,A.93),

$$\min_{p\in P} p \cdot u\Big(\big(\gamma_s f^*_s + (1-\gamma_s) g^*_s\big)_{s\in S}\Big) = \min_{p\in P} p \cdot u\Big(\big(\gamma_{\pi(s)} f^*_s + (1-\gamma_{\pi(s)}) g^*_s\big)_{s\in S}\Big)$$

which shows Axiom 1.9.

□

A.6 State Dependent Bewley Representation

This section axiomatizes the existence of a state dependent Bewley representation and characterizes its uniqueness.

First, we present the five characterizing axioms. The first axiom is the standard assumption that $\succsim$ is a nontrivial partial order.

Axiom 1$^{\mathbf{B}}$ $\succsim$ is nontrivial, reflexive and transitive.

The next one is the continuity axiom as before.

Axiom 2$^{\mathbf{B}}$ For all $f, g, h \in H$, the sets $\{\alpha \in [0,1] | \alpha f + (1-\alpha) g \succsim h\}$ and $\{\alpha \in [0,1] | h \succsim \alpha f + (1-\alpha) g\}$ are closed in $\mathbb{R}$.

The next axiom assumes that there is no uncertainty about tastes for the consequences but instead when the acts differ only in one state they can always be compared. This also assumes that there is no ambiguity about a single state.

Axiom 3$^{\mathbf{B}}$ For all $s \in S, f \in H, x_s, y_s \in \Delta(X_s)$,

$$(x_s, f_{-s}) \succsim (y_s, f_{-s}) \text{ or } (y_s, f_{-s}) \succsim (x_s, f_{-s}).$$

The next axiom is the second part of the assumption that there is no ambiguity about a single state. It assumes that preferences over consequences in a single state does not depend on what

common act is received in other states. This axiom guarantees that statewise preferences are well-defined.

Axiom 4$^{\mathbf{B}}$ For all $s \in S, f, g \in H, x_s, y_s \in \Delta(X_s)$,

$$(x_s, f_{-s}) \succsim (y_s, f_{-s}) \Longrightarrow (x_s, g_{-s}) \succsim (y_s, g_{-s}).$$

The last axiom is the standard independence axiom.

Axiom 5$^{\mathbf{B}}$ For all $f, g \in H, \alpha \in (0, 1)$,

$$f \succsim g \Longleftrightarrow \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h.$$

The next result states that the previous five axioms characterize the existence of a state dependent Bewley representation and characterizes the uniqueness of the representation. The uniqueness part states that every statewise positive affine transformations for the utilities are allowed but for each utility function there is a unique set of probabilities.

Theorem A.81 (Bewley Representation) $\succsim$ satisfies Axioms 1^{B}-5^{B} if and only if there exists (u, C) that is a state dependent Bewley representation for $\succsim$.

Additionally, let (u, C) be a state dependent Bewley representation for $\succsim$, $\widetilde{u} = (\widetilde{u}_s)_{s \in S}$ be such that for all $s \in S$, $\widetilde{u}_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\widetilde{C} \subseteq \Delta(S)$ be closed and convex. Then $(\widetilde{u}, \widetilde{C})$ is a state dependent Bewley representation for $\succsim$ if and only if there exist $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that

$$\widetilde{C} = \left\{ \left(\frac{a_s^{-1} p_s}{\sum_{s' \in S} a_{s'}^{-1} p_{s'}} \right)_{s \in S} \middle| p \in C \right\}.$$

and for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = a_s u_s + b_s.$$

Next, we consider the relation to state independent Bewley representation. Assume a state independent setting with for all $s \in S$, $X_s = X$. We show that the only difference between state independent and state dependent Bewley representations is the following monotonicity with between state comparisons.

Axiom 6$^{\mathbf{B}}$ For all $s, s' \in S, f, g \in H, x, y \in \Delta(X)$,

$$(x_s, f_{-s}) \succ (y_s, f_{-s}) \Longrightarrow (x_{s'}, g_{-s'}) \succsim (y_{s'}, g_{-s'}),$$

The next result shows that this gives state independent Bewley representation with the same utility in every state of the world. Additionally, we show that any statewise positive affine transformation of utilities gives an alternative state dependent Bewley representation.

Theorem A.82 (State Independent Bewley Representation) $\succsim$ satisfies Axioms 1^{B}-6^{B} if and only if there exists (u, C) that is a state independent Bewley representation for $\succsim$.

Additionally, let (u, C) be a state independent Bewley representation for $\succsim$, $\widetilde{u} = (\widetilde{u}_s)_{s\in S}$ be such that for all $s \in S$, $\widetilde{u}_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\widetilde{C} \subseteq \Delta(S)$ be closed and convex. Then $(\widetilde{u}, \widetilde{C})$ is a state dependent Bewley representation for $\succsim$ if and only if there exist $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that

$$\widetilde{C} = \left\{ \left(\frac{a_s^{-1} p_s}{\sum_{s' \in S} a_{s'}^{-1} p_{s'}} \right)_{s\in S} \middle| p \in C \right\}.$$

and for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = a_s u + b_s.$$

A.6.1 Bewley Existence Proof

For all $s \in S$, define $\succsim_s$ and $\succ_s$ as in Section 1.2.3. Our first lemma shows that statewise preferences have an expected utility representation.

Lemma A.83 Let $\succsim$ satisfy Axioms 1^{B}-5^{B}. Then for each $s \in S$, there exists $u_s : \Delta(X_s) \to \mathbb{R}$ that is affine such that for all $x_s, y_s \in \Delta(X_s)$,

$$x_s \succsim_s y_s \iff u_s(x_s) \geq u_s(y_s).$$

Proof. Let $s \in S$. We will show that $\succsim_s$ is complete, transitive, mixture continuous, and satisfy Independence. Let $x_s, y_s, z_s \in \Delta(X_s), f, g \in H$.

1. Complete: By Axiom 3^{B}, either $(x_s, f_{-s}) \succsim (y_s, f_{-s})$ or $(y_s, f_{-s}) \succsim (x_s, f_{-s})$. Hence, by Axiom 4^{B}, respectively either for all $g \in H$, $(x_s, g_{-s}) \succsim (y_s, g_{-s})$ or for all $g \in H$, $(y_s, g_{-s}) \succsim (x_s, g_{-s})$. Thus by the definition, either $x_s \succsim_s y_s$ or $y_s \succsim_s x_s$ respectively.
2. Transitive: Let $x_s \succsim_s y_s$ and $y_s \succsim_s z_s$. By the definition, for all $f \in H$, $(x_s, f_{-s}) \succsim (y_s, f_{-s})$ and $(y_s, f_{-s}) \succsim (z_s, f_{-s})$. Thus by Axiom 1^{B}, for all $f \in H$, $(x_s, f_{-s}) \succsim (y_s, f_{-s})$ and hence $x_s \succsim_s y_s$.
3. Mixture Continuous: Let $x_s \succsim_s y_s$ and $y_s \succsim_s z_s$. By definition, there exist $f, g \in H$ such that $(x_s, f_{-s}) \succ (y_s, f_{-s}) \succsim (z_s, f_{-s})$ and $(x_s, g_{-s}) \succsim (y_s, g_{-s}) \succ (z_s, g_{-s})$. By Axiom 2^{B}, the sets $\{\alpha \in [0,1] | (y_s, f_{-s}) \succsim (\alpha x_s + (1-\alpha) z_s, f_{-s})\}$ and $\{\alpha \in [0,1] | (\alpha x_s + (1-\alpha) z_s, g_{-s}) \succsim (y_s, g_{-s})\}$ are closed and hence by Axiom 3^{B}, the sets $\{\alpha \in [0,1] | (\alpha x_s + (1-\alpha) z_s, f_{-s}) \succ (y_s, f_{-s})\}$ and $\{\alpha \in [0,1] | (y_s, g_{-s}) \succ (\alpha x_s + (1-\alpha) z_s, g_{-s})\}$ are open as the complements of the previous sets.

Additionally, they are nonempty since $1 \in \{\alpha \in [0,1] | (\alpha x_s + (1-\alpha) z_s, f_{-s}) \succ (y_s, f_{-s})\}$ and $0 \in \{\alpha \in [0,1] | (y_s, g_{-s}) \succ (\alpha x_s + (1-\alpha) z_s, g_{-s})\}$. Thus, there exist $\alpha, \beta \in (0,1)$ such that

$$(\alpha x_s + (1-\alpha) z_s, f_{-s}) \succ (y_s, f_{-s}) \text{ and } (y_s, g_{-s}) \succ (\beta x_s + (1-\beta) z_s, g_{-s}).$$

Additionally, by Axiom 4^{B}, for all $h \in H$, $(\alpha x_s + (1-\alpha) z_s, h_{-s}) \succsim (y_s, h_{-s})$ and $(y_s, h_{-s}) \succsim (\beta x_s + (1-\beta) z_s, h_{-s})$. Thus by the definition of $\succsim_s$, $\alpha x_s + (1-\alpha) z_s \succ_s y_s$ and $y_s \succ_s \beta x_s + (1-\beta) z_s$.

4. Independence: Let $\alpha \in (0,1)$. For all $f \in H$, by Axiom 5^{B},

$$(x_s, f_{-s}) \succsim (y_s, f_{-s}) \iff (\alpha x_s + (1-\alpha) z_s, f_{-s}) \succsim (\alpha y_s + (1-\alpha) z_s, f_{-s}).$$

Thus by the definition of $\succsim_s$,

$$x_s \succsim_s y_s \iff \alpha x_s + (1-\alpha) z_s \succsim_s \alpha y_s + (1-\alpha) z_s.$$

By von Neumann–Morgenstern utility theorem (Fishburn, 1970), there exists affine function $u_s : \Delta(X_s) \to \mathbb{R}$ such that for all $x_s, y_s \in \Delta(X_s)$,

$$x_s \succsim_s y_s \iff u_s(x_s) \geq u_s(y_s).$$

□

Next, we consider the induced preferences in the utility space. Additionally, we extend these induced preferences linearly to $\mathbb{R}^{S^{\mathcal{P}}}$. In this case, they have a state independent Bewley representation. We extend the preferences linearly starting from some interior act h^*. Let $h^* \in H$ be such that for all $s \in S^{\mathcal{P}}$, there exist $x_s, y_s \in \Delta(X_s)$ such that $x_s \succ_s h^*_s \succ_s y_s$. Define the extended preferences in the utility space $\widehat{\succsim}$ on $\mathbb{R}^{S^{\mathcal{P}}}$ by for all $\varphi, \psi \in \mathbb{R}^{S^{\mathcal{P}}}$,

$$\varphi \,\widehat{\succsim}\, \psi \iff \exists\, \alpha \in (0,1), f, g \in H, \alpha\varphi + (1-\alpha) u(h^*) = u(f), \alpha\psi + (1-\alpha) u(h^*) = u(g), f \succsim g,$$

where we embed φ to $\mathbb{R}^S$ as $\varphi_s = u_s(h^*_s)$ for all $s \notin S^{\mathcal{P}}$.

The next lemma shows that in the definition of $\widehat{\succsim}$, we can use for all quantifier for α by linearity.

Lemma A.84 Let $\succsim$ satisfy Axioms 1^{B}-5^{B}. For all $\varphi, \psi \in \mathbb{R}^{S^{\mathcal{P}}}$,

$$\varphi \,\widehat{\succsim}\, \psi \iff$$

$$[\forall\, \alpha \in (0,1], f, g \in H, \alpha\varphi + (1-\alpha) u(h^*) = u(f), \alpha\psi + (1-\alpha) u(h^*) = u(g) \Rightarrow f \succsim g].$$

Proof. First, we show the if direction. Let $\varphi \,\widehat{\succsim}\, \psi, \alpha \in (0,1], f, g \in H$ be such that $\alpha\varphi + (1-\alpha) u(h^*) = u(f), \alpha\psi + (1-\alpha) u(h^*) = u(g)$. We show that $f \succsim g$. Since $\varphi \,\widehat{\succsim}\, \psi$, there exist $\alpha^* \in (0,1), f^*, g^* \in H$ such that $f^* \succsim g^*$ and $\alpha^*\varphi + (1-\alpha^*) u(h^*) = u(f^*), \alpha^*\psi + (1-\alpha^*) u(h^*) = u(g^*)$.

We consider two cases: 1) $\alpha \geq \alpha^*$: Now by the linearity of u, $u\big({}^{\alpha^*}\!/_{\alpha} f + (1 - {}^{\alpha^*}\!/_{\alpha})h^*\big) = u(f^*)$ and $\big({}^{\alpha^*}\!/_{\alpha} g + (1 - {}^{\alpha^*}\!/_{\alpha})h^*\big) = u(g^*)$. Since u_s represents $\succsim_s$ and by Lemma A.13,

$$ {}^{\alpha^*}\!/_{\alpha} f + (1 - {}^{\alpha^*}\!/_{\alpha})h^* \sim f^* \succsim g^* \sim {}^{\alpha^*}\!/_{\alpha} g + (1 - {}^{\alpha^*}\!/_{\alpha})h^*.$$

Thus by Axioms 1^{B} and 5^{B}, $f \succsim g$.

2) $\alpha^* > \alpha$: Now by the linearity of u, $u\big({}^{\alpha}\!/_{\alpha^*} f^* + (1 - {}^{\alpha}\!/_{\alpha^*})h^*\big) = u(f)$ and $\big({}^{\alpha}\!/_{\alpha^*} g^* + (1 - {}^{\alpha}\!/_{\alpha^*})h^*\big) = u(g)$. Since u_s represents $\succsim_s$ and by Lemma A.13 and Axiom 5^{B},

$$ f \sim {}^{\alpha}\!/_{\alpha^*} f^* + (1 - {}^{\alpha}\!/_{\alpha^*})h^* \succsim {}^{\alpha}\!/_{\alpha^*} g^* + (1 - {}^{\alpha}\!/_{\alpha^*})h^* \sim g.$$

Thus by Axiom 1^{B}, $f \succsim g$.

Next, we show the only if direction. Now $\varnothing \neq \{\alpha' \in (0,1] | \alpha' \varphi + (1-\alpha') u(h^*), \alpha' \psi + (1-\alpha') u(h^*) \in H\}$ since $u(h^*) \in \operatorname{int} \operatorname{pr}_{S^{\mathcal{P}}} u(H)$, hence there exist $\alpha \in (0,1), f, g \in H$ such that $\alpha\varphi + (1-\alpha)u(h^*) = u(f), \alpha\psi + (1-\alpha)u(h^*) = u(g)$. By assumption, $f \succsim g$ and hence by the definition, $\varphi \mathrel{\widehat{\succsim}} \psi$.

□

The next lemma shows that $\widehat{\succsim}$ satisfies the axioms for state independent Bewley representation.

Lemma A.85 Let $\succsim$ satisfy Axioms 1^{B}-5^{B}. $\widehat{\succsim}$ is reflexive, nontrivial, transitive, continuous, Independence, and monotonic.

Proof.

1. Nontrivial: By Axiom 1^{B}, there exist $f, g \in H$ such that $f \succ g$ and hence by Lemma A.84, $\operatorname{pr}_{S^{\mathcal{P}}} u(f) \mathrel{\widehat{\succ}} \operatorname{pr}_{S^{\mathcal{P}}} u(g)$.
2. Reflexive: Let $\varphi \in \mathbb{R}^{S^{\mathcal{P}}}$. Since $u(h^*) \in \operatorname{int} u(H)$, there exist $\alpha \in (0,1), f \in H$ such that $\alpha\varphi + (1-\alpha)u(h^*) = u(f)$. Since $\succsim$ is reflexive, $f \succsim f$ and hence $\varphi \mathrel{\widehat{\succsim}} \varphi$.
3. Transitive: Let $\varphi \mathrel{\widehat{\succsim}} \psi$ and $\psi \mathrel{\widehat{\succsim}} \theta$. Let $\alpha^1, \alpha^2 \in (0,1], f, g, g', h \in H$ be such that $\alpha^1\varphi + (1-\alpha^1)u(h^*) = u(f), \alpha^1\psi + (1-\alpha^1)u(h^*) = u(g), \alpha^2\psi + (1-\alpha^2)u(h^*) = u(g'), \alpha^2\theta + (1-\alpha^2)u(h^*) = u(h)$ and $f \succsim g$, $g' \succsim h$. Assume without loss of generality that $\alpha^1 \geq \alpha^2$. By Axiom 5^{B}, ${}^{\alpha^2}\!/_{\alpha^1} f + (1 - {}^{\alpha^2}\!/_{\alpha^1})h^* \succsim {}^{\alpha^2}\!/_{\alpha^1} g + (1 - {}^{\alpha^2}\!/_{\alpha^1})h^*$. Additionally by the linearity of u, $u\big({}^{\alpha^2}\!/_{\alpha^1} g + (1 - {}^{\alpha^2}\!/_{\alpha^1})h^*\big) = u(g')$. Since u_s represents $\succsim_s$ and by Lemma A.13, ${}^{\alpha^2}\!/_{\alpha^1} g + (1 - {}^{\alpha^2}\!/_{\alpha^1})h^* \sim g'$. By Axioms 1^{B} and 5^{B},

$$ {}^{\alpha^2}\!/_{\alpha^1} f + (1 - {}^{\alpha^2}\!/_{\alpha^1})h^* \succsim {}^{\alpha^2}\!/_{\alpha^1} g + (1 - {}^{\alpha^2}\!/_{\alpha^1})h^* \succsim g' \succsim h$$

and $\alpha^2\varphi + (1-\alpha^2)u(h^*) = u\big({}^{\alpha^2}\!/_{\alpha^1} f + (1 - {}^{\alpha^2}\!/_{\alpha^1})h^*\big)$ and $\alpha^2\theta + (1-\alpha^2)u(h^*) = u(h)$. Hence $\varphi \mathrel{\widehat{\succsim}} \theta$.

4. Mixture Continuous: We show that upper contour set is closed. The lower contour set follows symmetrically. Let $\varphi, \psi, \theta \in \mathbb{R}^{S^{\mathcal{P}}}$. Let $\alpha^* \in (0,1), f, g, h \in H$ be such that $\alpha^*\varphi + (1-\alpha^*)u(h^*) = u(f), \alpha^*\psi + (1-\alpha^*)u(h^*) = u(g), \alpha^*\theta + (1-\alpha^*)u(h^*) = u(h)$. Now for all $\alpha \in [0,1]$,

$$\alpha^*\big(\alpha\varphi + (1-\alpha)\psi\big) + (1-\alpha^*)u(h^*) = u\big(\alpha f + (1-\alpha)g\big).$$

By Lemma A.84, for all $\alpha \in [0,1]$,

$$\alpha\varphi + (1-\alpha)\psi \mathrel{\widehat{\succsim}} \theta \iff \alpha f + (1-\alpha)g \succsim h.$$

Thus

$$\{\alpha \in [0,1] | \alpha\varphi + (1-\alpha)\psi \mathrel{\widehat{\succsim}} \theta\} = \{\alpha \in [0,1] | \alpha f + (1-\alpha)g \succsim h\}.$$

By Axiom 2^{B}, $\{\alpha \in [0,1] | h \succsim \alpha f + (1-\alpha)g\}$ is closed and hence $\{\alpha \in [0,1] | \alpha\varphi + (1-\alpha)\psi \mathrel{\widehat{\succsim}} \theta\}$ is closed.

5. Independence: Let $\varphi, \psi, \theta \in \mathbb{R}^{S^{\mathcal{P}}}, \alpha \in (0,1)$. Let $\alpha^* \in (0,1), f, g, h \in H$ be such that $\alpha^*\varphi + (1-\alpha^*)u(h^*) = u(f), \alpha^*\psi + (1-\alpha^*)u(h^*) = u(g), \alpha^*\theta + (1-\alpha^*)u(h^*) = u(h)$. By Lemma A.84 and Axiom 5^{B},

$\varphi \mathrel{\widehat{\succsim}} \psi \iff f \succsim g \iff \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h \iff \alpha\varphi + (1-\alpha)\theta \mathrel{\widehat{\succsim}} \alpha\psi + (1-\alpha)\theta.$

6. Monotonic: Let $\varphi, \psi \in \mathbb{R}^{S^{\mathcal{P}}}$ be such that for all $s \in S^{\mathcal{P}}$, $\varphi_s \geq \psi_s$. Let $\alpha^* \in (0,1), f, g \in H$ be such that $\alpha^*\varphi + (1-\alpha^*)u(h^*) = u(f), \alpha^*\psi + (1-\alpha^*)u(h^*) = u(g)$. Now, for all $s \in S^{\mathcal{P}}$,

$$u_s(f) = \alpha^*\varphi_s + (1-\alpha^*)u_s(h^*) \geq \alpha^*\psi_s + (1-\alpha^*)u_s(h^*) = u_s(g).$$

Since for all $s \in S$, u_s represents $\succsim_s$, we have for all $s \in S$,

$$f_s \succsim_s g_s.$$

Thus by Lemma A.13, $f \succsim g$ and hence $\varphi \mathrel{\widehat{\succsim}} \psi$.

□

Proposition A.86 $\succsim$ satisfies Axioms 1^{B}-5^{B} iff there exists (u, C) that is a state dependent Bewley representation for $\succsim$.

Proof. By Ghirardato et al.'s (2004) Theorem A.2, Gilboa et al.'s (2010) Lemma 3, and Lemma A.85, there exist closed and convex set $C \subseteq \Delta(S^{\mathcal{P}})$ and affine $v : \mathbb{R} \to \mathbb{R}$ such that for all $\varphi, \psi \in \mathbb{R}^{S^{\mathcal{P}}}$,

$$\varphi \mathrel{\widehat{\succsim}} \psi \iff \forall\, p \in C, v(\varphi) \cdot p \geq v(\psi) \cdot p.$$

Since v is affine and increasing on $\mathbb{R}$, there exist $a > 0, b \in \mathbb{R}$ such that $v = a\,\mathrm{Id} + b$, where Id is the identity function. Thus especially by a positive affine transformation of utility, we have for all $\varphi, \psi \in \mathbb{R}^{S^{\mathcal{P}}}$,

$$\varphi \mathrel{\widehat{\succsim}} \psi \iff \forall\, p \in C, \varphi \cdot p \geq \psi \cdot p.$$

Finally, we can embed C to $\Delta(S)$ with $\tilde{C}$ by defining for all $p \in C, s \notin S^{\mathcal{P}}$, $p_s = 0$. Thus $(u, \tilde{C})$ gives a Bewley representation for $\succsim$ since for all $f, g \in H$,

$$f \succsim g \overset{\text{Lem. A.13}}{\iff} (f_{S^{\mathcal{P}}}, h^*_{-S^{\mathcal{P}}}) \succsim (g_{S^{\mathcal{P}}}, h^*_{-S^{\mathcal{P}}}) \overset{\text{Lem. A.84}}{\iff} \mathrm{pr}_{S^{\mathcal{P}}}\, u(f) \mathrel{\widehat{\succsim}} \mathrm{pr}_{S^{\mathcal{P}}}\, u(g)$$

$$\iff \forall\, p \in C, \mathrm{pr}_{S^{\mathcal{P}}}\, u(f) \cdot p \geq \mathrm{pr}_{S^{\mathcal{P}}}\, u(g) \cdot p \iff \forall\, p \in \tilde{C}, u(f) \cdot p \geq u(g) \cdot p.$$

Additionally, by non-triviality, $\tilde{C} \neq \varnothing$.

Next, we show the only if direction.

1. Nontrivial: Assume, per contra, for all $f, g \in H$, $f \sim h$. Now by the definition of null-states, $S^{\mathcal{P}} = \varnothing$ and hence for all $s \in S, p \in C$, $p_s = 0$ which is a contradiction since $C \neq \varnothing$.
2. Reflexive: Let $f \in H$. Now for all $p \in C$, $p \cdot u(f) = p \cdot u(f)$ and hence $f \succsim f$.
3. Transitive: Let $f, g, h \in H$, $f \succsim g$, and $g \succsim h$. Now by the definition of $\succsim$, for all $p \in C$,

   $$p \cdot u(f) \geq p \cdot u(g) \text{ and } p \cdot u(g) \geq p \cdot u(h)$$

   and hence by the transitivity of $\geq$,

   $$p \cdot u(f) \geq p \cdot u(h).$$

   Thus $f \succsim h$.
4. Axiom 2^{B}: Let $f, g, h \in H$. Define for all $p \in C$, an operator $I_p : [0,1] \to \mathbb{R}$,

   $$\alpha \overset{I_p}{\longmapsto} \alpha\Big(p \cdot \big(u(f) - u(g)\big)\Big) + p \cdot u(g) - p \cdot u(h).$$

   As a linear function for all $p \in C$, I_p is continuous. Now

   $$\{\alpha \in [0,1] | \alpha f + (1-\alpha) g \succsim h\} = \big\{\alpha \in [0,1] \big| \forall\, p \in C, p \cdot u\big(\alpha f + (1-\alpha) g\big) \geq p \cdot u(h)\big\}$$

   $$= \big\{\alpha \in [0,1] \big| \forall\, p \in C, \alpha p \cdot \big(u(f) - u(g)\big) + p \cdot u(g) - p \cdot u(h) \geq 0\big\}$$

   $$= \{\alpha \in [0,1] | \forall\, p \in C, I_p(\alpha) \geq 0\} = \bigcap_{p \in C} I_p^{-1}[0, \infty)$$

   and symmetrically

   $$\{\alpha \in [0,1] | h \succsim \alpha f + (1-\alpha) g\} = \bigcap_{p \in C} I_p^{-1}(-\infty, 0]$$

   are closed as the intersections of preimages of a continuous function over closed sets.

5. Axiom 3^{B}: Let $s^* \in S, f \in H, x_{s^*}, y_{s^*} \in \Delta(X_{s^*})$. Assume w.l.o.g. $u_{s^*}(x_{s^*}) \geq u_{s^*}(y_{s^*})$. Now for all $p \in C$,

$$p \cdot (x_{s^*}, f_{-s^*}) = \sum_{s \in S \setminus \{s^*\}} p_s u_s(f_s) + p_{s^*} u_{s^*}(x_{s^*}) \geq \sum_{s \in S \setminus \{s^*\}} p_s u_s(f_s) + p_{s^*} u_{s^*}(y_{s^*}) = p \cdot (y_{s^*}, f_{-s^*}).$$

Thus $(x_{s^*}, f_{-s^*}) \succsim (y_{s^*}, f_{-s^*})$.

6. Axiom 4^{B}: Let $s^* \in S, f, g \in H, x_{s^*}, y_{s^*} \in \Delta(X_{s^*})$ and $(x_s^*, f_{-s^*}) \succsim (y_s^*, f_{-s^*})$. If $s^* \notin S^{\mathcal{P}}$, then by definition, $(x_s^*, g_{-s^*}) \succsim (y_s^*, g_{-s^*})$. Thus assume $s^* \in S^{\mathcal{P}}$. Thus since $C \neq \varnothing$, there exists $p^* \in C$ such that $p^*_{s^*}$. Since $(x_s^*, f_{-s^*}) \succsim (y_s^*, f_{-s^*})$, especially

$$p^* \cdot (x_s^*, f_{-s^*}) \geq p^* \cdot (y_s^*, f_{-s^*}).$$

Hence

$$\sum_{s \in S \setminus \{s^*\}} p_s^* u_s(f_s) + p_{s^*}^* u_{s^*}(x_{s^*}) \geq \sum_{s \in S \setminus \{s^*\}} p_s^* u_s(f_s) + p_{s^*}^* u_{s^*}(y_{s^*})$$
$$\iff p_{s^*}^* u_{s^*}(x_{s^*}) \geq p_{s^*}^* u_{s^*}(y_{s^*}) \overset{p^*_{s^*}>0}{\iff} u_{s^*}(x_{s^*}) \geq u_{s^*}(y_{s^*}).$$

Thus by the above proof for Axiom 3^{B},

$$(x_{s^*}, g_{-s^*}) \succsim (y_{s^*}, g_{-s^*}).$$

7. Axiom 5^{B}: Let $f, g, h \in H, \alpha \in (0,1)$. By the linearity of u, we have

$$\begin{aligned}
f \succsim g \quad &\iff \quad \forall\, p \in C, \sum_{s \in S} p_s u_s(f_s) \geq \sum_{s \in S} p_s u_s(g_s) \\
&\overset{\alpha>0}{\iff} \quad \forall\, p \in C, \sum_{s \in S} p_s \alpha u_s(f_s) + \sum_{s \in S} p_s (1-\alpha) u_s(h_s) \geq \sum_{s \in S} p_s \alpha u_s(g_s) + \sum_{s \in S} p_s (1-\alpha) u_s(h_s) \\
&\overset{u \text{ linear}}{\iff} \quad \forall\, p \in C, \sum_{s \in S} p_s u_s(\alpha f_s + (1-\alpha) h_s) \geq \sum_{s \in S} p_s \alpha u_s(\alpha g_s + (1-\alpha) h_s) \\
&\iff \quad \alpha f_s + (1-\alpha) h_s \succsim \alpha g_s + (1-\alpha) h_s
\end{aligned}$$

□

A.6.2 Bewley Uniqueness Proof

Next we show the uniqueness of the representation. First some notation. Let $f^*, g^* \in H$ be such that for all $s \in S^{\mathcal{P}}$, $f_s^* \succ_s g_s^*$ which exist by Axiom 1^{B} and denote $h^* \coloneqq {}^1\!/_2 f^* + {}^1\!/_2 g^*$. The next simple lemma shows that states are not proper iff they have zero probability.

Lemma A.87 Let $\succsim$ be an order on H, (u, C) be a state dependent Bewley representation for $\succsim$. $s^* \notin S^{\mathcal{P}}$ if and only if for all $p \in C$, $p_{s^*} = 0$.

Proof. First, we show the if direction. Let $s^* \in S$ be such that for all $p \in C$, $p_{s^*} = 0$. Let $x_{s^*}, y_{s^*} \in \Delta(X_{s^*}), f \in H$. Now for all $p \in C$,

$$p \cdot u\big((x_{s^*}, f_{-s^*})\big) = \sum_{s \in S \setminus \{s^*\}} p_s u_s(f_s) + p_{s^*} u_{s^*}(x_{s^*})$$
$$\overset{p_{s^*}=0}{=} \sum_{s \in S \setminus \{s^*\}} p_s u_s(f_s) + p_{s^*} u_{s^*}(y_{s^*}) = p \cdot u\big((y_{s^*}, f_{-s^*})\big).$$

and thus by the representation,

$$(x_{s^*}, f_{-s^*}) \sim (y_{s^*}, f_{-s^*}).$$

Hence by definition, s^* is null.

The only if direction follows directly from the definition of a Bewley representation. □

The next three lemmas establish that u_s represents $\succsim_s$. First, we show a weak representation for the strict part.

Lemma A.88 Let (u, C) be a state dependent Bewley representation for $\succsim$, $s^* \in S^{\mathcal{P}}, x_{s^*}, y_{s^*} \in \Delta(X_{s^*})$. If $x_{s^*} \succ_{s^*} y_{s^*}$, then $u_{s^*}(x_{s^*}) > u_{s^*}(y_{s^*})$.

Proof. Assume, per contra, $u_{s^*}(y_{s^*}) \geq u_{s^*}(x_{s^*})$. Now for all $p \in C$, $f \in H$

$$\sum_{s \in S \setminus \{s^*\}} p_s u_s(f_s) + p_{s^*} u_{s^*}(y_{s^*}) \geq \sum_{s \in S \setminus \{s^*\}} p_s u_s(f_s) + p_{s^*} u_{s^*}(x_{s^*}).$$

Thus by the representation, for all $f \in H$,

$$(y_{s^*}, f_{-s^*}) \succsim (x_{s^*}, f_{-s^*}).$$

On the other hand, by the definition of $\succsim_{s^*}$, since $x_{s^*} \succ_{s^*} y_{s^*}$ i.e. $x_{s^*} \succsim_{s^*} y_{s^*}$ and $y_{s^*} \not\succsim_{s^*} x_{s^*}$, for all $f \in H$, $(x_{s^*}, f_{-s^*}) \succsim (y_{s^*}, f_{-s^*})$ and there exists $f^0 \in H$ such that $(y_{s^*}, f^0_{-s^*}) \not\succsim (x_{s^*}, f^0_{-s^*})$ which is a contradiction with the above. □

Second, we show that $u(h^*)$ is in the interior of utility space.

Corollary A.89 Let (u, C) be a state dependent Bewley representation for $\succsim$. Then $\operatorname{pr}_{S^{\mathcal{P}}} u(h^*) \in \operatorname{int} \operatorname{pr}_{S^{\mathcal{P}}} u(H)$.

Proof. By Proposition A.86, $\succsim$ satisfies Axioms 1^{B}-5^{B}. Let $s^* \in S^{\mathcal{P}}$. By Lemma A.83, $f^*_{s^*} \succ_{s^*} h^*_{s^*} \succ_{s^*} g^*_{s^*}$. Hence by Lemma A.88, $u_{s^*}(f^*_{s^*}) > u_{s^*}(h^*_{s^*}) > u_{s^*}(g^*_{s^*})$. Thus by the linearity of u_{s^*}, $u_{s^*}(h^*_{s^*}) \in \operatorname{int} u_{s^*}\big(\Delta(X_{s^*})\big)$. Since $s^* \in S^{\mathcal{P}}$ was arbitrary, $\operatorname{pr}_{S^{\mathcal{P}}} u(h^*) \in \operatorname{int} \operatorname{pr}_{S^{\mathcal{P}}} u(H)$. □

Third, we show that u_s represents indifferences.

Lemma A.90 Let (u, C) be a state dependent Bewley representation for $\succsim$, $s^* \in S^{\mathcal{P}}, x_{s^*}, y_{s^*} \in \Delta(X_{s^*})$. If $x_{s^*} \sim_{s^*} y_{s^*}$, then $u_{s^*}(x_{s^*}) = u_{s^*}(y_{s^*})$.

Proof. Assume, per contra without loss of generality, $u_{s^*}(y_{s^*}) > u_{s^*}(x_{s^*})$. Since s^* is not null, by Lemma A.87, there exists $p^0 \in C$ such that $p^0_{s^*} > 0$. Let $f^0 \in H$. Now

$$\begin{aligned} p^0 \cdot u\big((x_{s^*}, f^0_{-s^*})\big) &= \sum_{s \in S \setminus \{s^*\}} p^0_s u_s(f^0_s) + p^0_{s^*} u_{s^*}(x_{s^*}) \\ &\overset{p^0_{s^*}>0}{<} \sum_{s \in S \setminus \{s^*\}} p^0_s u_s(f^0_s) + p^0_{s^*} u_{s^*}(y_{s^*}) = p^0 \cdot u\big((y_{s^*}, f^0_{-s^*})\big). \end{aligned}$$

Thus by the representation,

$$(x_{s^*}, f^0_{-s^*}) \not\succsim (y_{s^*}, f^0_{-s^*})$$

which is a contradiction since $x_{s^*} \sim_{s^*} y_{s^*}$ and hence especially for all $f \in H$,

$$(x_{s^*}, f_{-s^*}) \succsim (y_{s^*}, f_{-s^*}).$$

Thus $u_{s^*}(x_{s^*}) = u_{s^*}(y_{s^*})$. □

The last result shows the uniqueness of the state dependent Bewley representation.

Proposition A.91 If (u, C) and $(\widetilde{u}, \widetilde{C})$ are state dependent Bewley representations for $\succsim$, then there exist $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that

$$\widetilde{C} = \left\{ \widetilde{p} \in \Delta(S) \,\middle|\, \exists\, p \in C, \forall\, s \in S, \widetilde{p}_s = \frac{a_s^{-1} p_s}{\sum_{s \in S} a_s^{-1} p_s} \right\}$$

and for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = a_s u_s + b_s.$$

Proof. Denote

$$u_{S^\neg} \coloneqq \operatorname{pr}_{S^{\mathcal{P}}} u \qquad \text{and} \qquad \widetilde{u}_{S^\neg} \coloneqq \operatorname{pr}_{S^{\mathcal{P}}} \widetilde{u}$$

and for all $D \subseteq \Delta(S)$,

$$D_{S^\neg} \coloneqq \operatorname{pr}_{S^{\mathcal{P}}} D.$$

By Lemmas A.88 and A.90, for all $s \in S^{\mathcal{P}}$, u_s and $\widetilde{u}_s$ represents $\succsim_s$. By Lemma A.83, $\succsim_s$ satisfies, completeness, transitivity, mixture continuity, and independence and u_s and $\widetilde{u}_s$ are affine representations, hence by von Neumann–Morgenstern utility theorem (Fishburn, 1970), there exist $a \in \mathbb{R}^{S^{\mathcal{P}}}_{++}$, $b \in \mathbb{R}^{S^{\mathcal{P}}}$ such that for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = a_s u_s + b_s.$$

Denote for all $s \in S \setminus S^{\mathcal{P}}$, $a_s = 1$. Define

$$\widehat{C} := \left\{ \left(\frac{a_s^{-1} p_s}{\sum_{s' \in S} a_{s'}^{-1} p_{s'}} \right)_{s \in S} \middle| p \in C \right\}.$$

Now $\widehat{C} \subseteq \Delta(S)$ since for all $s \in S$, $a_s > 0$. Additionally, for all $\widehat{p} \in \widehat{C}, \widetilde{p} \in \widetilde{C}, s \notin S^{\mathcal{P}}$,

$$\widehat{p}_s = 0 = \widetilde{p}_s. \tag{A.94}$$

Thus $\widehat{C}_{S^\neg}, \widetilde{C}_{S^\neg} \subseteq \Delta(S^{\mathcal{P}})$.

Define $\succsim^{\widehat{C}}$ and $\succsim^{\widetilde{C}}$ on $\mathbb{R}^{S^{\mathcal{P}}}$ by for all $\varphi, \psi \in \mathbb{R}^{S^{\mathcal{P}}}$,

$$\varphi \succsim^{\widehat{C}} \psi \iff \forall\, \widehat{p} \in \widehat{C}_{S^\neg}, p \cdot \varphi \geq p \cdot \psi \text{ and}$$

$$\varphi \succsim^{\widetilde{C}} \psi \iff \forall\, \widetilde{p} \in \widetilde{C}_{S^\neg}, p \cdot \varphi \geq p \cdot \psi.$$

Let $\varphi, \psi \in \mathbb{R}^{S^{\mathcal{P}}}$. We show that $\varphi \succsim^{\widehat{C}} \psi \iff \varphi \succsim^{\widetilde{C}} \psi$. Denote $a^{-1}\varphi = (a_s^{-1}\varphi_s)_{s \in S^{\mathcal{P}}}$ and similarly for ψ. Let $\alpha^\star \in (0, 1]$ be such that $\alpha^\star a^{-1}\varphi + (1 - \alpha^\star) u_{S^\neg}(h^*), \alpha^\star a^{-1}\psi + (1 - \alpha^\star) u_{S^\neg}(h^*) \in u_{S^\neg}(H)$, which exists by Corollary A.89 and let $f, g \in H$ be such that

$$\alpha^\star A^{-1}\varphi + (1 - \alpha^\star) u_{S^\neg}(h^*) = u_{S^\neg}(f) \text{ and } \alpha^\star A^{-1}\psi + (1 - \alpha^\star) u_{S^\neg}(h^*) = u_{S^\neg}(g).$$

Now we have when multiplications with a are define statewise

$$\begin{aligned}
&\varphi \succsim^{\widehat{C}} \psi \iff \forall\, p \in \widehat{C}_{S^\neg}, p \cdot \varphi \geq p \cdot \psi \\
\iff &\quad \forall\, \widehat{p} \in \widehat{C}_{S^\neg}, \alpha^\star \widehat{p} \cdot \varphi + (1 - \alpha^\star) \widehat{p} \cdot a u_{S^\neg}(h^*) \geq \alpha^\star \widehat{p} \cdot \psi + (1 - \alpha^\star) \widehat{p} \cdot a u_{S^\neg}(h^*) \\
\iff &\quad \forall\, \widehat{p} \in \widehat{C}_{S^\neg}, \widehat{p} \cdot a\Big(\alpha^\star a^{-1}\varphi + (1 - \alpha^\star) u_{S^\neg}(h^*)\Big) \geq \widehat{p} \cdot a\Big(\alpha^\star a^{-1}\psi + (1 - \alpha^\star) u_{S^\neg}(h^*)\Big) \\
\iff &\quad \forall\, p \in C_{S^\neg}, \frac{a^{-1} p}{\sum_{s \in S} p_s a_s^{-1}} \cdot a\Big(\alpha^\star a^{-1}\varphi + (1 - \alpha^\star) u_{S^\neg}(h^*)\Big) \\
&\qquad\qquad \geq \frac{a^{-1} p}{\sum_{s \in S} p_s a_s^{-1}} \cdot a\Big(\alpha^\star a^{-1}\psi + (1 - \alpha^\star) u_{S^\neg}(h^*)\Big) \\
\iff &\quad \forall\, p \in C_{S^\neg}, p \cdot \Big(\alpha^\star a^{-1}\varphi + (1 - \alpha^\star) u_{S^\neg}(h^*)\Big) \geq p \cdot \Big(\alpha^\star a^{-1}\psi + (1 - \alpha^\star) u_{S^\neg}(h^*)\Big) \\
\iff &\ \forall\, p \in C_{S^\neg}, p \cdot u_{S^\neg}(f) \geq p \cdot u_{S^\neg}(g) \iff \forall\, p \in C, p \cdot u(f) \geq p \cdot u(g) \iff f \succsim g \\
\iff &\ \forall\, \widetilde{p} \in \widetilde{C}, \widetilde{p} \cdot \widetilde{u}(f) \geq \widetilde{p} \cdot \widetilde{u}(g) \iff \varphi \succsim^{\widetilde{C}} \psi
\end{aligned}$$

where the last equivalence follows symmetrically to the first part.

Thus $\succsim^{\widetilde{C}}$ has state independent Bewley representations with $(\mathrm{Id}, \widetilde{C}_{S^\neg})$ and $(\mathrm{Id}, \widehat{C}_{S^\neg})$. By Ghirardato et al.'s (2004) Theorem A.2,

$$\widetilde{C}_{S^\neg} = \widehat{C}_{S^\neg}.$$

Finally, by Lemma A.87 and (A.94),

$$\widetilde{C} = \widehat{C}.$$

This proves the uniqueness. □

Putting all together, we have the following theorem.

Theorem A.81 (Bewley Representation) $\succsim$ satisfies Axioms 1^{B}-5^{B} if and only if there exists (u, C) that is a state dependent Bewley representation for $\succsim$.

Additionally, let (u, C) be a state dependent Bewley representation for $\succsim$, $\widetilde{u} = (\widetilde{u}_s)_{s\in S}$ be such that for all $s \in S$, $\widetilde{u}_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\widetilde{C} \subseteq \Delta(S)$ be closed and convex. Then $(\widetilde{u}, \widetilde{C})$ is a state dependent Bewley representation for $\succsim$ if and only if there exist $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that

$$\widetilde{C} = \left\{ \left(\frac{a_s^{-1} p_s}{\sum_{s'\in S} a_{s'}^{-1} p_{s'}} \right)_{s\in S} \middle| p \in C \right\}.$$

and for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = a_s u_s + b_s.$$

Proof. Follows from Propositions A.86 and A.91. □

A.6.3 State Independent Bewley Proof

Assume state independent setting with for all $s \in S$, $X_s = X$. The next lemma shows that under Axiom 6^{B}, all state dependent utilities are positive affine transformations of each other.

Lemma A.92 Assume that $\succsim$ satisfies Axioms 1^{B}-6^{B} and has a state dependent Bewley expected utility representation with (u, C).

Then there exist an affine $u^* : \Delta(X) \to \mathbb{R}$ and $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that for all $s \in S^{\mathcal{P}}$,

$$u_s = a_s u^* + b_s.$$

Proof. By Lemmas A.88 and A.90, for all $s \in S^{\mathcal{P}}$, u_s represents $\succsim_s$. By the affinity of u, for all $s \in S, x, y, c \in \Delta(X), \alpha \in (0, 1)$,

$$x \succ_s y \Rightarrow \alpha x + (1-\alpha)c \succ_s \alpha y + (1-\alpha)c. \tag{A.95}$$

Additionally, by Axiom 6^{B}, for all $s, s' \in S, x, y \in \Delta(X)$,

$$x \succ_s y \Rightarrow x \succsim_{s'} y. \tag{A.96}$$

Let $s \in S^{\mathcal{P}}$. Now there exist $x^*, y^* \in \Delta(X)$ such that $u_s(x^*) > u_s(y^*)$. Since u_s represents $\succsim_s$, we have

$$x^* \succ_s \frac{1}{2}x^* + \frac{1}{2}y^* \succ_s y^*. \tag{A.97}$$

Let $s' \in S^{\mathcal{P}}$. By (A.95,A.96,A.97) and Lemma A.36, $u_{s'}$ gives a weak affine representation for $\succsim_s$. By the affine representation, $\succsim_s$ satisfies Axioms 1.1, 1.2, 4′. Additionally, u_s gives a weak affine representation for $\succsim_s$. By Proposition A.12, there exist $a_{s'} > 0, b_{s'} \in \mathbb{R}$ such that

$$u_{s'} = a_{s'}u_s + b_{s'}.$$

Since $s' \in S^{\mathcal{P}}$ was arbitrary, this shows the claim. □

This gives us state independent Bewley expected utility.

Theorem A.82 (State Independent Bewley Representation) $\succsim$ satisfies Axioms 1^{B}-6^{B} if and only if there exists (u, C) that is a state independent Bewley representation for $\succsim$.

Additionally, let (u, C) be a state independent Bewley representation for $\succsim$, $\widetilde{u} = (\widetilde{u}_s)_{s\in S}$ be such that for all $s \in S$, $\widetilde{u}_s : \Delta(X_s) \to \mathbb{R}$ is affine and $\widetilde{C} \subseteq \Delta(S)$ be closed and convex. Then $(\widetilde{u}, \widetilde{C})$ is a state dependent Bewley representation for $\succsim$ if and only if there exist $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that

$$\widetilde{C} = \left\{ \left(\frac{a_s^{-1} p_s}{\sum_{s'\in S} a_{s'}^{-1} p_{s'}} \right)_{s\in S} \middle| p \in C \right\}.$$

and for all $s \in S^{\mathcal{P}}$,

$$\widetilde{u}_s = a_s u + b_s.$$

Proof. We first show the sufficiency of the axioms. By Theorem A.81, there exists (u, C) that is a state dependent Bewley expected utility for $\succsim$. By Lemma A.92, there exist $u^* : \Delta(X) \to \mathbb{R}$ and $a \in \mathbb{R}^S_{++}, b \in \mathbb{R}^S$ such that for all $s \in S^{\mathcal{P}}$,

$$u_s = a_s u^* + b_s.$$

By the uniqueness in Theorem A.81, there exists $C^* \subseteq \Delta(S)$ such that (u^*, C^*) is a state independent Bewley expected utility for $\succsim$ since changing the utility to u^* for $s \in S \setminus S^{\mathcal{P}}$ does not affect the representation.

Next we show the necessity of the axioms. By Theorem A.81, we only show Axiom 6^{B}. This follows from Lemmas A.88 and A.90 and the state independent utility.

Finally, the uniqueness follows from Theorem A.81 and the existence of state independent Bewley representation.

□

Appendix B

Appendix for Chapter 2

This appendix is organized as follows. In Appendix B.1, we reconsider expected utility and show Theorems 2.2 and 2.3. Here, we show these results extend to any interval X. Additionally, we show that Theorem 2.3 extends to Borel lotteries on any interval and Theorem 2.2 extends to Borel lotteries on a compact interval.

In Appendix B.2, we consider general preferences that satisfy Restricted Common Consequence Independence, continuity in prizes, and mixture continuity. We show that there exist a common utility and common degree for power probability weighting such that lotteries sharing the same co-probability are compared using power weighted expected utility. This provides common results for all the representations in the paper.

In Appendix B.3, we show Theorem 2.5 by building on Appendix B.2. Here we add Branch Cancellation to the previous axioms.

In Appendix B.4, we start the proof for Theorem 2.1. Here, we consider Common Consequence Independence and establish that Axioms 2.1-2.5 characterize Entropy Modified Expected Utility and Axioms 2.1-2.4 characterize power weighted expected utility with generalized entropy cost function of H^λ as defined in Section 2.5. These results build on Appendices B.2 and B.3.

In Appendix B.5, we show the uniqueness for the power weighted variational expected utility. This shows Theorems 2.6 and 2.7 and Corollary 2.8 and finishes the proof for Theorem 2.1.

In Appendix B.6, we show Theorem 2.4. Here, we consider preferences that satisfy Restricted Common Consequence Independence and CE Bounds. This section builds on Appendix B.2.

In Appendix B.7 we consider alternative continuity axiom using weighted weak convergence as defined in Remark 2.1. Here we show Remarks 2.1 and 2.3.

First, we start off with a definition for non-diagonal different prizes given a co-probability and an observation of how continuity conditional on a co-probability is equivalent to continuity in prizes. For a co-probability $p \in \Delta(\mathbb{N})$, denote

$$X^{\operatorname{supp} p} = \{(x_i)_{i \in \operatorname{supp} p} | \text{for all } i \in \operatorname{supp} p, x_i \in X\}.$$

Second, for a co-probability $p \in \Delta(\mathbb{N})$, denote the off-diagonal of $X^{\operatorname{supp} p}$ by

$$\operatorname{NDiag}(X^p) = \{x \in X^{\operatorname{supp} p} | \, \forall \, i, j \in \operatorname{supp} p, i \neq j, x_i \neq x_j\}.$$

We endow $\operatorname{NDiag}(X^p)$ with the Euclidean product topology.

Our first result shows how weak convergence continuity conditional on a co-probability, Axiom 2.2, is equivalent to continuity in prizes off-diagonal conditional on a co-probability. We will use this form of continuity in the proof.

Proposition B.1 The following two are equivalent:

1. $\succsim$ satisfies Axiom 2.2
2. For all lotteries P, co-probabilities $p \in \Delta(\mathbb{N})$, and convergent sequences $(x^i)_{i=1}^{\infty} \subseteq \operatorname{NDiag}(X^p)$ with a limit $x \in \operatorname{NDiag}(X^p)$,

$$\text{if } \sum_{l \in \operatorname{supp} p} p_l x_l^i \succsim P \text{ for all } i \in \mathbb{N}, \text{ then } \sum_{l \in \operatorname{supp} p} p_l x_l \succsim P$$

and

$$\text{if } P \succsim \sum_{l \in \operatorname{supp} p} p_l x_l^i \text{ for all } i \in \mathbb{N}, \text{ then } P \succsim \sum_{l \in \operatorname{supp} p} p_l x_l.$$

Proof. This follows from the following observation since for any lottery P with a co-probability p, there exists $x \in \operatorname{NDiag}(X^p)$ such that $P = \sum_{l \in \operatorname{supp} p} p_l x_l$. Let $p \in \Delta(\mathbb{N})$ and $(x^i)_{i=1}^{\infty} \subseteq \operatorname{NDiag}(X^p)$, $x \in \operatorname{NDiag}(X^p)$. Then $\sum_{l \in \operatorname{supp} p} p_l x_l^i \to \sum_{l \in \operatorname{supp} p} p_l x_l$ weakly iff $x^i \to x$.

First, assume that $\sum_{l \in \operatorname{supp} p} p_l x_l^i \to \sum_{l \in \operatorname{supp} p} p_l x_l$ weakly. Let $l \in \operatorname{supp} p$. Denote

$$\varepsilon = \min_{n,m \in \operatorname{supp} p, n \neq m} |x_n - x_m| > 0$$

where positivity follows from $x \in \operatorname{NDiag}(X^p)$. Let $f^l : X \to \mathbb{R}$ be a continuous function such that $f^l(x) = 1$ if $|x - x_l| < \frac{1}{3}\varepsilon$ and for each $k \in \operatorname{supp} p \setminus \{l\}$, $f^l(x) = 0$ if $|x - x_k| < \frac{1}{3}\varepsilon$. f^l exists by the choice of ε. By weak convergence,

$$p_l x_l^i = \sum_{k \in \operatorname{supp} p} f^l(x_k^i) p_k \to \sum_{k \in \operatorname{supp} p} f^l(x_k) p_k = p_l x_l.$$

Thus $x^i \to x$.

Second, assume that $x^i \to x$. Let f be a continuous function and $\varepsilon > 0$. Since $\operatorname{supp} p$ is finite, there exists $\delta > 0$ such that for each $l \in \operatorname{supp} p$, if $|x - x_l| < \delta$, then $|f(x) - f(x_l)| < \varepsilon$. Now there exists $i^0 \in \mathbb{N}$ such that for all $i > i^0$, $\|x^i - x\| < \delta$ and so especially for all $l \in \operatorname{supp} p$, $|x^i_l - x_l| < \delta$. Thus for all $i > i^0$,

$$\Big| \sum_{k \in \operatorname{supp} p} f(x^i_k) p_k - \sum_{k \in \operatorname{supp} p} f(x_k) p_k \Big| \leq \sum_{k \in \operatorname{supp} p} p_k |f(x^i_k) - f(x_k)| < \sum_{k \in \operatorname{supp} p} p_k \varepsilon = \varepsilon.$$

Thus $\sum_{k \in \operatorname{supp} p} f(x^i_k) p_k \to \sum_{k \in \operatorname{supp} p} f(x_k) p_k$ that shows weak convergence. □

B.1 Expected Utility

For this section, we consider a more general setup that shows Theorems 2.2 and 2.3 as simple corollaries. We allow $X \subseteq \mathbb{R}$ to be any interval. Denote by $\Delta^{\mathrm{Bor}}(X)$ the set of all Borel probability measures on X. We consider preferences $\succsim$ on $\Delta^{\mathrm{Bor}}(X)$ and endow $\Delta^{\mathrm{Bor}}(X)$ with the topology of weak convergence. We define co-probabilities for simple lotteries as before. Our first axiom is as before that the preferences on the general lotteries are complete, transitive, and monotone on degenerate lotteries.

Axiom 1′EU (Weak Order and Monotonicity) $\succsim$ is complete and transitive and for all $x > y$, $\delta_x \succ \delta_y$.

Restricted Branch Cancellation is as before and only applies to simple lotteries.

Axiom 2′EU (Restricted Branch Cancellation) If two simple lotteries $P, Q \in \Delta(X)$ have the same co-probabilities and if $x, y \notin \operatorname{supp} P \cup \operatorname{supp} Q$ and $\alpha \in (0,1)$, then

$$\alpha P + (1-\alpha)\delta_x \succsim \alpha Q + (1-\alpha)\delta_x \Longrightarrow \alpha P + (1-\alpha)\delta_y \succsim \alpha Q + (1-\alpha)\delta_y.$$

For the possibility of compact intervals, we will use censored shifts that keep the shifted lottery within the prize interval by censoring.

Definition For a simple lottery $P \in \Delta(X)$ and $\varepsilon \in \mathbb{R}$, define $P +_{|X} \varepsilon$ by for all $x \in X$

$$P +_{|X} \varepsilon(x) = \begin{cases} P(x - \varepsilon), & \text{if } x \notin \{\inf X, \sup X\} \\ \displaystyle\sum_{y \in \operatorname{supp} P, y \geq x - \varepsilon} P(y), & \text{if } x = \sup X \\ \displaystyle\sum_{y \in \operatorname{supp} P, y \leq x - \varepsilon} P(y), & \text{if } x = \inf X. \end{cases}$$

Additionally, denote $P -_{|X} \varepsilon = P +_{|X} (-\varepsilon)$.

Here, we are shifting the probabilities of the prizes of P by ε. Hence, the probability of prize x for $P +_{|X} \varepsilon$ after shifting is the probability of the prize $x - \varepsilon$ for P. Additionally, the shifted lottery is censored above $\sup X$ and below $\inf X$ by accumulating all the shifted values going over at these points.

First, if X is an open interval, then for a small enough ε, we have $P + \varepsilon = P +_{|X} \varepsilon$. Second, for any interval X, we have for a small enough ε, $P +_{|X} \varepsilon \in \Delta(X)$. Third, for a small enough of ε, at most either the best prize or the worst prize of the lottery is affected by the censoring. The next axiom generalizes our shift continuity to a general interval and to Borel lotteries using censored shifts.

Axiom 3′EU (Censored Prize Shift Continuity) For all simple lotteries $P \in \Delta(X)$ and Borel lotteries $Q \in \Delta^{\text{Bor}}(X)$,
if $P \succ Q$, then there exists $\varepsilon > 0$ such that

$$P -_{|X} \varepsilon \in \Delta(X) \text{ and } P -_{|X} \varepsilon \succ Q$$

and if $Q \succ P$, then there exists $\varepsilon > 0$ such that

$$P +_{|X} \varepsilon \in \Delta(X) \text{ and } Q \succ P +_{|X} \varepsilon.$$

This continuity allows for rank dependent discontinuities. Without Borel lotteries, for an open interval of prizes, this is equivalent to our simple Prize Shift Continuity, Axiom 2.7.

Finally, we generalize First Order Stochastic Dominance to apply for all lotteries.

Axiom 4′EU (FOSD) For all lotteries $P, Q \in \Delta^{\text{Bor}}(X)$, if P first order stochastically dominates Q, then $P \succsim Q$.

And, generalize Weak Convergence Continuity to apply for all lotteries.

Axiom 5′EU (Weak Convergence Continuity) For each $Q \in \Delta^{\text{Bor}}(X)$, the sets $\{P \in \Delta^{\text{Bor}}(X) | P \succsim Q\}$ and $\{P \in \Delta^{\text{Bor}}(X) | Q \succsim P\}$ are closed.

In this general context our first result considers the case when X is bounded and closed interval. In this case, Weak Order, Restricted Branch Cancellation and either Weak Convergence Continuity or Continuity in Prizes and FOSD are equivalent to preferences having an expected utility representation for all lotteries. Since X is compact, the continuous utility u will be bounded in this case.

Theorem B.2 Assume that X is a closed and bounded interval. Then the following three conditions are equivalent:

1. $\succsim$ satisfies Axioms $1'^{\text{EU}}$ to $4'^{\text{EU}}$.
2. $\succsim$ satisfies Axioms $1'^{\text{EU}}$, $2'^{\text{EU}}$, and $5'^{\text{EU}}$.
3. There exists a strictly increasing and continuous $u : X \to \mathbb{R}$ such that for all lotteries $P, Q \in \Delta^{\text{Bor}}(X)$,
$$P \succsim Q \iff \int_X u\, dP \geq \int_X u\, dQ.$$

Second, we consider the case of general interval. First, in this case the results under Weak Convergence Continuity remain but give bounded utility.

Theorem B.3 Assume that X is an interval. Then the following two conditions are equivalent:

1. $\succsim$ satisfies Axioms $1'^{\text{EU}}$, $2'^{\text{EU}}$, and $5'^{\text{EU}}$.
2. There exists a strictly increasing, bounded, and continuous $u : X \to \mathbb{R}$ such that for all lotteries $P, Q \in \Delta^{\text{Bor}}(X)$,
$$P \succsim Q \iff \int_X u\, dP \geq \int_X u\, dQ.$$

Third, with FOSD and general interval, we only get a representation for simple lotteries due to unbounded utilities.

Theorem B.4 Assume that X is an interval. If $\succsim$ satisfies Axioms $1'^{\text{EU}}$ to $4'^{\text{EU}}$, then there exists a strictly increasing and continuous $u : X \to \mathbb{R}$ such that for all simple lotteries $P, Q \in \Delta(X)$,
$$P \succsim Q \iff \int_X u\, dP \geq \int_X u\, dQ.$$

Our proof is sketched in Section 2.6.2. The proof for this result follows in three steps. First, we show that preferences conditional on a co-probability can be extended continuously to on-diagonal. Then we consider dyadic lotteries that have uniform 2^{-n} co-probability for some n. Second, we show that dyadic lotteries with the same co-probability satisfy first order stochastic dominance under Weak Convergence Continuity. Third, using this we get an expected utility representation for dyadic lotteries. Finally, we can approximate any general lotteries using these dyadic lotteries either with first order stochastic dominance or weak convergence continuity.

B.1.1 Continuity On-Diagonal

This section shows that we can extend preferences conditional on a co-probability continuously to on-diagonal.

First, we define a stronger continuity than Axiom 2.2 that we will use in the proof. This is continuity when the limit prizes might be on-diagonal but the converging sequence is off-diagonal.

Axiom 6′EU (On-Diagonal Continuity) For all lotteries $P \in \Delta^{\mathrm{Bor}}(X)$ and co-probabilities $p \in \Delta(\mathbb{N})$, if $(x^i)_{i=1}^{\infty} \subseteq \mathrm{NDiag}(X^p)$ is a convergent sequence and if for all $i \in \mathbb{N}$,

$$\sum_{l \in \operatorname{supp} p} p_l x_l^i \succsim P, \text{ then } \sum_{l \in \operatorname{supp} p} p_l \lim_{i \to \infty} x_l^i \succsim P$$

and if for all $i \in \mathbb{N}$,

$$P \succsim \sum_{l \in \operatorname{supp} p} p_l x_l^i, \text{ then } P \succsim \sum_{l \in \operatorname{supp} p} p_l \lim_{i \to \infty} x_l^i.$$

The next two results show that weak convergence continuity or our weak continuity in money, Axiom 3′EU, and first order stochastic dominance give the above on-diagonal continuity. First, we consider weak convergence continuity.

Lemma B.5 If $\succsim$ satisfies Axioms 1′EU and 5′EU, then it satisfies Axiom 6′EU.

Proof. Let $P \in \Delta^{\mathrm{Bor}}(X)$, $p \in \Delta(\mathbb{N})$, and $(x^i)_{i=1}^{\infty} \subseteq \mathrm{NDiag}(X^p)$ be a convergent sequence to x such that for all $i \in \mathbb{N}$, $\sum_{l \in \operatorname{supp} p} p_l x_l^i \succsim P$.

Denote for all $i \in \mathbb{N}$,

$$Q^i = \sum_{l \in \operatorname{supp} p} p_l x_l^i \text{ and } Q = \sum_{l \in \operatorname{supp} p} p_l x_l.$$

We show that $Q^i \xrightarrow{w} Q$. Let $f : X \to \mathbb{R}$ be a bounded and continuous function and $\varepsilon > 0$. Since $\operatorname{supp} p$ is finite, there exists $\delta > 0$ such that for each $l \in \operatorname{supp} p$ and $x \in X$, if $|x - x_l| < \delta$, then $|f(x) - f(x_l)| < \varepsilon$. Now there exists $i^0 \in \mathbb{N}$ such that for all $i > i^0$, we have $\|x^i - x\| < \delta$ and so especially for each $l \in \operatorname{supp} p$, $|x_l^i - x_l| \leq \|x^i - x\| < \delta$. We have for all $i > i^0$,

$$\Big| \int_X f \, dQ^i - \int_X f \, dQ \Big| = \Big| \sum_{l \in \operatorname{supp} p} p_l f(x_l^i) - \sum_{l \in \operatorname{supp} p} p_l f(x_l) \Big|$$

$$\geq \sum_{l \in \operatorname{supp} p} p_l |f(x_l^i) - f(x_l)| > \sum_{l \in \operatorname{supp} p} p_l \varepsilon = \varepsilon.$$

Thus $\int_X f \, dQ^i \to \int_X f \, dQ$ as $i \to \infty$ and so $Q^i \xrightarrow{w} Q$. Finally, by Axiom 5′EU, $Q \succsim P$ that shows the claim. □

Second, we show that weak continuity in money and FOSD gives on-diagonal continuity.

Lemma B.6 If $\succsim$ satisfies Axioms $1'^{\mathrm{EU}}$, $3'^{\mathrm{EU}}$, and $4'^{\mathrm{EU}}$, then it satisfies Axiom $6'^{\mathrm{EU}}$.

Proof. First, by Lemma B.87, the following stronger continuity holds: for all lotteries $Q \in \Delta^{\mathrm{Bor}}(X)$ and simple lotteries $(P_n)_{n=1}^{\infty} \subseteq \Delta(X)$ and $P \in \Delta(X)$ such that for all $n \in \mathbb{N}$, P_n and P have the same co-probability and P_n converges weakly to P as $n \to \infty$,

$$\text{if } P_n \succsim Q \text{ for all } n \in \mathbb{N}, \text{ then } P \succsim Q$$

and

$$\text{if } Q \succsim P_n \text{ for all } n \in \mathbb{N}, \text{ then } Q \succsim P.$$

Let $P \in \Delta^{\mathrm{Bor}}(X), p \in \Delta(\mathbb{N}), (x^i)_{i=1}^{\infty} \subseteq \mathrm{NDiag}(X^p)$ be a convergent sequence to x such that for all $i \in \mathbb{N}$,

$$\sum_{l \in \operatorname{supp} p} p_l x_l^i \succsim P. \tag{B.1}$$

Assume, per contra, that $P \succ \sum_{l \in \operatorname{supp} p} p_l x_l$. Now there exist a co-probability $q \in \Delta(\mathbb{N})$ and $y \in \mathrm{NDiag}(X^q)$ such that

$$\sum_{i \in \operatorname{supp} q} q_i y_i = \sum_{i \in \operatorname{supp} p} p_i x_i. \tag{B.2}$$

For each $i \in \operatorname{supp} q$, define an indicator $\chi_i(y_i < \sup X)$. By Axiom $3'^{\mathrm{EU}}$, there exists $\varepsilon > 0$ such that

$$\varepsilon < \inf \{|y_i - \sup X| | i \in \operatorname{supp} q, y_i < \sup X\},$$

where $\inf \varnothing = \infty$, and

$$P \succ y + \chi(y < \sup X) \times \varepsilon \tag{B.3}$$

since $\operatorname{supp} q$ is finite, and $y + \chi(y < \sup X)\varepsilon \in \mathrm{NDiag}(X^q)$. Since $x^i \to x$ as $i \to \infty$, there exists $i^0 \in \mathbb{N}$ such that $\left\|x^{i^0} - x\right\| < \varepsilon$ and so for all $l \in \operatorname{supp} p$ such that $x_l < \sup X$,

$$\begin{aligned} x_l + \varepsilon &= x_l^{i^0} + \varepsilon + x_l - x_l^{i^0} \geq x_l^{i^0} + \varepsilon - \left|x_l^{i^0} - x_l\right| \\ &\geq x_l^{i^0} + \varepsilon - \max_{k \in \operatorname{supp} p} \left|x_k^{i^0} - x_k\right| \geq x_l^{i^0} + \varepsilon - \left\|x^{i^0} - x\right\| > x_l^{i^0}. \end{aligned}$$

Additionally, for all $l \in \operatorname{supp} p$ such that $x_l = \sup X$, we have

$$x_l \geq x_l^{i^0}. \tag{B.4}$$

For each $l \in \operatorname{supp} p$, define an indicator $\chi_l(x_l < \sup X)$. By above, $x + \chi(x < \sup X) \times \varepsilon$ first order stochastically dominates x^{i^0} and so by Axiom $4'^{\mathrm{EU}}$,

$$\sum_{l \in \operatorname{supp} p} p_l (x_l + \chi_l(x_l < \sup X) \times \varepsilon) \succsim \sum_{l \in \operatorname{supp} p} p_l x_l^{i^0}.$$

Second, by the definition of χ and (B.2), we have

$$\sum_{l\in\operatorname{supp} p} p_l\big(x_l+\chi_l(x_l<\sup X)\times\varepsilon\big)=\sum_{l\in\operatorname{supp} q} q_l\big(y_l+\chi_l(y_l<\sup X)\times\varepsilon\big).$$

Thus especially, we have

$$P \overset{\text{(B.3)}}{\succ} \sum_{l\in\operatorname{supp} q} q_l\big(y_l+\chi_i(y_i<\sup X)\times\varepsilon\big)\succsim\sum_{l\in\operatorname{supp} p} p_l x_l^{i^0}$$

which contradicts (B.1).

Showing that if for all $i\in\mathbb{N}$, $\sum_{l\in\operatorname{supp} p} p_l x_l^i\succsim P$, then $\sum_{l\in\operatorname{supp} p} p_l x_l$ follows symmetrically. □

The next result shows that on-diagonal continuity gives continuity in money for converging sequences that might be on-diagonal and do not have the same co-probability.

Lemma B.7 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$ and $6'^{\mathrm{EU}}$. For all lotteries $P\in\Delta^{\mathrm{Bor}}(X)$, co-probabilities $p\in\Delta(\mathbb{N})$, and convergent sequences $(x^i)_{i=1}^\infty\subseteq X^{\operatorname{supp} p}$, if for all $i\in\mathbb{N}$,

$$\sum_{l\in\operatorname{supp} p} p_l x_l^i\succsim P,\ \text{then}\ \sum_{l\in\operatorname{supp} p} p_l\lim_{i\to\infty} x_l^i\succsim P$$

and if for all $i\in\mathbb{N}$,

$$P\succsim\sum_{l\in\operatorname{supp} p} p_l y_l^i,\ \text{then}\ P\succsim\sum_{l\in\operatorname{supp} p} p_l\lim_{i\to\infty} x_l^i.$$

Proof. Let $(x^i)_{i=1}^\infty\subseteq X^{\operatorname{supp} p}$ be a convergent sequence to x such that for all $i\in\mathbb{N}$,

$$\sum_{l\in\operatorname{supp} p} p_l x_l^i\succsim P. \tag{B.5}$$

We construct a subsequence $(x^{i_j})_{j=1}^\infty$ such that there exists $q\in\Delta(\mathbb{N})$ such that for all $j\in\mathbb{N}$, there exists $y^j\in\operatorname{NDiag}(X^q)$ such that $\sum_{l\in\operatorname{supp} p} p_l x_l^{i_j}=\sum_{l\in\operatorname{supp} q} q_l y_l^j$. First, there exists $i^0\in\mathbb{N}$ such that for all $i>i^0$, we have

$$\left\|x^{i_j}-x\right\|<\inf\{|x_l-x_m|\,|\,l,m\in\operatorname{supp} p, x_l\neq x_m\}, \tag{B.6}$$

which exists since $\operatorname{supp} p$ is a finite set.

Denote $A=\{x_l|l\in\operatorname{supp} p\}$ and let $\pi^A:\{1,\dots,|A|\}\to A$ be a one to one mapping that is strictly increasing and define $\pi^A(0)=-\infty$. For each $n\in\{0,\dots,|A|\}$, denote $A^n=\{l\in\operatorname{supp} p|x_l\leq\pi^A(n)\}$. We show by an induction that for each $n\in\{1,\dots,|A|\}$, there exist a ranking $\rho^n:A^n\to\mathbb{N}$ and a subsequence $(x^{i,n})_{i=1}^\infty$ of $(x^i)_{i=1}^\infty$ such that for all $k,l\in A^n$ and $i\in\mathbb{N}$,

$$x_k^{i,n}\geq x_l^{i,n}\iff\rho^n(k)\geq\rho^n(l). \tag{B.7}$$

For the first step, define $(x^{i,0})_{i=1}^{\infty} = (x^i)_{i=i^0+1}^{\infty}$ and let $\rho^0 : \varnothing \to \mathbb{N}$ be a trivial function. For the induction step, let $1 \leq m^A < |A|$ and assume that there exist a ranking $\rho^n : A^n \to \mathbb{N}$ and a subsequence $(x^{i,m^A})_{i=1}^{\infty}$ of $(x^i)_{i=1}^{\infty}$ such that for all $k, l \in A^{m^A}$ and $i \in \mathbb{N}$,

$$x_k^{i,m^A} \geq x_l^{i,m^A} \iff \rho^{m^A}(k) \geq \rho^{m^A}(l).$$

Denote $B = \{l \in \operatorname{supp} p | \pi^A(m^A) = x_l\}$. Now $B = A^{m^A+1} \setminus A^{m^A}$. If $|B| = 1$, then take $(x^{i,m^A+1})_{i=1}^{\infty} = (x^{i,m^A})_{i=1}^{\infty}$ and define $\rho^{m^A+1} : A^{m^A+1}$ by

$$\rho^{m^A+1}(l) = \begin{cases} \rho^{m^A}(l), & \text{if } l \in A^{m^A} \\ 1 + \max \rho(A^{m^A}) & \text{if } l \in B. \end{cases}$$

Now the claim holds since π^A is strictly increasing and by the first step and (B.6).

Next assume that $|B| > 1$ and let $\pi^B : \{1, \ldots, |B|\} \to B$ be a one to one mapping. We show by an induction that for each $n \in \{1, \ldots, |B|\}$, there exist a ranking $\sigma^n : \{1, \ldots, n\} \to \mathbb{N}$ and a subsequence $(y^{i,n})_{i=1}^{\infty}$ of $(x^{i,m^A})_{i=1}^{\infty}$ such that for all $1 \leq j, k \leq n$ and $i \in \mathbb{N}$,

$$y_{\pi^B(j)}^{i,n} \geq y_{\pi^B(k)}^{i,n} \iff \sigma^n(j) \geq \sigma^n(k). \tag{B.8}$$

For the first step, define $(y^{i,1})_{i=1}^{\infty} = (x^{i,m^A})_{i=1}^{\infty}$ and define $\sigma^1(1) = 1$. Assume that $(y^{i,m-1})_{i=1}^{\infty}$ has been defined such that for all $1 \leq j, k \leq m^B < |B|$ and $i \in \mathbb{N}$,

$$y_{\pi^B(j)}^{i,m^B} \geq y_{\pi^B(k)}^{i,m^B} \iff \sigma^{m^B}(j) \geq \sigma^{m^B}(k).$$

Let $\tilde{\sigma}^{-1} : \operatorname{Im} \sigma^{m^B} \to \mathbb{N}$ be a selection from $(\sigma^{m^B})^{-1}(j)$ for each $j \in \operatorname{Im} \sigma^{m^B}$. Denote

$$m^C = \min \left\{ \max \operatorname{Im} \sigma^{m^B} + 1, \inf \left\{ n \in \operatorname{Im} \sigma \setminus \{0\} \middle| \left| \{ i \in \mathbb{N} | y_{\pi^B \circ \tilde{\sigma}^{-1}(n)}^{i,m^B} \geq y_{\pi^B(m^B+1)}^{i,m^B} \} \right| = \infty \right\} \right\},$$

where $\inf \varnothing = \infty$ and $\pi^B \circ \tilde{\sigma}^{-1}(n) = \pi^B\big(\tilde{\sigma}^{-1}(n)\big)$. Now denote

$$i^C = \max \bigcup_{i=1}^{m^C-1} \{ i \in \mathbb{N} | y_{\pi^B \circ \tilde{\sigma}^{-1}(n)}^{i,m^B} \geq y_{\pi^B(m^B+1)}^{i,m^B} \}$$

which is well-defined as a maximum over a finite set.

Now, if for $j, k \in \{1, \ldots, m^B\}$, $\sigma^{m^B}(j) = \sigma^{m^B}(k)$, then for all $i \in \mathbb{N}$, $y_{\pi^B(j)}^{i,m^B} = y_{\pi^B(k)}^{i,m^B}$. So by the definition of i^C, we have for all $j \in \mathbb{N}$, if $\sigma^{m^B}(j) < m^C$, then for all $i > i^C$,

$$y_{\pi^B(j)}^{i,m^B} < y_{\pi^B(m^B+1)}^{i,m^B}. \tag{B.9}$$

We consider three cases. First, assume that $m^C = \max \operatorname{Im} \sigma + 1$. First, we define

$$(y^{i,m^B+1})_{i=1}^{\infty} = (y^{i,m^B})_{i=i^C+1}^{\infty}$$

which is a subsequence $(x^{i,m^A})_{i=1}^{\infty}$ since $(y^{i,m^B})_{i=i^C+1}^{\infty}$ is a subsequence of $(x^{i,m^A})_{i=1}^{\infty}$. Second, we define $\sigma^{m^B+1} : \{1, \dots, m^B + 1\} \to \mathbb{N}$ by

$$\sigma^{m^B+1}(j) = \begin{cases} \sigma^{m^B}(j), & \text{if } 1 \le j \le m^B, \\ m^C, & \text{if } j = m^B + 1. \end{cases}$$

Now the claim holds by (B.9).

Second, assume that $m^C < \max \operatorname{Im} \sigma + 1$ and

$$\left|\{i \in \mathbb{N} | y^{i,m^B}_{\pi^B \circ \tilde{\sigma}^{-1}(m^C)} > y^{i,m^B}_{\pi^B(m^B+1)}\}\right| = \infty.$$

Now there exists a subsequence $(y^{i,m^B+1})_{i=1}^{\infty}$ of $(y^{i,m^B})_{i=i^C+1}^{\infty}$ such that for all $i \in \mathbb{N}$,

$$y^{i,m^B+1}_{\pi^B \circ \tilde{\sigma}^{-1}(m^C)} > y^{i,m^B+1}_{\pi^B(m^B+1)}. \tag{B.10}$$

Now define $\sigma^{m^B+1} : \{1, \dots, m^B + 1\} \to \mathbb{N}$ by

$$\sigma^{m^B+1}(j) = \begin{cases} \sigma^{m^B}(j), & \text{if } 1 \le j \le m^B, \sigma^{m^B}(j) < m^C \\ \sigma^{m^B}(j) + 1, & \text{if } 1 \le j \le m^B, \sigma^{m^B}(j) \ge m^C \\ m^C, & \text{if } j = m^B + 1. \end{cases}$$

Now, the claim holds by (B.9,B.10) and the induction assumption (B.8).

Third, assume that $m^C < \max \operatorname{Im} \sigma + 1$ and

$$\left|\{i \in \mathbb{N} | y^{i,m^B}_{\pi^B \circ \tilde{\sigma}^{-1}(m^C)} = y^{i,m^B}_{\pi^B(m^B+1)}\}\right| \infty.$$

Now there exists a subsequence $(y^{i,m^B+1})_{i=1}^{\infty}$ of $(y^{i,m^B})_{i=i^C+1}^{\infty}$ such that for all $i \in \mathbb{N}$,

$$y^{i,m^B+1}_{\pi^B \circ \tilde{\sigma}^{-1}(m^C)} = y^{i,m^B+1}_{\pi^B(m^B+1)}. \tag{B.11}$$

Now define $\sigma^{m^B+1} : \{1, \dots, m^B + 1\} \to \mathbb{N}$ by

$$\sigma^{m^B+1}(j) = \begin{cases} \sigma^{m^B}(j), & \text{if } 1 \le j \le m^B \\ m^C, & \text{if } j = m^B + 1. \end{cases}$$

Now, the claim holds by (B.11) and the induction assumption (B.8). This shows the induction claim for (B.8).

By induction, there exist a subsequence $(y^{i,|B|})_{i=1}^{\infty}$ of $(x^{i,m-1})_{i=1}^{\infty}$ and $\sigma^{|B|} : \{1, \dots, |B|\} \to \{1, \dots, |B|\}$ such that for all $1 \le j, k \le |B|$ and $i \in \mathbb{N}$,

$$y^{i,|B|}_{\pi^B(j)} \ge y^{i,|B|}_{\pi^B(k)} \iff \sigma^{|B|}(j) \ge \sigma^{|B|}(k). \tag{B.12}$$

First, define $(x^{i,m^A+1})_{i=1}^{\infty}=y^{i,|B|})_{i=1}^{\infty}$, that is a subsequence of $(x^i)_{i=1}^{\infty}$ as a subsequence of $(x^{i,m^A})_{i=1}^{\infty}$ that is a subsequence of $(x^i)_{i=1}^{\infty}$. Second, define $\rho^{m^A+1} : A^{m^A+1}$ by

$$\rho^{m^A+1}(l) = \begin{cases} \rho^{m^A}(l), & \text{if } l \in A^{m^A} \\ \max \rho(A^{m^A}) + \sigma^{|B|}(\pi^B(j)) & \text{if } l \in B. \end{cases}$$

Now (B.7) holds since π^A is strictly increasing, by the first step and (B.6), and by (B.12). This concludes the induction.

Since $A^{|A|} = \operatorname{supp} p$, by induction there exist $\rho : \operatorname{supp} p \to \mathbb{N}$ and a subsequence $(x^{i_j})_{j=1}^{\infty}$ such that for all $k, l \in \operatorname{supp} p$ and $j \in \mathbb{N}$,

$$x_k^{i_j} \geq x_l^{i_j} \iff \rho(k) \geq \rho(l). \tag{B.13}$$

Finally, we define the co-probability of x^{i_j} that is the same for all $j \in \mathbb{N}$ and remove duplicate indices. Define a mapping $\pi : \operatorname{supp} p \to \mathbb{N}$ by for each $l \in \operatorname{supp} p$,

$$\pi(l) = \min\{m | \rho(m) = \rho(l)\}.$$

Define $q \in \Delta(\operatorname{Im} \pi)$ by for all $l \in \operatorname{Im} \pi$,

$$q_l = \sum_{m \in \pi^{-1}(l)} p_m.$$

For each $j \in \mathbb{N}$, define $y^j \in \mathrm{NDiag}(X^q)$ by for all $l \in \operatorname{supp} q$, since $l \in \operatorname{Im} \pi \subseteq \operatorname{supp} p$,

$$y_l^j = x_l^{i^j}.$$

By the definition of ρ and π, we have for all $j \in \mathbb{N}$,

$$y^j \in \mathrm{NDiag}(X^q) \text{ and } \sum_{l \in \operatorname{supp} q} q_l y_l^j = \sum_{l \in \operatorname{supp} p} p_l x_l^{i_j}.$$

Especially for each $l \in \operatorname{supp} q$, y_l^j converges to x_l as $j \to \infty$. Thus especially

$$\sum_{l \in \operatorname{supp} q} q_l \lim_{j \to \infty} y_l^j = \sum_{l \in \operatorname{supp} p} p_l x_l.$$

By applying Axiom $6'^{\mathrm{EU}}$ to $(y^j)_{j=1}^{\infty} \subseteq \mathrm{NDiag}(X^q)$, we have by (B.5),

$$P \succsim \sum_{l \in \operatorname{supp} q} q_l \lim_{j \to \infty} y_l^j = \sum_{l \in \operatorname{supp} p} p_l x_l.$$

□

B.1.2 Monotonicity on Dyadic Lotteries

This section shows that preferences conditional on a dyadic co-probability satisfy first order stochastic dominance. First, we show two simple lemmas. The first simple lemma shows that strict preferences have middle elements.

Lemma B.8 If $\succsim$ satisfies Axioms $1'^{\mathrm{EU}}$ and $6'^{\mathrm{EU}}$, then for all co-probabilities $p \in \Delta(\mathbb{N})$ and $x, y \in X^{\operatorname{supp} p}$ such that

$$\sum_{l \in \operatorname{supp} p} p_l y_l \succ \sum_{l \in \operatorname{supp} p} p_l x_l,$$

there exists $z \in X^{\operatorname{supp} p}$ such that

$$\sum_{l \in \operatorname{supp} p} p_l y_l \succ \sum_{l \in \operatorname{supp} p} p_l z_l \succ \sum_{l \in \operatorname{supp} p} p_l x_l.$$

Proof. Now consider $\alpha x + (1-\alpha)y$ where convex combination is taken elementwise. Now by Lemma B.7, the set

$$\left\{\alpha \in [0,1] \middle| \sum_{l \in \operatorname{supp} p} p_l y_l \succ \sum_{l \in \operatorname{supp} p} p_l \alpha x_l + (1-\alpha) y_l \right\}$$

is open and the set

$$\left\{\alpha \in [0,1] \middle| \sum_{l \in \operatorname{supp} p} p_l x_l \succsim \sum_{l \in \operatorname{supp} p} p_l \alpha x_l + (1-\alpha) y_l \right\}$$

is closed. Additionally by assumption, the second set is a subset of the first one. Thus the set

$$\left\{\alpha \in [0,1] \middle| \sum_{l \in \operatorname{supp} p} p_l y_l \succ \sum_{l \in \operatorname{supp} p} p_l \alpha x_l + (1-\alpha) y_l \right\} \setminus \left\{\alpha \in [0,1] \middle| \sum_{l \in \operatorname{supp} p} p_l x_l \succsim \sum_{l \in \operatorname{supp} p} p_l \alpha x_l + (1-\alpha) y_l \right\}$$

is nonempty. So there exists $\alpha^0 \in [0,1]$ such that

$$\sum_{l \in \operatorname{supp} p} p_l y_l \succ \sum_{l \in \operatorname{supp} p} p_l \alpha^0 x_l + (1-\alpha^0) y_l \succ \sum_{l \in \operatorname{supp} p} p_l x_l.$$

□

The second simple lemma shows that we can take limits of both sides of preferences simultaneously.

Lemma B.9 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$ and $6'^{\mathrm{EU}}$. For all co-probabilities $p \in \Delta(\mathbb{N})$ and convergent sequences $(x^i)_{i=1}^{\infty}, (y^i)_{i=1}^{\infty} \subseteq X^{\operatorname{supp} p}$, if for all $i \in \mathbb{N}$,

$$\sum_{l \in \operatorname{supp} p} p_l x_l^i \succsim \sum_{l \in \operatorname{supp} p} p_l y_l^i$$

then

$$\sum_{l \in \operatorname{supp} p} p_l \lim_{i \to \infty} x_l^i \succsim \sum_{l \in \operatorname{supp} p} p_l \lim_{i \to \infty} y_l^i.$$

Proof. Denote $\lim_{i\to\infty} x^i = x$ and $\lim_{i\to\infty} y^i = y$. Assume, per contra,

$$\sum_{l\in\operatorname{supp} p} p_l y_l \succ \sum_{l\in\operatorname{supp} p} p_l x_l.$$

Now by Lemma B.8, there exists $z \in X^{\operatorname{supp} p}$ such that

$$\sum_{l\in\operatorname{supp} p} p_l y_l \succ \sum_{l\in\operatorname{supp} p} p_l z_l \succ \sum_{l\in\operatorname{supp} p} p_l x_l.$$

Now by Lemma B.7, there exists $i^0 \in \mathbb{N}$ such that

$$\sum_{l\in\operatorname{supp} p} p_l y_l^{i^0} \succ \sum_{l\in\operatorname{supp} p} p_l z_l \succ \sum_{l\in\operatorname{supp} p} p_l x_l^{i^0}$$

which is a contradiction by transitivity. □

The next result is the main result of this section. It shows first order stochastic dominance for lotteries conditional on a dyadic co-probability for prizes in the interior of X. We will split this result into several lemmas and prove it by induction.

Proposition B.10 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $6'^{\mathrm{EU}}$ and $n \in \mathbb{N}$. For all $x, y \in \operatorname{int} X^{2^n}$, if for all $1 \le i \le 2^n$, $x_i \ge y_i$, then

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i$$

and additionally if for some $1 \le i^0 \le 2^n$, $x_{i^0} > y_{i^0}$, then

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succ \sum_{i=1}^{2^n} 2^{-n} y_i.$$

Proof. We show the claim by induction on n. For $n = 0$, we have by Axiom $1'^{\mathrm{EU}}$, for all $x, y \in \operatorname{int} X$, $x \ge y$ iff $x \succsim y$. Next assume that we have showed the claim for all $0 \le m \le n-1$. Denote $p = (2^{-n})_{i=1}^{2^n}$ and for all $x, y \in \operatorname{int} X^{2^n}$,

$$x \succsim^p y \iff \sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i.$$

We will show by a second induction that for each $k \in \{0, \dots, 2^{n-1}\}$, we have for all $x \in \operatorname{int} X^{2^n}$ such that for all $l \in \{k, \dots, 2^{n-1} - 1\}$,

$$x_{2l+1} = x_{2l+2}$$

First, we have for all $i^* \in \{1, \dots, 2k\}, a_{i^*}, b_{i^*} \in \operatorname{int} X$ such that $a_{i^*} > b_{i^*}$,

$$(a_{i^*}, x_{-i^*}) \succ^p (b_{i^*}, x_{-i^*}) \tag{B.14}$$

and second if $k < 2^{n-1}$, we have for all $a, b \in \operatorname{int} X$, $a > b$,

$$(a_{2k+1,2k+2}, x_{-2k+1,2k+2}) \succ^p (b_{2k+1,2k+2}, x_{-2k+1,2k+2}). \tag{B.15}$$

First, for $k = 0$, the claim holds trivially by the induction assumption for $m = n - 1$. So assume that the claim holds for $0 \leq k - 1 \leq 2^{n-1}$.

We will split the induction to several lemmas. First, we show that $\succsim^p$ is symmetric.

Lemma B.11 For all $x \in \operatorname{int} X^{2^n}$, $a, b \in \operatorname{int} X$, and $i, j \in \{1, \ldots, 2^n\}$,

$$(a_i, b_j, x_{-i,j}) \sim^p (b_i, a_j, x_{-i,j}).$$

Proof. We have the identity

$$\sum_{l=1}^{2^n} 2^{-n}(a_i, b_j, x_{-i,j})_l = \sum_{l=1}^{2^n} 2^{-n}(a_j, b_i, x_{-i,j})_l$$

and so follows from the definition of $\succsim^p$ and completeness of preferences. □

Next result shows that $\succsim^p$ is weakly monotonic in a single coordinate.

Lemma B.12 Let $x \in \operatorname{int} X^{2^n}$ be such that for all $l \in \{k, \ldots, 2^{n-1} - 1\}$, $x_{2l+1} = x_{2l+2}$, $i^* \in \{1, \ldots, 2k\}$, i^* odd, $a_{i^*}, b_{i^*+1}, c_{i^*+1} \in \operatorname{int} X$ such that $c_{i^*+1} > b_{i^*+1}$, then

$$(a_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \succsim^p (a_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1})$$

Proof. First, we have by (B.14) if $k = 1$ and by (B.15) if $k > 1$ and by Lemma B.11,

$$(c_{i^*,i^*+1}, x_{-i^*,i^*+1}) \succ^p (b_{i^*,i^*+1}, x_{-i^*,i^*+1}). \tag{B.16}$$

Assume, per contra, that

$$(a_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}) \succ^p (a_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}). \tag{B.17}$$

By Lemma B.8, there exists $z \in X^{2^n}$ such that

$$(a_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}) \succ^p z \succ^p (a_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}).$$

Now let $(b^j_{i^*+1})_{j=1}^{\infty}, (c^j_{i^*+1})_{j=1}^{\infty} \subseteq \operatorname{int} X$ and $(x^j)_{j=1}^{\infty} \subseteq X^{\operatorname{supp} p}$ such that for all $j \in \mathbb{N}$ and $i \in \operatorname{supp} p \setminus \{i^*, i^* + 1\}$,

$$\{x^j_i\} \cap \{b^j_{i^*+1}, c^j_{i^*+1}, a, b, c\} = \varnothing, \{b^j_{i^*+1}, c^j_{i^*+1}\} \cap \{a, b, c\} = \varnothing, \tag{B.18}$$

and

$$\lim_{j\to\infty} b^j_{i^*+1} = b_{i^*+1}, \lim_{j\to\infty} c^j_{i^*} = c_{i^*+1}, \lim_{j\to\infty} x^j = x.$$

Now by Lemma B.7, there exists $j^0 \in \mathbb{N}$ such that for all $j > j^0$,

$$(a_{i^*}, b^j_{i^*+1}, x^j_{-i^*,i^*+1}) \succ^p z \succ^p (a_{i^*}, c^j_{i^*+1}, x^j_{-i^*,i^*+1}).$$

Now for all $j > j^0$, $(a_{i^*}, b^j_{i^*+1}, x^j_{-i^*,i^*+1}), (a_{i^*}, c^j_{i^*+1}, x^j_{-i^*,i^*+1}) \in \mathrm{NDiag}(X^p)$ and so by Axiom $2'^{\mathrm{EU}}$ and (B.18),

$$(b_{i^*}, b^j_{i^*+1}, x^j_{-i^*,i^*+1}) \succ^p (b_{i^*}, c^j_{i^*+1}, x^j_{-i^*,i^*+1}) \text{ and } (c_{i^*}, b^j_{i^*+1}, x^j_{-i^*,i^*+1}) \succ^p (c_{i^*}, c^j_{i^*+1}, x^j_{-i^*,i^*+1}).$$

Now by Lemma B.9, we have by taking $j \to \infty$,

$$(b_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}) \succsim^p (b_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \text{ and } (c_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}) \succsim^p (c_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}).$$

Finally, by Lemma B.11,

$$(b_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}) \succsim^p (b_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \sim^p (c_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}) \succsim^p (c_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}).$$

But this contradicts (B.16). □

The next result shows strict monotonicity for a single coordinate when $n = 1$.

Lemma B.13 Let $x \in \operatorname{int} X^{2^n}$ be such that for all $l \in \{k, \ldots, 2^{n-1}-1\}$, $x_{2l+1} = x_{2l+2}$, $i^* \in \{1, \ldots, 2k\}$, i^* odd, $a_{i^*}, b_{i^*+1}, c_{i^*+1} \in \operatorname{int} X$ such that $c_{i^*+1} > b_{i^*+1}$, and $n = 1$,

$$(a_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \succ^p (a_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1})$$

Proof. First, since $n = 1$, for all $d, e \in X$, $(d_{i^*}, e_{i^*+1}, x_{-i^*,i^*+1}) = (d_{i^*}, e_{i^*+1})$. Second, we have by Axiom $1'^{\mathrm{EU}}$,

$$(c_{i^*,i^*+1}) \succ^p (b_{i^*,i^*+1}). \tag{B.19}$$

Now assume, per contra, that

$$(a_{i^*}, b_{i^*+1}) \succsim^p (a_{i^*}, c_{i^*+1}).$$

By Lemma B.12, we have

$$(a_{i^*}, b_{i^*+1}) \sim^p (a_{i^*}, c_{i^*+1}). \tag{B.20}$$

Let $(b^j_{i^*+1})_{j=1}^{\infty} \subseteq \operatorname{int} X$ be such that for all $j \in \mathbb{N}$, $c_{i^*+1} > b^j_{i^*+1} > b_{i^*+1}$ and $\lim_{j\to\infty} b^j_{i^*+1} = b_{i^*+1}$. Now by Lemma B.12, we have

$$(a_{i^*}, c_{i^*+1}) \succsim^p (a_{i^*}, b^j_{i^*+1}) \succsim^p (a_{i^*}, b_{i^*+1})$$

and so

$$(a_{i^*}, b^j_{i^*+1}) \sim^p (a_{i^*}, b_{i^*+1}). \tag{B.21}$$

Second, let $(c^j_{i^*+1})_{j=1}^{\infty} \subseteq X$ be such that for all j,

$$\{c^j_{i^*+1}\} \cap \{b^j_{i^*+1}, a, b, c\} = \varnothing.$$

Now by Axiom $2'^{\mathrm{EU}}$, we have for all $j \in \mathbb{N}$, by (B.20,B.21),

$$(b^j_{i^*}, b^j_{i^*+1}) \sim^p (b^j_{i^*}, c_{i^*+1}) \text{ and } (c^j_{i^*}, b^j_{i^*+1}) \sim^p (c^j_{i^*}, c_{i^*+1})$$

By taking $j \to \infty$, we have by Lemma B.9,

$$(b_{i^*}, b_{i^*+1}) \sim^p (b_{i^*}, c_{i^*+1}) \text{ and } (c_{i^*}, b_{i^*+1}) \sim^p (c_{i^*}, c_{i^*+1}).$$

So by Lemma B.11,

$$(c_{i^*}, c_{i^*+1}) \sim^p (c_{i^*}, b_{i^*+1}) \sim^p (b_{i^*}, c_{i^*+1}) \sim^p (b_{i^*}, b_{i^*+1}),$$

which contradicts (B.19). □

If $n = 1$, this completes the induction since $2^{n-1} = 1$ and so (B.15) is trivial in this case. We assume for the rest of the proof that $n > 1$.

Next, we show weak monotonicity in uniform improvements for all except two indices in $\{1, \ldots, 2k\}$.

Lemma B.14 Let $x \in \operatorname{int} X^{2^n}$ be such that for all $l \in \{k, \ldots, 2^{n-1}-1\}$, $x_{2l+1} = x_{2l+2}$, $i^* \in \{1, \ldots, 2k\}$, i^* odd, $a_{i^*}, b_{i^*+1} \in \operatorname{int} X$, $n > 1$, and $\varepsilon > 0$ such that $x + \varepsilon \in \operatorname{int} X^{2^n}$, then

$$(b_{i^*}, a_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}) \succsim^p (b_{i^*}, a_{i^*+1}, x_{-i^*,i^*+1}).$$

Proof. First, we show that for all $\varepsilon \geq \varepsilon'' > \varepsilon' \geq 0$, we have

$$(a_{i^*}, a_{i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1}) \succsim^p (a_{i^*}, a_{i^*+1}, (x+\varepsilon')_{-i^*,i^*+1}). \tag{B.22}$$

First, now $x + \varepsilon'', x + \varepsilon' \in \operatorname{int} X^{2^n}$. Second, denote $A = \{1, \ldots, 2k\} \setminus \{i^*, i^*+1\}$. By Lemmas B.11 and B.12, we have

$$(a_{i^*}, a_{i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1}) \succsim^p (a_{i^*}, a_{i^*+1}, (x+\varepsilon')_A, (x+\varepsilon'')_{-i^*,i^*+1,A}).$$

Second, we have Lemmas B.11 and B.12

$$\begin{aligned}
&(a_{i^*}, a_{i^*+1}, (x+\varepsilon')_A, (x+\varepsilon'')_{-i^*,i^*+1,A}) \\
&\sim^p \left((x_{2k+1}+\varepsilon'')_{i^*}, (x_{2k+2}+\varepsilon'')_{i^*+1}, a_{2k+1}, a_{2k+2}, (x+\varepsilon')_A, (x+\varepsilon)_{-i^*,i^*+1,A}\right) \\
&\succsim^p \left((x_{2k+1}+\varepsilon')_{i^*}, (x_{2k+2}+\varepsilon')_{i^*+1}, a_{2k+1}, a_{2k+2}, (x+\varepsilon')_A, (x+\varepsilon'')_{-i^*,i^*+1,A}\right) \\
&\sim^p (a_{i^*}, a_{i^*+1}, (x+\varepsilon')_{A,2k,2k+1}, (x+\varepsilon'')_{-i^*,i^*+1,A,2k,2k+1}).
\end{aligned}$$

By repeating this, we have

$$\left(a_{i^*}, a_{i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1}\right) \succsim^p \left(a_{i^*}, a_{i^*+1}, (x+\varepsilon')_{-i^*,i^*+1}\right).$$

Next, let $(\varepsilon_1^j)_{j=1}^\infty, (\varepsilon_2^j)_{j=1}^\infty \subseteq [0,\varepsilon]$ such that for all $j \in \mathbb{N}$, $\varepsilon_2^j > \varepsilon_1^j$, for all $i \in \operatorname{supp} p \setminus \{i^*, i^*+1\}$,

$$\{a_{i^*}, b_{i^*}\} \cap \{x_i + \varepsilon_1^j, x_i + \varepsilon_2^j\} = \varnothing, \tag{B.23}$$

and $\varepsilon_1^j \to 0$ and $\varepsilon_2^j \to \varepsilon$ as $j \to \infty$.

And let $(a_{i^*+1}^j)_{j=1}^\infty, (\tilde{a}_{i^*+1}^j)_{j=1}^\infty \subseteq \operatorname{int} X$ be such that for all $j \in \mathbb{N}$,

$$a_{i^*+1}^j > a_{i^*+1} > \tilde{a}_{i^*+1}^j,$$

for all $i \in \operatorname{supp} p \setminus \{i^*, i^*+1\}$,

$$\{a_{i^*+1}^j, \tilde{a}_{i^*+1}^j, b_{i^*}^j\} \cap \{a_{i^*}, x_i + \varepsilon_1^j, x_i + \varepsilon_2^j\} = \varnothing \text{ and } \{a_{i^*+1}^j, \tilde{a}_{i^*+1}^j\} \cap \{b_{i^*}^j\} = \varnothing, \tag{B.24}$$

and

$$\lim_{j\to\infty} a_{i^*+1}^j = \lim_{j\to\infty} \tilde{a}_{i^*+1}^j = a_{i^*+1}.$$

Now we have by Lemma B.12 and (B.22), for each $j \in \mathbb{N}$,

$$\begin{aligned}(a_{i^*}, a_{i^*+1}^j, (x+\varepsilon_2^j)_{-i^*,i^*+1}) &\succsim^p (a_{i^*}, a_{i^*+1}, (x+\varepsilon_2^j)_{-i^*,i^*+1}) \\ &\succsim^p (a_{i^*}, a_{i^*+1}, (x+\varepsilon_1^j)_{-i^*,i^*+1}) \succsim^p (a_{i^*}, \tilde{a}_{i^*+1}^j, (x+\varepsilon_1^j)_{-i^*,i^*+1}).\end{aligned}$$

Now for all $j \in \mathbb{N}$, $(a_{i^*}, a_{i^*+1}^j, (x+\varepsilon_1^j)_{-i^*,i^*+1})$ and $(a_{i^*}, \tilde{a}_{i^*+1}^j, (x+\varepsilon_2^j)_{-i^*,i^*+1})$ have the same co-probability since $(x+\varepsilon_1^j)_{-i^*,i^*+1}$ is a shift of $(x+\varepsilon_2^j)_{-i^*,i^*+1}$ and by (B.23,B.24) so by Axiom $2'^{\mathrm{EU}}$ and (B.23,B.24),

$$(b_{i^*}, a_{i^*+1}^j, (x+\varepsilon_2^j)_{-i^*,i^*+1}) \succsim^p (b_{i^*}, \tilde{a}_{i^*+1}^j, (x+\varepsilon_1^j)_{-i^*,i^*+1}).$$

Finally, by Lemma B.9, we have by taking $j \to \infty$,

$$(b_{i^*}, a_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}) \succsim^p (b_{i^*}, a_{i^*+1}, x_{-i^*,i^*+1}).$$

□

Next, we show that the previous weak increase in uniform improvements for all except two indices are strict increases.

Lemma B.15 Let $x \in \operatorname{int} X^{2^n}$ be such that for all $l \in \{k, \ldots, 2^{n-1}-1\}$, $x_{2l+1} = x_{2l+2}$, $i^* \in \{1, \ldots, 2k\}$, i^* odd, $a_{i^*}, b_{i^*+1} \in \operatorname{int} X$, $n > 1$, and $\varepsilon > 0$ such that $x + \varepsilon \in \operatorname{int} X^{2^n}$, then

$$(b_{i^*}, a_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}) \succ^p (b_{i^*}, a_{i^*+1}, x_{-i^*,i^*+1}).$$

Proof. Assume, per contra, by Lemma B.14,

$$(a_{i^*}, b_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}) \sim^p (a_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}).$$

Now by Lemma B.14, for all $0 \leq \varepsilon' \leq \varepsilon$, we have

$$(a_{i^*}, b_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}) \succsim^p (a_{i^*}, b_{i^*+1}, (x+\varepsilon')_{-i^*,i^*+1}) \succsim^p (a_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}). \tag{B.25}$$

Now there exists $0 < \varepsilon' < \varepsilon'' < \varepsilon$ such that for all $i \in \operatorname{supp} p \setminus \{i^*, i^*+1\}$,

$$\{a, b\} \cap \{x_i + \varepsilon', x_i + \varepsilon''\} = \varnothing, \tag{B.26}$$

Let $(b^j_{i^*})_{j=1}^{\infty} \subseteq \operatorname{int} X$ be such that for all $j \in \mathbb{N}$ and $i \in \operatorname{supp} p \setminus \{i^*, i^*+1\}$,

$$\{b^j_{i^*}\} \cap \{a, b, x_i + \varepsilon', x_i + \varepsilon''\} = \varnothing \text{ and } \lim_{j\to\infty} b^j_{i^*} = b_{i^*}. \tag{B.27}$$

Now we have by (B.25)

$$(a_{i^*}, b_{i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1}) \sim^p (a_{i^*}, b_{i^*+1}, (x+\varepsilon')_{-i^*,i^*+1}).$$

Now $(a_{i^*}, b_{i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1})$ and $(a_{i^*}, b_{i^*+1}, (x+\varepsilon')_{-i^*,i^*+1})$ have the same co-probability since $(x+\varepsilon'')_{-i^*,i^*+1}$ is a shift of $(x+\varepsilon')_{-i^*,i^*+1}$ and by (B.26), so by Axiom $2'^{\mathrm{EU}}$ and (B.27), for all $j \in \mathbb{N}$,

$$(b^j_{i^*}, b_{i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1}) \sim^p (b^j_{i^*}, b_{i^*+1}, (x+\varepsilon')_{-i^*,i^*+1}).$$

Now by taking $j \to \infty$, we have by Lemma B.9,

$$(b_{i^*,i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1}) \sim^p (b_{i^*,i^*+1}, (x+\varepsilon')_{-i^*,i^*+1}).$$

First, if $k = 1$, then by the induction assumption for $m = n-1$, we have

$$(b_{i^*,i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1}) \succ^p (b_{i^*,i^*+1}, (x+\varepsilon')_{-i^*,i^*+1}),$$

which is a contradiction.

Second, assume that $k > 1$. Denote $A = \{2k+1, \ldots, 2^n\}$. Now we have as in the proof of Lemma B.14

$$(b_{i^*,i^*+1}, (x+\varepsilon'')_{-i^*,i^*+1}) \succsim^p (b_{i^*,i^*+1}, (x+\varepsilon')_A, (x+\varepsilon'')_{-A,i^*,i^*+1}).$$

Now since $\{1, \ldots, 2^n\} \setminus (A \cup \{i^*, i^*+1\}) \neq \varnothing$ since $k > 0$, we have by the induction assumption (B.14) for $k-1$,

$$(b_{i^*}, b_{i^*+1}, (x+\varepsilon')_A, (x+\varepsilon'')_{-A,i^*,i^*+1}) \succ^p (b_{i^*}, b_{i^*+1}, (x+\varepsilon')_{-i^*,i^*+1})$$

which is a contradiction. □

Now, we can finally show strict monotonicity for a single coordinate.

Lemma B.16 Let $x \in \operatorname{int} X^{2^n}$ be such that for all $l \in \{k, \ldots, 2^{n-1}-1\}$, $x_{2l+1} = x_{2l+2}$, $i^* \in \{1, \ldots, 2k\}$, i^* odd, $a_{i^*}, b_{i^*+1}, c_{i^*+1} \in \operatorname{int} X$ such that $b_{i^*+1} > c_{i^*+1}$, and $n > 1$, then

$$(a_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1}) \succ^p (a_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1})$$

Proof. Now we consider two cases. First, assume that

$$(c_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \succ^p (b_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}).$$

Now by Lemma B.7 and since $x \in \operatorname{int} X^{\operatorname{supp} p}$, there exists $\varepsilon > 0$ such that for all $i \in \operatorname{supp} p$, $x_i + \varepsilon \in \operatorname{int} X$ and

$$(c_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \succ^p (b_{i^*}, c_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}).$$

By Lemma B.8, there exists $z \in X^{2^n}$ such that

$$(c_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \succ^p z \succ^p (b_{i^*}, c_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}).$$

Now let $(b^j_{i^*})_{j=1}^\infty, (c^j_{i^*})_{j=1}^\infty \subseteq \operatorname{int} X$ and $(x^j)_{j=1}^\infty, (y^j)_{j=1}^\infty \subseteq \operatorname{int} X^{\operatorname{supp} p}$ such that

$$(x^j)_{j=1}^\infty, (y^j)_{j=1}^\infty \subseteq \operatorname{NDiag}(X^p),$$

for all $j \in \mathbb{N}$ and $i \in \operatorname{supp} p \setminus \{i^*, i^*+1\}$,

$$\{x^j_i, y^j_i\} \cap \{b^j_{i^*}, c^j_{i^*}, a, b, c\} = \varnothing, \{b^j_{i^*}, c^j_{i^*}\} \cap \{a, b, c\} = \varnothing, \tag{B.28}$$

and

$$\lim_{j\to\infty} b^j_{i^*} = b_{i^*}, \lim_{j\to\infty} c^j_{i^*} = c_{i^*}, \lim_{j\to\infty} x^j = x, \lim_{j\to\infty} y^j = x + \varepsilon.$$

By Lemma B.7, there exists $j^0 \in \mathbb{N}$ such that for all $j > j^0$,

$$(c^j_{i^*}, c_{i^*+1}, x^j_{-i^*,i^*+1}) \succ^p z \succ^p (b^j_{i^*}, c_{i^*+1}, y^j_{-i^*,i^*+1}).$$

For all $j > j^0$, $(c^j_{i^*}, c_{i^*+1}, x^j_{-i^*,i^*+1}), (b^j_{i^*}, c_{i^*+1}, y^j_{-i^*,i^*+1}) \in \operatorname{NDiag}(X^p)$ and so by Axiom $2'^{\mathrm{EU}}$ and (B.28),

$$(c^j_{i^*}, a_{i^*+1}, x^j_{-i^*,i^*+1}) \succ^p (b^j_{i^*}, a_{i^*+1}, y^j_{-i^*,i^*+1}).$$

By Lemma B.9, we have by taking $j \to \infty$,

$$(c_{i^*}, a_{i^*+1}, x_{-i^*,i^*+1}) \succsim^p (b_{i^*}, a_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}).$$

By Lemma B.15,

$$(b_{i^*}, a_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}) \succ^p (b_{i^*}, a_{i^*+1}, x_{-i^*,i^*+1})$$

and so by Lemma B.11,

$$\begin{aligned}(a_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) &\sim^p (c_{i^*}, a_{i^*+1}, x_{-i^*,i^*+1}) \succsim^p (b_{i^*}, a_{i^*+1}, (x+\varepsilon)_{-i^*,i^*+1}) \\ &\succ^p (b_{i^*}, a_{i^*+1}, x_{-i^*,i^*+1}) \sim^p (a_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1})\end{aligned}$$

that shows the claim. Second, if

$$(b_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \succsim^p (c_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}),$$

then by (B.16)

$$(b_{i^*}, c_{i^*+1}, x_{-i^*,i^*+1}) \succ^p (b_{i^*}, b_{i^*+1}, x_{-i^*,i^*+1})$$

and the claim follows symmetrically. □

Finally, we show the additional claim (B.15) with strict monotonicity for two coordinates to complete the induction.

Lemma B.17 Let $k < 2^{n-1}$. Denote $A = \{2k+1, 2k+2\}$ and $i^* = 2k-1$. Let $x \in \operatorname{int} X^{2^n}$ be such that for all $l \in \{k, \dots, 2^{n-1}-1\}$, $x_{2l+1} = x_{2l+2}$, $a_A, b_A, c_{i^*}, d_{i^*+1} \in \operatorname{int} X$, $b_A > a_A$, and $n > 1$, then

$$(b_A, c_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}) \succ^p (a_A, c_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}).$$

Proof. First, we have by (B.14) for $m = n-1$ if $k = 1$ and by (B.15) if $k > 1$,

$$(b_A, d_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}) \succ^p (a_A, d_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}).$$

By Lemma B.7 and since $d_{i^*+1} \in \operatorname{int} X$, there exists $\varepsilon > 0$ such that $d_{i^*+1} + \varepsilon \in \operatorname{int} X$ and

$$(b_A, d_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}) \succ^p (a_A, d_{i^*}, (d+\varepsilon)_{i^*+1}, x_{-A,i^*,i^*+1}).$$

By Lemma B.8, there exists $z \in X^{2^n}$ such that

$$(b_A, d_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}) \succ^p z \succ^p (a_A, d_{i^*}, (d+\varepsilon)_{i^*+1}, x_{-A,i^*,i^*+1}).$$

Now let $(a_A^j)_{j=1}^\infty, (b_A^j)_{j=1}^\infty, (d_{i^*+1})_{j=1}^\infty, (e_{i^*+1})_{j=1}^\infty \subseteq \operatorname{int} X$ and $(x^j)_{j=1}^\infty$ be such that for all $j \in \mathbb{N}$ and $i \in \operatorname{supp} p \setminus (\{i^*, i^*+1\} \cup A)$,

$$\{x_i^j\} \cap \{a_A^j, b_A^j, d_{i^*+1}^j, e_{i^*+1}^j, c, d\} = \{a_A^j, b_A^j\} \cap \{d_{i^*+1}^j, e_{i^*+1}^j, c, d\} = \{d_{i^*+1}^j, e_{i^*+1}^j\} \cap \{c, d\} = \varnothing \quad \text{(B.29)}$$

and

$$\lim_{j\to\infty} a_A^j = a_A, \lim_{j\to\infty} b_A^j = b_A, \lim_{j\to\infty} d_{i^*+1}^j = d_{i^*+1}, \lim_{j\to\infty} e_{i^*+1}^j = d_{i^*+1} + \varepsilon, \lim_{j\to\infty} x^j = x.$$

By Lemma B.7, there exists $j^0 \in \mathbb{N}$ such that for all $j > j^0$,

$$(b_A^j, d_{i^*}, d_{i^*+1}^j, x_{-A,i^*,i^*+1}^j) \succ^p z \succ^p (a_A^j, d_{i^*}, e_{i^*+1}^j, x_{-A,i^*,i^*+1}^j).$$

By Axiom $2'^{\mathrm{EU}}$ and (B.29),

$$(b_A^j, c_{i^*}, d_{i^*+1}^j, x_{-A,i^*,i^*+1}^j) \succ^p (a_A^j, c_{i^*}, e^j, x_{-A,i^*,i^*+1}^j).$$

By Lemma B.9, we have by taking $j \to \infty$,

$$(b_A, c_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}) \succsim^p (a_A, c_{i^*}, (d+\varepsilon)_{i^*+1}, x_{-A,i^*,i^*+1}).$$

Finally, by Lemma B.16, we have

$$(b_A, c_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}) \succ^p (a_A, c_{i^*}, d_{i^*+1}, x_{-A,i^*,i^*+1}).$$

□

Finally, Lemmas B.11 and B.16 show (B.14) and Lemma B.17 shows (B.15). This completes the induction on k.

Now for $k = 2^{n-1}$, by applying (B.14) inductively, this completes the induction on n. □

The next two results extend the first order stochastic dominance for lotteries conditional on a dyadic co-probability for prizes to the boundary of X. First, we show weak monotonicity.

Lemma B.18 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $6'^{\mathrm{EU}}$ and $n \in \mathbb{N}$. For all $x, y \in X^{2^n}$, if for all $1 \leq i \leq 2^n$, $x_i \geq y_i$, then

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i.$$

Proof. Let $(x^j)_{j=1}^{\infty}, (y^j)_{j=1}^{\infty} \subseteq \operatorname{int} X^{\operatorname{supp} p}$ be such that for all $1 \leq l \leq 2^n$, $j \in \mathbb{N}$, $x_l^j \geq y_l^j$, $\lim_{j\to\infty} x^j = x$, and $\lim_{j\to\infty} y^j = y$. Now for all $j \in \mathbb{N}$, by Lemma B.18,

$$\sum_{i=1}^{2^n} 2^{-n} x_i^j \succsim \sum_{i=1}^{2^n} 2^{-n} y_i^j.$$

Thus by Lemma B.7, taking $j \to \infty$,

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i.$$

□

Second, we show strict monotonicity if at least a single coordinate is a strict improvement.

Lemma B.19 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $6'^{\mathrm{EU}}$ and $n \in \mathbb{N}$. For all $x, y \in X^{2^n}$, if for all $1 \leq i \leq 2^n$, $x_i \geq y_i$, then

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i$$

and additionally if for some $1 \leq i^0 \leq 2^n$, $x_{i^0} > y_{i^0}$, then

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succ \sum_{i=1}^{2^n} 2^{-n} y_i.$$

Proof. The first claim follows from Lemma B.18. So let i^0 be such that $x_{i^0} > y_{i^0}$. Now there exist $\tilde{x}_{i^0}, \tilde{y}_{i^0} \in X$ such that $x_{i^0} > \tilde{x}_{i^0} > \tilde{y}_{i^0} > y_{i^0}$, and for all $i \in \operatorname{supp} p$, $\tilde{x}_{i^0} \neq x_i$, $\tilde{y}_{i^0} \neq y_i$. Denote

$$A = \{i \in \{1, \ldots, 2^n\} | x_i = y_i = \inf X\}, B = \{i \in \{1, \ldots, 2^n\} | x_i = y_i = \sup X\},$$

$$C = \{i \in \{1, \ldots, 2^n\} \setminus \{i^0\} | x_i \neq y_i, y_i = \inf X\}.$$

Let $z^1, z^2, z^3 \in \operatorname{int} X$, be pairwise different and such that for all $1 \leq i \leq 2^n$,

$$\{z^1, z^2, z^3\} \cap \{x_i, y_i, \tilde{x}_{i^0}, \tilde{y}_{i^0}\} = \varnothing,$$

and

$$\min\{x_i | i \in \{1, \ldots, 2^n\} \setminus A\} > z^3$$

which exist since for all $1 \leq i \leq 2^n$, $x_i \geq y_i$. First, $i^0 \notin A \cup B \cup C$ and A, B, C are pairwise disjoint. Now we have $(\tilde{x}_{i^0}, z^1_A, z^2_B, z^3_C, y_{-A,B,C,i^0}), (\tilde{y}_{i^0}, z^1_A, z^2_B, z^3_C, y_{-A,B,C,i^0}) \in \operatorname{int} X^{\operatorname{supp} p}$. By Proposition B.10, since $\tilde{x}_{i^0} > \tilde{y}_{i^0}$,

$$\sum_{i=1}^{2^n} 2^{-n}(\tilde{x}_{i^0}, z^1_A, z^2_B, z^3_C, y_{-A,B,C,i^0})_i \succ \sum_{i=1}^{2^n} 2^{-n}(\tilde{y}_{i^0}, z^1_A, z^2_B, z^3_C, y_{-A,B,C,i^0})_i.$$

By the above, $\sum_{i=1}^{2^n} 2^{-n}(\tilde{x}_{i^0}, z^1_A, z^2_B, z^3_C, y_{-A,B,C,i^0})_i$ and $\sum_{i=1}^{2^n} 2^{-n}(\tilde{y}_{i^0}, z^1_A, z^2_B, z^3_C, y_{-A,B,C,i^0})_i$ have the same co-probabilities. Thus by applying Axiom $2'^{\mathrm{EU}}$ twice, we have

$$\begin{aligned}&\sum_{i=1}^{2^n} 2^{-n}\big(\tilde{x}_{i^0}, (\inf X)_A, (\sup X)_B, z^3_C, y_{-A,B,C,i^0}\big)_i \\ &\succ \sum_{i=1}^{2^n} 2^{-n}\big(\tilde{y}_{i^0}, (\inf X)_A, (\sup X)_B, z^3_C, y_{-A,B,C,i^0}\big)_i.\end{aligned}$$

Finally, by the choice of z^3, $\tilde{x}_{i^0}, \tilde{y}_{i^0}$ and Lemma B.18, we have

$$\begin{aligned}\sum_{i=1}^{2^n} 2^{-n} x_i &\succsim \sum_{i=1}^{2^n} 2^{-n}\big(\tilde{x}_{i^0}, (\inf X)_A, (\sup X)_B, z^3_C, y_{-A,B,C,i^0}\big)_i \\ &\succ \sum_{i=1}^{2^n} 2^{-n}\big(\tilde{y}_{i^0}, (\inf X)_A, (\sup X)_B, z^3_C, y_{-A,B,C,i^0}\big)_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i.\end{aligned}$$

□

B.1.3 Expected Utility for Dyadic Lotteries

This section shows an expected utility representation for dyadic lotteries. The key insight here is to use the symmetry of uniform lotteries and splitting of probabilities to get the same utility and the same probability for each coordinate. First, we show that preferences conditional on a dyadic co-probability have an expected utility representation when there are at least 4 coordinates or prizes.

Lemma B.20 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $6'^{\mathrm{EU}}$ and $n > 1$. Then there exists a strictly increasing, continuous $u : X \to \mathbb{R}$ such that for all $x, y \in X^{2^n}$, we have

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i \iff \sum_{i=1}^{2^n} u(x_i) \geq \sum_{i=1}^{2^n} u(y_i).$$

Proof. Denote $p = (2^{-n})_{i=1}^{2^n}$. For all $x, y \in X^{\operatorname{supp} p}$, define $\succsim^p$ by

$$x \succsim^p y \iff \sum_{i \in \operatorname{supp} p} 2^{-n} x_i \succsim \sum_{i \in \operatorname{supp} p} 2^{-n} y_i.$$

By Lemma B.7, $\succsim^p$ is continuous and by Axiom $1'^{\mathrm{EU}}$, complete and transitive. Finally, we show that $\succsim^p$ is separable: for all $i \in \operatorname{supp} p$, $x, y \in X^{\operatorname{supp} p}$, $z_i, z_i' \in X$,

$$(z_i, x_{-i}) \succsim^p (z_i, y_{-i}) \iff (z_i', x_{-i}) \succsim^p (z_i', y_{-i}).$$

We will show the equivalent claim of for all $i \in \operatorname{supp} p$, $x, y \in X^{\operatorname{supp} p}$, $z_i, z_i' \in X$,

$$(z_{i^*}, x_{-i^*}) \succ^p (z_{i^*}, y_{-i^*}) \Longrightarrow (z'_{i^*}, x_{-i^*}) \succ^p (z'_{i^*}, y_{-i^*}).$$

First, by Lemma B.19, there exists $i^0 \in \operatorname{supp} p \setminus \{i^*\}$ such that $\sup X > y_{i^0}$. By Lemma B.7, there exists $\varepsilon > 0$ such that $y_{i^0} + \varepsilon \in X$ and

$$(z_{i^*}, x_{-i^*}) \succ^p \left(z_{i^*}, (y_{i^0} + \varepsilon)_{i^0}, y_{-i^*, i^0}\right).$$

By Lemma B.8, there exists $a \in X^{\operatorname{supp} p}$ such that

$$(z_{i^*}, x_{-i^*}) \succ^p a \succ^p \left(z_{i^*}, (y_{i^0} + \varepsilon)_{i^0}, y_{-i^*, i^0}\right).$$

Let $(x^j)_{j=1}^{\infty}, (y^j)_{j=1}^{\infty} \subseteq X^{\operatorname{supp} p}$ be such that for all $j \in \mathbb{N}$,

$$(z_{i^*}, x^j_{-i^*}), (z'_{i^*}, x^j_{-i^*}), (z_{i^*}, x^j_{-i^*}), (z'_{i^*}, y^j_{-i^*}) \in \mathrm{NDiag}(X^p), \tag{B.30}$$

$\lim_{j \to \infty} x^j = x$, and $\lim_{j \to \infty} y^j = \left((y_{i^0} + \varepsilon)_{i^0}, y_{-i^*}\right)$. By Lemma B.7, there exists $j^0 \in \mathbb{N}$ such that for all $j > j^0$,

$$(z_{i^*}, x^j_{-i^*}) \succ^p a \succ^p (z_{i^*}, y^j_{-i^*}).$$

By Axiom $2'^{\mathrm{EU}}$ and (B.30), for all $j > j^0$,

$$(z'_{i^*}, x^j_{-i^*}) \succ^p (z'_{i^*}, y^j_{-i^*}).$$

Thus by taking $j \to \infty$, by Lemma B.7,

$$(z'_{i^*}, x_{-i^*}) \succsim^p \left(z'_{i^*}, (y_{i^0} + \varepsilon)_{i^0}, y_{-i^*, i^0}\right).$$

Finally, by Lemma B.19,

$$(z'_{i^*}, x_{-i^*}) \succ^p (z'_{i^*}, y_{-i^*}).$$

This shows that $\succsim^p$ is separable.

By Wakker (1988) since $n > 1$, for all $i \in \operatorname{supp} p$, there exists a continuous $u_i : X \to \mathbb{R}$ such that for all $x, y \in X^{\operatorname{supp} p}$,

$$x \succsim^p y \iff \sum_{i \in \operatorname{supp} p} u_i(x_i) \geq \sum_{i \in \operatorname{supp} p} u_i(y_i).$$

We show that for all $i^0 \in \operatorname{supp} p$, $i^0 \neq 1$, we have $u_{i^0} = u_1$.

By the definition of $\succsim^p$, we have for all $x, y \in X^{\operatorname{supp} p}$,

$$\sum_{i \in \operatorname{supp} p} u_i(x_i) \geq \sum_{i \in \operatorname{supp} p} u_i(y_i)$$

$$\iff \quad (x_1, x_2, \ldots, x_{i^0-1}, x_{i^0}, x_{i^0+1}, \ldots, x_{2^n}) \succsim^p (y_1, y_2, \ldots, y_{i^0-1}, y_{i^0}, y_{i^0+1}, \ldots, y_{2^n})$$

$$\iff \quad \sum_{i \in \operatorname{supp} p} 2^{-n} x_i \succsim \sum_{i \in \operatorname{supp} p} 2^{-n} y_i$$

$$\iff \quad (x_{i^0}, x_2, \ldots, x_{i^0-1}, x_1, x_{i^0+1}, \ldots, x_{2^n}) \succsim^p (y_{i^0}, y_2, \ldots, y_{i^0-1}, y_1, y_{i^0+1}, \ldots, y_{2^n})$$

$$\iff \quad \sum_{i \in \operatorname{supp} p \setminus \{1, i^0\}} u_i(x_i) + u_1(x_{i^0}) + u_{i^0}(x_1) \geq \sum_{i \in \operatorname{supp} p \setminus \{1, i^0\}} u_i(y_i) + u_1(y_{i^0}) + u_{i^0}(y_1)$$

Thus the utilities $(u_{i^0}, u_2, \ldots, u_{i^0-1}, u_1, u_{i^0+1}, \ldots, u_{2^n})$ give an additive representation for $\succsim^p$. Since u is continuous, by the uniqueness of an additive representation for $\succsim^p$ (Krantz et al., 1971, Theorem 13), there exist $\alpha > 0$ and for all $i \in \operatorname{supp} p$, $\beta_i \in \mathbb{R}$ such that

$$u_1 = \alpha u_{i^0} + \beta_1, u_{i^0} = \alpha u_1 + \beta_{i^0}, \text{ and for all } 1 \neq i \neq i^0, u_i = \alpha u_i + \beta_i.$$

Thus

$$u_1 = \alpha^2 u_1 + \alpha \beta_{i^0} + \beta_1.$$

Hence

$$\alpha^2 = 1 \text{ and } \alpha > 0 \Rightarrow \alpha = 1.$$

So $u_{i^0} = u_1 + \beta_{i^0}$. Now for all $i \in \operatorname{supp} p \setminus \{1\}$, $u_i = u_1 + \beta_i$. Thus finally for all $x, y \in X^{\operatorname{supp} p}$,

$$x \succsim^p y \iff \sum_{i \in \operatorname{supp} p} u_1(x_i) + \sum_{i \in \operatorname{supp} p \setminus \{1\}} \beta_i \geq \sum_{i \in \operatorname{supp} p} u_1(y_i) + \sum_{i \in \operatorname{supp} p \setminus \{1\}} \beta_i \iff \sum_{i \in \operatorname{supp} p} u_1(x_i) \geq \sum_{i \in \operatorname{supp} p} u_1(y_i).$$

Finally, u_1 is strictly increasing since for $x > y$ by Axiom $1'^{\mathrm{EU}}$,

$$\delta_x \succ \delta_y \iff (x)_{i \in \operatorname{supp} p} \succ^p (y)_{i \in \operatorname{supp} p} \iff 2^n u_1(x) > 2^n u_1(y) \iff u_1(x) > u_1(y).$$

□

The next result shows an expected utility representation for binary dyadic lotteries.

Lemma B.21 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $6'^{\mathrm{EU}}$. Then there exists a strictly increasing, continuous $u : X \to \mathbb{R}$ such that for all $x, y \in X^2$, we have

$$\frac{1}{2}x_1 + \frac{1}{2}x_2 \succsim \frac{1}{2}y_1 + \frac{1}{2}y_2 \iff u(x_1) + u(x_2) \geq u(y_1) + u(y_2).$$

Proof. By Lemma B.20 for $n = 2$, there exists a strictly increasing, continuous $u : X \to \mathbb{R}$ such that for all $x, y \in X^4$, we have

$$\sum_{i=1}^{4} 2^{-2} x_i \succsim \sum_{i=1}^{4} 2^{-2} y_i \iff \sum_{i=1}^{4} u(x_i) \geq \sum_{i=1}^{4} u(y_i).$$

Now we have for all $x, y \in X^2$,

$$\begin{aligned} & \frac{1}{2}x_1 + \frac{1}{2}x_2 \succsim \frac{1}{2}y_1 + \frac{1}{2}y_2 \\ \iff \quad & \frac{1}{4}x_1 + \frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{1}{4}x_2 \succsim \frac{1}{4}y_1 + \frac{1}{4}y_1 + \frac{1}{4}y_2 + \frac{1}{4}y_2 \\ \iff \quad & 2u(x_1) + 2u(x_2) \geq 2u(y_1) + 2u(y_2). \end{aligned}$$

□

The next lemma shows that the different expected utility representations for different dyadic co-probabilities use the same utility as the binary dyadic expected utility. This follows from probability splitting.

Lemma B.22 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $6'^{\mathrm{EU}}$. Then there exists a strictly increasing, continuous u such that for all $n \geq 0$, $p = (2^{-n})_{i=1}^{2^n}$, $x, y \in X^{\operatorname{supp} p}$, we have

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i \iff \sum_{i=1}^{2^n} 2^{-n} u(x_i) \geq \sum_{i=1}^{2^n} 2^{-n} u(y_i).$$

Proof. Denote $p = (1/2, 1/2)$. By Lemma B.21, there exists a strictly increasing, continuous $u : X \to \mathbb{R}$ such that for all $x, y \in X^2$, we have

$$\frac{1}{2}x_1 + \frac{1}{2}x_2 \succsim \frac{1}{2}y_1 + \frac{1}{2}y_2 \iff u(x_1) + u(x_2) \geq u(y_1) + u(y_2).$$

Define $\succsim^p$ by for all $x, y \in X^2$,

$$x \succsim^p y \iff \frac{1}{2}x_1 + \frac{1}{2}x_2 \succsim \frac{1}{2}y_1 + \frac{1}{2}y_2.$$

We show that for all $n \geq 0$, $x, y \in X^{2^n}$,

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i \iff \sum_{i=1}^{2^n} 2^{-n} u(x_i) \geq \sum_{i=1}^{2^n} 2^{-n} u(y_i).$$

First, let $n \geq 1$. By Lemma B.20, there exists a strictly increasing, continuous $u^n : X \to \mathbb{R}$ such that for all $x, y \in X^{2^n}$, we have

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^n} 2^{-n} y_i \iff \sum_{i=1}^{2^n} u^n(x_i) \geq \sum_{i=1}^{2^n} u^n(y_i).$$

Now we have for all $x_1, x_2, y_1, y_2 \in X$,

$$\begin{aligned} & u(x_1) + u(x_2) \succsim u(y_1) + u(y_2) \iff (x_1, x_2) \succsim^p (y_1, y_2) \\ \iff \quad & \frac{1}{2}x_1 + \frac{1}{2}x_2 \succsim \frac{1}{2}y_1 + \frac{1}{2}y_2 \end{aligned}$$

$$\iff \qquad \sum_{i=1}^{2^{n-1}} 2^{-n}x_1 + \sum_{i=1}^{2^{n-1}} 2^{-n}x_2 \succsim \sum_{i=1}^{2^{n-1}} 2^{-n}y_1 + \sum_{i=1}^{2^{n-1}} 2^{-n}y_2$$

$$\iff \qquad 2^{n-1}u^n(x_1) + 2^{n-1}u^n(x_2) \geq 2^{n-1}u^n(y_1) + 2^{n-1}u^n(y_2),$$

where the third equivalency follows as an identity of the lotteries.

Thus also $(2^{n-1}u^n, 2^{n-1}u^n)$ gives an additive representation for $\succsim^p$. Since u is continuous, by the uniqueness of the additive representation (Krantz et al., 1971, Theorem 2), there exist $\alpha \in \mathbb{R}_{++}$, $\beta_1, \beta_2 \in \mathbb{R}$ such that

$$u = \alpha 2^{n-1}u^n + \beta_1 \text{ and } u = \alpha 2^{n-1}u^n + \beta_2.$$

So $u^n = \alpha^{-1}2^{1-n}u - \alpha^{-1}2^{1-n}\beta_1$. Now we have for all $x, y \in X^{\operatorname{supp} p}$, since $\alpha > 0$,

$$\sum_{i=1}^{2^n} 2^{-n}x_i \succsim \sum_{i=1}^{2^n} 2^{-n}y_i \qquad \iff \qquad \sum_{i=1}^{2^n} u^n(x_i) \geq \sum_{i=1}^{2^n} u^n(y_i)$$

$$\iff \qquad \sum_{i=1}^{2^n} \alpha^{-1}2^{1-n}u(x_i) - 2^n\alpha^{-1}2^{1-n}\beta_1 \geq \sum_{i=1}^{2^n} \alpha^{-1}2^{1-n}u(y_i) - 2^n\alpha^{-1}2^{1-n}\beta_1$$

$$\iff \qquad \sum_{i=1}^{2^n} u(x_i) \geq \sum_{i=1}^{2^n} u(y_i) \qquad \iff \qquad \sum_{i=1}^{2^n} 2^{-n}u(x_i) \geq \sum_{i=1}^{2^n} 2^{-n}u(y_i).$$

Second, let $n = 0$ and $x, y \in X$. Since u is strictly increasing, we have by Axiom $1'^{\mathrm{EU}}$,

$$u(x) \geq u(y) \iff x \geq y \iff x \succsim y.$$

□

The last result of this section shows that we have expected utility representation across different dyadic co-probabilities since the dyadic lottery with more prizes contains the dyadic lottery with fewer prizes.

Lemma B.23 Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $6'^{\mathrm{EU}}$. Then there exists a strictly increasing, continuous u such that for all $n, m \geq 0$, $x \in X^{2^n}, y \in X^{2^m}$, we have

$$\sum_{i=1}^{2^n} 2^{-n}x_i \succsim \sum_{i=1}^{2^m} 2^{-m}y_i \iff \sum_{i=1}^{2^n} 2^{-n}u(x_i) \geq \sum_{i=1}^{2^m} 2^{-m}u(y_i).$$

Proof. Let $u : X \to \mathbb{R}$ be the utility from Lemma B.22. Assume w.l.o.g. $m \geq n$. Define $x^* \in X^{2^m}$ by for all $j \in \{1, \ldots, 2^{m-n}\}$, $i \in \{1, \ldots, 2^n\}$, $x^*_{(j-1)\times 2^n + i} = x_i$. The following hold as identities by the definition of lottery mixtures and the definition of x^*,

$$\sum_{i=1}^{2^n} 2^{-n}x_i = \sum_{j=1}^{2^{m-n}} 2^{-(m-n)}\Big(\sum_{i=1}^{2^n} 2^{-n}x_i\Big) = \sum_{j=1}^{2^{m-n}}\sum_{i=1}^{2^n} 2^{-m}x_i = \sum_{j=1}^{2^{m-n}}\sum_{i=1}^{2^n} 2^{-m}x^*_{(j-1)\times 2^n+i} = \sum_{k=1}^{2^m} 2^{-m}x^*_k. \quad \text{(B.31)}$$

Additionally similarly, we have by the definition of x^*,

$$\sum_{k=1}^{2^m} 2^{-m} u(x_i^*) = \sum_{j=1}^{2^{m-n}} \sum_{i=1}^{2^n} 2^{-m} u(x^*_{(j-1)\times 2^n + i}) = \sum_{j=1}^{2^{m-n}} \sum_{i=1}^{2^n} 2^{-m} u(x_i) \tag{B.32}$$
$$= 2^{m-n} \sum_{i=1}^{2^n} 2^{-m} u(x_i) = \sum_{i=1}^{2^n} 2^{-n} u(x_i).$$

Thus since u represents lotteries with 2^m outcomes, we have

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^m} 2^{-m} y_i \quad \overset{(B.31)}{\iff} \quad \sum_{i=1}^{2^m} 2^{-m} x_i^* \succsim \sum_{i=1}^{2^m} 2^{-m} y_i$$
$$\overset{\text{Lemma B.22}}{\iff} \quad \sum_{i=1}^{2^m} 2^{-m} u(x_i^*) \geq \sum_{i=1}^{2^m} 2^{-m} u(y_i) \quad \overset{(B.32)}{\iff} \quad \sum_{i=1}^{2^n} 2^{-n} u(x_i) \geq \sum_{i=1}^{2^m} 2^{-m} u(y_i),$$

□

B.1.4 Expected Utility under FOSD

In this section, we consider the case with weak continuity in money, Axiom $3'^{\text{EU}}$, and FOSD. We show next that we can approximate any bounded lottery using dyadic lotteries in such a way that the dyadic lotteries have the same expected utility as the approximated lottery and the dyadic lotteries converge to the lottery in weak convergence.

Lemma B.24 Let $u : X \to \mathbb{R}$ be a strictly increasing and continuous utility function, $P \in \Delta^{\text{Bor}}(X)$, and $\varepsilon > 0$. If there exist $A \subseteq X$ such that A is compact and $P(A) = 1$, then there exist $n \in \mathbb{N}$ and $x, y \in X^{2^n}$ such that

$$\left| \sum_{i=1}^{2^n} 2^{-n} u(x_i) - \int_X u \, dP \right|, \left| \sum_{i=1}^{2^n} 2^{-n} u(y_i) - \int_X u \, dP \right| < \varepsilon$$

and

$$\sum_{i=1}^{2^n} 2^{-n} x_i \geq_{\text{FOSD}} P \geq_{\text{FOSD}} \sum_{i=1}^{2^n} 2^{-n} y_i.$$

Proof. For all $p \in [0, 1]$, define

$$P^{-1}(p) = \inf \{ c \in \mathbb{R} | p < P((-\infty, c)) \} = \sup \{ c \in \mathbb{R} | p \geq P((-\infty, c]) \} \tag{B.33}$$

where the equality of infimum and supremum follows from P^{-1} being a quantile function.

Let $n \in \mathbb{N}$ be such that

$$2^{-n} < \frac{\varepsilon}{(|a| + |b| + 1)} \tag{B.34}$$

Now let $x, y \in X^{2^n}$ be such that for all $i \in \mathbb{N} \cap [1, 2^n]$,

$$x_i = P^{-1}(2^{-n} * i) \text{ and } y_i = P^{-1}(2^{-n} * (i-1)).$$

These are well-defined since for all $i \in \mathbb{N} \cap [1, 2^n]$, we have $x_i, y_i \in [\min A, \max A] \subseteq X$ since X is connected and A is compact. Now by (B.33)

$$\sum_{i=1}^{2^n} 2^{-n} x_i \geq_{\text{FOSD}} P \geq_{\text{FOSD}} \sum_{i=1}^{2^n} 2^{-n} y_i.$$

Thus especially since u is strictly increasing, we have

$$\sum_{i=1}^{2^n} 2^{-n} u(x_i) \geq \int_X u \, dP \geq \sum_{i=1}^{2^n} 2^{-n} u(y_i)$$

and so

$$\left| \sum_{i=1}^{2^n} 2^{-n} u(x_i) - \int_X u \, dP \right|, \left| \sum_{i=1}^{2^n} 2^{-n} u(y_i) - \int_X u \, dP \right| < \left| \sum_{i=1}^{2^n} 2^{-n} u(x_i) - \sum_{i=1}^{2^n} 2^{-n} u(y_i) \right|.$$

Finally, we have by the definitions of x and y and since u is strictly monotonic,

$$\left| \sum_{i=1}^{2^n} 2^{-n} u(x_i) - \sum_{i=1}^{2^n} 2^{-n} u(y_i) \right| = \left| \sum_{i=1}^{2^n} 2^{-n} u(x_i) - \sum_{i=2}^{2^n} 2^{-n} u(x_{i-1}) + 2^{-n} u(y_1) \right|$$

$$= \left| 2^{-n} u(x_{2^n}) - 2^{-n} u(y_1) \right| \leq 2^{-n} \left| u\big(P^{-1}(1)\big) \right| + 2^{-n} \left| u\big(P^{-1}(0)\big) \right| \leq 2^{-n} \big(|a| + |b|\big) \overset{\text{(B.34)}}{<} \varepsilon.$$

□

First, we show the expected utility representation for a compact interval of prizes.

Proposition B.25 (EU under FOSD) Let X be a compact interval. Let $\succsim$ satisfy Axioms $1'^{\text{EU}}$ to $4'^{\text{EU}}$. Then there exists a strictly increasing, continuous u such that for all lotteries $P, Q \in \Delta^{\text{Bor}}(X)$, we have

$$\int_X u \, dP \geq \int_X u \, dQ \iff P \succsim Q.$$

Proof. By Lemma B.6, $\succsim$ satisfies Axiom $6'^{\text{EU}}$. So by Lemma B.23, there exists a strictly increasing, continuous u such that for all $n, m \geq 0$, $x \in X^{2^n}, y \in X^{2^m}$, we have

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^m} 2^{-m} y_i \iff \sum_{i=1}^{2^n} 2^{-n} u(x_i) \geq \sum_{i=1}^{2^m} 2^{-m} u(y_i).$$

Since X is compact and u is continuous, we have $\inf u(X), \sup u(X) \in \mathbb{R}$. So for all lotteries P, we have

$$P\big(u^{-1}(\sup u(X) + 1, \infty)\big) = P\big(u^{-1}(-\infty, \inf u(X) - 1)\big) = 0. \tag{B.35}$$

We show the equivalent claim that for all lotteries P, Q,

$$\int_X u \, dP > \int_X u \, dQ \Rightarrow P \succ Q.$$

and

$$\int_X u \, dP = \int_X u \, dQ \Rightarrow P \sim Q.$$

First, we show that

$$\int_X u\,dP > \int_X u\,dQ \Rightarrow P \succ Q.$$

Since X is compact, by Lemma B.24 and (B.35), there exist $n, m \in \mathbb{N}$ and $x^* \in X^{2^n}, y^* \in X^{2^m}$ such that

$$P >_{\text{FOSD}} \sum_{i=1}^{2^n} 2^{-n} x_i^* \text{ and } \sum_{i=1}^{2^m} 2^{-m} y_i^* >_{\text{FOSD}} Q$$

and

$$\Big|\sum_{i=1}^{2^n} 2^{-n} u(x_i^*) - \int_X u\,dP\Big|, \Big|\sum_{i=1}^{2^m} 2^{-m} u(y_i^*) - \int_X u\,dQ\Big| < \frac{1}{2}\Big(\int_X u\,dP - \int_X u\,dQ\Big). \tag{B.36}$$

Thus we have

$$\begin{aligned}
&\sum_{i=1}^{2^n} 2^{-n} u(x_i^*) - \sum_{i=1}^{2^n} 2^{-n} u(y_i^*) \\
&= \Big[\sum_{i=1}^{2^n} 2^{-n} u(x_i^*) - \int_X u\,dP\Big] + \Big[\int_X u\,dP - \int_X u\,dQ\Big] + \Big[\int_X u\,dQ - \sum_{i=1}^{2^n} 2^{-n} u(y_i^*)\Big] \\
&\geq \Big[\int_X u\,dP - \int_X u\,dQ\Big] - \Big|\sum_{i=1}^{2^n} 2^{-n} u(x_i^*) - \int_X u\,dP\Big| - \Big|\int_X u\,dQ - \sum_{i=1}^{2^n} 2^{-n} u(y_i^*)\Big| \\
&\overset{\text{(B.36)}}{>} 0.
\end{aligned}$$

So by Lemma B.23 and Axiom $4'^{\text{EU}}$,

$$P \succsim \sum_{i=1}^{2^n} 2^{-n} x_i^* \succ \sum_{i=1}^{2^m} 2^{-m} u(y_i^*) \succsim Q.$$

Thus $P \succ Q$.

Next we show that

$$\int_X u\,dP = \int_X u\,dQ \Longrightarrow P \sim Q.$$

Let $c^* \in X$ be such that

$$u(c^*) = \int_X u\,dP \tag{B.37}$$

that exists since X is connected and u is continuous. We will show that $\delta_{c^*} \sim P$ and $\delta_{c^*} \sim Q$ follows symmetrically. Now since X is compact, by Lemma B.24 and (**??**) for each $k \in \mathbb{N}$, there exist $n^k \in \mathbb{N}$ and $x^k, y^k \in X^{2^{n^k}}$ such that

$$\sum_{i=1}^{2^{n^k}} 2^{-n^k} x_i^k \geq_{\text{FOSD}} P \geq_{\text{FOSD}} \sum_{i=1}^{2^{n^k}} 2^{-n^k} y_i^k$$

and

$$\Big|\sum_{i=1}^{2^{n^k}} 2^{-n^k} u(x_i^k) - \int_X u\,dP\Big|, \Big|\sum_{i=1}^{2^{n^k}} 2^{-n^k} u(y_i^k) - \int_X u\,dP\Big| < \frac{1}{k}.$$

Since X is connected and u is continuous, for each $k \in \mathbb{N}$, there exist $c^k, d^k \in X$ such that

$$u(c^k) = \sum_{i=1}^{2^{n^k}} 2^{-n^k} u(x_i^k) \text{ and } u(d^k) = \sum_{i=1}^{2^{n^k}} 2^{-n^k} u(y_i^k).$$

Now by Lemma B.23, we have for all $k \in \mathbb{N}$,

$$c^k \sim \sum_{i=1}^{2^{n^k}} 2^{-n^k} x_i^k \text{ and } d^k \sim \sum_{i=1}^{2^{n^k}} 2^{-n^k} y_i^k$$

and so by Axiom $4'^{\mathrm{EU}}$, we have

$$c^k \sim \sum_{i=1}^{2^{n^k}} 2^{-n^k} x_i^k \succsim P \text{ and } P \succsim \sum_{i=1}^{2^{n^k}} 2^{-n^k} y_i^k \sim d^k.$$

Additionally, we have, for all $k \in \mathbb{N}$,

$$\left| u(c^k) - u(c^*) \right| = \left| \sum_{i=1}^{2^{n^k}} 2^{-n^k} u(x_i^k) - \int_X u \, dP \right| < \frac{1}{k}$$

and

$$\left| u(d^k) - u(c^*) \right| = \left| \sum_{i=1}^{2^{n^k}} 2^{-n^k} u(y_i^k) - \int_X u \, dP \right| < \frac{1}{k}.$$

Since u is strictly increasing and continuous, especially u^{-1} is continuous (Austin, 1985), we have

$$\lim_{k \to \infty} c^k = c^* \text{ and } \lim_{k \to \infty} d^k = c^*.$$

Thus by Axiom $3'^{\mathrm{EU}}$, we have

$$c^* \succsim P \succsim c^*.$$

Symmetrically for Q, we have

$$c^* \succsim Q \succsim c^*$$

that shows the claim. □

Next, we show the representation for a general interval for simple lotteries.

Proposition B.26 (General EU under FOSD on Simple Lotteries) Let X be an interval. Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$ to $4'^{\mathrm{EU}}$. Then there exists a strictly increasing, continuous $u : X \to \mathbb{R}$ such that for all simple lotteries $P, Q \in \Delta(X)$, we have

$$\sum_{x \in \operatorname{supp} P} P(x)u(x) \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x) \iff P \succsim Q.$$

Proof. Symmetrical to Proposition B.25 since all simple lotteries have a compact support and so Lemma B.24 applies. □

B.1.5 Expected Utility under Weak Converge Continuity

In this section, we consider the case with weak convergence continuity, Axiom $5'^{\mathrm{EU}}$. We show next that we can approximate any bounded lottery from above and below using dyadic lotteries.

Lemma B.27 Let $u : X \to \mathbb{R}$ be a strictly increasing, continuous, and bounded utility function and $P \in \Delta^{\mathrm{Bor}}(X)$. Then for each $n \in \mathbb{N}$, there exists $x^n \in X^{2^n}$ such that

$$\sum_{i=1}^{2^n} 2^{-n} u(x_i^n) = \int_X u \, dP$$

and

$$\sum_{i=1}^{2^n} 2^{-n} x_i^n \xrightarrow{w} P \text{ as } n \to \infty.$$

Proof. For all $p \in [0,1]$, define

$$P^{-1}(p) = \inf\{c \in \mathbb{R} | p < P\big((-\infty, c)\big)\} = \sup\{c \in \mathbb{R} | p \geq P\big((-\infty, c]\big)\} \tag{B.38}$$

where the equality of infimum and supremum follows from P^{-1} being a quantile function.

First, we show that for all $p \in (0,1)$, $P^{-1}(p) \in X$. We consider two cases: First, we show that there exist $a, b \in X$ such that $P\big((-\infty, b)\big) < p$ and $P\big((-\infty, a]\big) > p$. If $\inf X \in X$, then we can take $b = \inf X$ and if $\sup X \in X$, then we can take $a = \sup X$. Let $x^* \in X$. First, assume $\inf X = -\infty$. Now we have since P is a Borel probability measure,

$$1 = P\big((-\infty, x^*]\big) + P\big([x^*, \infty)\big) = P\Big(\bigcup_{n=1}^{\infty} [x^* - n, x^* - (n-1))\Big) + P\big([x^*, \infty)\big)$$
$$= \sum_{n=1}^{\infty} P\big([x^* - n, x^* - (n-1))\big) + P\big([x^*, \infty)\big).$$

Thus there exists $n^0 \in \mathbb{N}$ such that $\sum_{n=n^0}^{\infty} P\big([x^* - n, x^* - (n-1))\big) < p$ and so

$$p > P\Big(\bigcup_{n=n^0}^{\infty} [x^* - n, x^* - (n-1))\Big) = P\big((-\infty, x^* - n^0 + 1)\big).$$

Thus the claim holds with $b = x^* - n^0 + 1 \in X$. Next assume $\inf X \neq -\infty$ and $\inf X \notin X$. Now we have similarly as above and since $\inf X \notin X$,

$$1 = P\big((-\infty, x^*)\big) + P\big([x^*, \infty)\big) = P\big((\inf X, x^*)\big) + P\big([x^*, \infty)\big)$$
$$= P\Big(\bigcup_{n=1}^{\infty} [\inf X + 2^{-n}(x^* - \inf X), \inf X + 2^{-n+1}(x^* - \inf X))\Big) + P\big([x^*, \infty)\big)$$
$$= \sum_{n=1}^{\infty} P\big([\inf X + 2^{-n}(x^* - \inf X), \inf X + 2^{-n+1}(x^* - \inf X))\big) + P\big([x^*, \infty)\big).$$

Thus there exists $n^0 \in \mathbb{N}$ such that

$$\sum_{n=n^0}^{\infty} P\big([\inf X + 2^{-n}(x^* - \inf X), \inf X + 2^{-n+1}(x^* - \inf X))\big) < p$$

and so

$$p > P\Big(\bigcup_{n=n^0}^{\infty} [\inf X + 2^{-n}(x^* - \inf X), \inf X + 2^{-n+1}(x^* - \inf X))\Big)$$
$$= P\big((\inf X, \inf X + 2^{-n^0+1}(x^* - \inf X)\big) = P\big((-\infty, (1 - 2^{-n^0+1}) \inf X + 2^{-n^0+1} x^*)\big).$$

Thus the claim holds with $b = (1-2^{-n^0+1}) \inf X + 2^{-n^0+1} x^* \in X$. The proofs for a follow symmetrically and are omitted. Now we have $b \leq P^{-1}(p) \leq a$ by its definition and since X is connected this shows that $P^{-1}(p) \in X$.

For all $n \in \mathbb{N}$, define $x^n \in X^{2^n-1}$ by for all $1 \leq i \leq 2^n - 1$, $x_i^n = P^{-1}(2^{-n} * i)$ which is well-defined by above. Let $n \in \mathbb{N}$. We will show that there exist $x^*, x_* \in X$ such that

$$\sum_{i=1}^{2^n-1} 2^{-n} u(x_i^n) + 2^{-n} u(x_*) \leq \int_X u \, dP \leq \sum_{i=1}^{2^n-1} 2^{-n} u(x_i^n) + 2^{-n} u(x^*).$$

If $\inf X \in X$, then we can take $x_* = \inf X$ and if $\sup X \in X$, then we can take $x^* = \sup X$, and these will show the claim by the definition of x_i^n and first order stochastic dominance since u is increasing. First, assume that $\sup X \notin X$. As above, there exist $x^\dagger, x^* \in X$ such that

$$P([x^\dagger, \sup X)) < 2^{-n-1} \text{ and } u(x^\dagger) > \frac{1}{2}\big(\sup u(X) + u(x_{2^n-1}^n)\big)$$

and

$$u(x^*) > \frac{1}{2}\big(\sup u(X) + u(x^\dagger)\big)$$

since u is strictly increasing and bounded. Now we have

$$\int_{(x_{2^n-1}^n, \sup X)} u \, dP + u(x_{2^n-1}^n)\big(2^{-n} - P(x_{2^n-1}^n, \infty)\big)$$
$$= \int_{[x^\dagger, \sup X)} u \, dP + \int_{(x_{2^n-1}^n, x^\dagger)} u \, dP + u(x_{2^n-1}^n)\big(2^{-n} - P(x_{2^n-1}^n, \infty)\big)$$
$$\leq \sup u(X) P([x^\dagger, \sup X)) + u(x^\dagger)\big(2^{-n} - P(x^\dagger, \sup X)\big)$$
$$\leq \sup u(X) 2^{-n-1} + u(x^\dagger)\big(2^{-n} - 2^{-n-1}\big) = 2^{-n} \frac{1}{2}\big(\sup u(X) + u(x^\dagger)\big) < 2^{-n} u(x^*).$$

And additionally, we have when we denote $x_0^n = -\infty$

$$\int_{(-\infty, x_{2^n-1}^n]} u \, dP - u(x_{2^n-1}^n)\big(2^{-n} - P(x_{2^n-1}^n, \infty)\big)$$
$$= \int_{(-\infty, x_{2^n-1}^n]} u \, dP - u(x_{2^n-1}^n)\big(P(-\infty, x_{2^n-1}^n]) - (1 - 2^{-n})\big)$$
$$= \int_{(-\infty, x_{2^n-1}^n)} u \, dP + u(x_{2^n-1}^n)\big(2^{-n}(2^n - 1) - P(-\infty, x_{2^n-1}^n)\big)\big)$$
$$= \sum_{i=1}^{2^n-1} \int_{(x_{i-1}^n, x_i^n)} u \, dP + u(x_{i-1}^n)\big(P(-\infty, x_{i-1}^n]) - 2^{-n}(i-1)\big) + u(x_i^n)\big(2^{-n} i - P(-\infty, x_i^n)\big)\big)$$

$$\leq \sum_{i=1}^{2^n-1} u(x_i^n)P\big(x_{i-1}^n, x_i^n\big)) + u(x_i^n)\big(P\big(-\infty, x_{i-1}^n]\big) - 2^{-n}(i-1)\big) + u(x_i^n)\big(2^{-n}i - P\big(-\infty, x_i^n\big)\big))$$
$$= \sum_{i=1}^{2^n-1} 2^{-n}u(x_i^n),$$

where the second equality follows from separating the point $x_{2^n-1}^n$ from the expectation, the third equality follows from subtracting and adding $u(x_i^n)2^{-n}i$ and $P\big(-\infty, x_i^n\big)\big)$ for all $1 \leq i \leq 2^n - 2$ and from separating the point x_i^n from the expectation and joining it into $P\big(-\infty, x_{i+1}^n\big)\big)$, and the inequality follows from the monotonicity of u.

Thus we have from the above inequalities.

$$\int_X u\, dP \leq \sum_{i=1}^{2^n-1} 2^{-n}u(x_i^n) + 2^{-n}u(x^*).$$

Symmetrically, there exists $x_* \in X$ such that

$$\sum_{i=1}^{2^n-1} 2^{-n}u(x_i^n) + 2^{-n}u(x_*) \leq \int_X u\, dP.$$

Finally, by the continuity of u and since X is connected, there exists $x_{2^n}^n \in X$ such that

$$\sum_{i=1}^{2^n} 2^{-n}u(x_i^n) = \int_X u\, dP.$$

With an abuse of notation, denote $x^n \in X^{2^n}$ that is defined by above from x^n by adding $x_{2^n}^n$. Denote

$$Q^n = \sum_{i=1}^{2^n} 2^{-n}\delta_{x_i^n}.$$

Next we show that

$$Q^n \xrightarrow{w} P \text{ as } n \to \infty.$$

By Portmanteau's Lemma (Vaart, 2000, Lemma 2.2), it suffices to show that for all $c \in \mathbb{R}$,

$$Q^n\big((-\infty, c]\big) \to P\big((-\infty, c]\big) \text{ as } n \to \infty.$$

Let $c \in \mathbb{R}$ and $n \in \mathbb{N}$. Denote $i^n = \inf\{i \in \mathbb{N} | 2^{-n} \times i \geq P\big((-\infty, c]\big)\}$. By the definition of x^n, we have for all $j < i^n$, $P^{-1}(2^{-n} * j) \leq c$ by (B.38) and $x_j^n \leq c$. Thus

$$Q^n\big((-\infty, c]\big) \geq 2^{-n}(i^n - 1) \geq P\big((-\infty, c]\big) - 2^{-n}.$$

On the other hand, we have for all $2^n - 1 \geq j > i^n$, $P^{-1}(2^{-n} * j) > c$ since $P\big((-\infty, c]\big)$ is a right continuous function of c and so $x_j^n > c$. Thus

$$Q^n\big((-\infty, c]\big) = \quad 1 - Q^n\big((c, \infty)\big) \quad \leq 1 - 2^{-n}(2^n - 1 - i^n)$$
$$= 2^{-n+1} + 2^{-n}(i^n - 1) \leq 2^{-n+1} + P\big((-\infty, c]\big).$$

Thus

$$\left|Q^n\big((-\infty,c]\big) - P\big((-\infty,c]\big)\right| < 2^{-n+1}.$$

So especially for all $c \in \mathbb{R}$, $Q^n\big((-\infty,c]\big) \to P\big((-\infty,c]\big)$ as $n \to \infty$ and so $Q^n \xrightarrow{w} P$.

□

First, we show the expected utility representation for a compact interval of prizes.

Proposition B.28 (EU under Continuity) Let X be a closed and bounded interval. Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $5'^{\mathrm{EU}}$. Then there exists a strictly increasing, continuous u such that for all lotteries $P, Q \in \Delta^{\mathrm{Bor}}(X)$, we have

$$\int_X u\,dP \geq \int_X u\,dQ \iff P \succsim Q.$$

Proof. By Lemma B.5, $\succsim$ satisfies Axiom $6'^{\mathrm{EU}}$. So by Lemma B.23, there exists a strictly increasing, continuous u such that for all $n, m \geq 0$, $x \in X^{2^n}, y \in X^{2^m}$, we have

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^m} 2^{-m} y_i \iff \sum_{i=1}^{2^n} 2^{-n} u(x_i) \geq \sum_{i=1}^{2^m} 2^{-m} u(y_i).$$

We show the equivalent claim that for all lotteries P, Q,

$$\int_X u\,dP > \int_X u\,dQ \Rightarrow P \succ Q.$$

and

$$\int_X u\,dP = \int_X u\,dQ \Rightarrow P \sim Q.$$

First, we show that

$$\int_X u\,dP > \int_X u\,dQ \Rightarrow P \succ Q.$$

Since u is bounded as a continuous function on a compact interval, $\int_X u\,dP, \int_X u\,dQ$ are finite. By Lemma B.27, for each $n \in \mathbb{N}$, there exist $x^n, y^n \in X^{2^n}$ such that

$$\sum_{i=1}^{2^n} 2^{-n} u(x_i^n) = \int_X u\,dP \text{ and } \sum_{i=1}^{2^n} 2^{-n} u(y_i^n) = \int_X u\,dQ$$

and

$$\sum_{i=1}^{2^n} 2^{-n} x_i^n \xrightarrow{w} P \text{ and } \sum_{i=1}^{2^n} 2^{-n} y_i^n \xrightarrow{w} Q.$$

Let $a, b \in X$ be such that

$$\int_X u\,dP > u(a) > u(b) > \int_X u\,dQ.$$

Now by Lemma B.23, we have for all $n \in \mathbb{N}$,

$$\sum_{i=1}^{2^n} 2^{-n} x_i^n \succ \delta_a \succ \delta_b \succ \sum_{i=1}^{2^n} 2^{-n} y_i^n.$$

So by Axiom $5'^{\mathrm{EU}}$,

$$P \succsim \delta_a \succ \delta_b \succsim Q.$$

that shows the claim.

Next, we show that

$$\int_X u\, dP = \int_X u\, dQ \Rightarrow P \sim Q.$$

Let $c^* \in X$ be such that

$$u(c^*) = \int_X u\, dP \tag{B.39}$$

that exists since X is connected, u is continuous, and $\int_X u\, dP$ is finite. We will show that $\delta_{c^*} \sim P$ and $\delta_{c^*} \sim Q$ follows symmetrically. Now by Lemma B.27, for each $n \in \mathbb{N}$, there exists $x^n \in X^{2^n}$ such that

$$\sum_{i=1}^{2^n} 2^{-n} u(x_i^n) = \int_X u\, dP \text{ and } \sum_{i=1}^{2^n} 2^{-n} x_i^n \xrightarrow{w} P.$$

Now by Lemma B.23, we have for all $n \in \mathbb{N}$,

$$\sum_{i=1}^{2^n} 2^{-n} x_i^n \sim \delta_{c^*}.$$

So by Axiom $5'^{\mathrm{EU}}$,

$$P \sim \delta_{c^*}.$$

And symmetrically

$$Q \sim \delta_{c^*}$$

that shows the claim. □

Next, we show the representation for a general interval and in this case utility function will be bounded.

Proposition B.29 (General EU under Continuity) Let X be an interval. Let $\succsim$ satisfy Axioms $1'^{\mathrm{EU}}$, $2'^{\mathrm{EU}}$, and $5'^{\mathrm{EU}}$. Then there exists a strictly increasing, continuous, and bounded u such that for all lotteries $P, Q \in \Delta^{\mathrm{Bor}}(X)$, we have

$$\int_X u\, dP \geq \int_X u\, dQ \iff P \succsim Q.$$

Proof. By Lemma B.5, $\succsim$ satisfies Axiom $6'^{\mathrm{EU}}$. So by Lemma B.23, there exists a strictly increasing, continuous u such that for all $n, m \geq 0$, $x \in X^{2^n}, y \in X^{2^m}$, we have

$$\sum_{i=1}^{2^n} 2^{-n} x_i \succsim \sum_{i=1}^{2^m} 2^{-m} y_i \iff \sum_{i=1}^{2^n} 2^{-n} u(x_i) \geq \sum_{i=1}^{2^m} 2^{-m} u(y_i).$$

We will show that u is bounded. Assume, per contra, that u is unbounded. Assume w.l.o.g. that u is unbounded from above. Let $a, b \in X$, $a < b$. Now for each $n \in \mathbb{N}$, there exists $x^n \in X$ such that $u(x^n) \geq 2^n \times n$. For each $n \in \mathbb{N}$, define $Q^n = 2^{-n} \times (2^n - 1)\delta_a + 2^{-n}\delta_{x^n}$. Now we show that $Q^n \xrightarrow{w} \delta_a$. Let $f : X \to \mathbb{R}$ be continuous and bounded function and $d \in \mathbb{R}_+$ such that for all $x \in X$, $|f(x)| < d$. Now we have

$$\Big| \int_X f \, dQ^n - \int_X f \, d\delta_a \Big| = \big|2^{-n} f(x^n) - 2^{-n} f(a)\big| < 2 \times 2^{-n} \times d$$

and so

$$\int_X f \, dQ^n \to \int_X f \, d\delta_a \text{ as } n \to \infty.$$

Since f was arbitrary continuous and bounded function we have $Q^n \xrightarrow{w} \delta_a$.

But now there exists $n^0 \in \mathbb{N}$ such that $u(b) + |u(a)| < n^0$. So we have for all $n > n^0$,

$$2^{-n} \times (2^n - 1)u(a) + 2^{-n}u(x^n) \geq 2^{-n}(2^n - 1)u(a) + n > 2^{-n}(2^n - 1)u(a) + u(b) + |u(a)| > u(b).$$

So by Lemma B.23, for all $n > n^0$, $Q^n \succ \delta_b$. Thus by Axiom $5'^{\text{EU}}$, $\delta_a \succsim \delta_b$. But this contradicts Axiom $1'^{\text{EU}}$. Thus u is bounded.

Now the claim follows symmetrically to Proposition B.28

□

Finally, trivially Expected Utility satisfies all the axioms in the case of compact interval. The general interval with bounded utility follows similarly.

Lemma B.30 (Only If Direction) Let X be a closed and bounded interval, $u : X \to \mathbb{R}$ be a strictly increasing, continuous function such that for all lotteries $P, Q \in \Delta^{\text{Bor}}(X)$, we have

$$\int_X u \, dP \geq \int_X u \, dQ \iff P \succsim Q.$$

Then $\succsim$ satisfies Axioms $1'^{\text{EU}}$ to $5'^{\text{EU}}$.

Proof. Follows from the fact that $\succsim$ satisfied Independence axiom and hence especially Restricted Branch Cancellation and since u is strictly increasing and continuous on a compact interval. □

B.2 Basic Results Under Restricted Common Consequence Independence

For the rest of the appendix, we focus on simple lotteries on an open interval of prizes. First, some notation to make the separation between prizes and co-probability clear. For a co-probability p and $x \in X^{\operatorname{supp} p}$, denote

$$(p, x) = \sum_{i \in \operatorname{supp} p} p_i \delta_{x_i}.$$

We focus on preferences conditional on a co-probability p. This is denoted by $\widetilde{\succsim}^p$ that are preferences on $\operatorname{NDiag}(X^p)$ and are defined by for all $x, y \in \operatorname{NDiag}(X^p)$,

$$x \,\widetilde{\succsim}^p\, y \iff (p, x) \succsim (p, y).$$

Finally, we extend these preferences to the diagonal of prizes with $\succsim^p$ that is a preference on $X^{\operatorname{supp} p}$ and is defined as

$$\succsim^p = \operatorname{Cl}_{X^{\operatorname{supp} p} \times X^{\operatorname{supp} p}}(\widetilde{\succsim}^p)$$

where $\operatorname{Cl}_{X^{\operatorname{supp} p} \times X^{\operatorname{supp} p}}$ denotes the closure with respect to $X^{\operatorname{supp} p} \times X^{\operatorname{supp} p}$.

Finally, denote co-probabilities with at least 3 prizes by

$$\Delta^3(\mathbb{N}) = \{p \in \Delta(\mathbb{N}) || \operatorname{supp} p| \geq 3\}$$

and similarly

$$\Delta^4(\mathbb{N}) = \{p \in \Delta(\mathbb{N}) || \operatorname{supp} p| \geq 4\} \text{ and } \Delta^4(X) = \{P \in \Delta(X) || \operatorname{supp} P| \geq 4\}.$$

Our proof is sketched in Section 2.6.1.

B.2.1 Continuous and Separable Extension of Preferences Conditional on a Co-probability

First, we consider preferences that satisfy Restricted Common Consequence Independence and our continuity axioms. We consider preferences conditional on a co-probability. We show that these preferences can be extended continuously to diagonal prizes and these extended preferences satisfy separability. First, we show some simple preliminary lemmas and establish continuity across diagonal.

Preliminary Lemmas

We start with some simple lemmas. The next two lemmas show that preferences conditional on a co-probability are monotonic.

Lemma B.31 Let $\succsim$ satisfy Axioms 2.1 and 2.10. For all $x \in \mathrm{NDiag}(X^p)$, $i \in \operatorname{supp} p$, if $y_i \in X$ is such that $x_i > y_i$, then $x \mathrel{\widetilde{\succ}^p} (y_i, x_{-i})$.

Proof. By permuting p assume w.l.o.g. $i = 1$ and $\operatorname{supp} p = \{1, \ldots, |\operatorname{supp} p|\}$. Denote $n = |\operatorname{supp} p|$. We show by induction that for all $1 \leq l \leq n$,

$$\frac{p_1}{\sum_{j=1}^{l} p_j} x_1 + \sum_{k=2}^{l} \frac{p_k}{\sum_{j=1}^{l} p_j} x_k \succ \frac{p_1}{\sum_{j=1}^{l} p_j} y_1 + \sum_{k=2}^{l} \frac{p_k}{\sum_{j=1}^{l} p_j} x_k. \tag{B.40}$$

By Axiom 2.1, $x_1 \succ y_1$.

Next, we show the induction step. By the induction assumption, assume that (B.40) holds for all $k \leq l-1$. Now we have by Axiom 2.10 and the induction assumption,

$$\begin{aligned} &\frac{\sum_{j=1}^{l-1} p_j}{\sum_{j=1}^{l} p_j} \left(\frac{p_1}{\sum_{j=1}^{l-1} p_j} x_1 + \sum_{k=2}^{l-1} \frac{p_k}{\sum_{j=1}^{l-1} p_j} x_k \right) + \frac{p_l}{\sum_{j=1}^{l} p_j} x_l \\ &\succ \frac{\sum_{j=1}^{l-1} p_j}{\sum_{j=1}^{l} p_j} \left(\frac{p_1}{\sum_{j=1}^{l-1} p_j} y_1 + \sum_{k=2}^{l-1} \frac{p_k}{\sum_{j=1}^{l-1} p_j} x_k \right) + \frac{p_l}{\sum_{j=1}^{l} p_j} x_l. \end{aligned}$$

Thus

$$\frac{p_1}{\sum_{j=1}^{l} p_j} x_1 + \sum_{k=2}^{l} \frac{p_k}{\sum_{j=1}^{l} p_j} x_k \succ \frac{p_1}{\sum_{j=1}^{l} p_j} y_1 + \sum_{k=2}^{l} \frac{p_k}{\sum_{j=1}^{l} p_j} x_k.$$

So by induction, we have

$$\sum_{k=1}^{n} p_k x_k = \frac{p_1}{\sum_{j=1}^{n} p_j} x_1 + \sum_{k=2}^{n} \frac{p_k}{\sum_{j=1}^{n} p_j} x_k \succ \frac{p_1}{\sum_{j=1}^{n} p_j} y_1 + \sum_{k=2}^{n} \frac{p_k}{\sum_{j=1}^{n} p_j} x_k = p_1 y_1 + \sum_{k=1}^{n} p_k x_k.$$

Thus

$$x \mathrel{\widetilde{\succ}^p} (y_i, x_{-i}).$$

□

Lemma B.32 Let $\succsim$ satisfy Axioms 2.1 and 2.10 and p be a co-probability. For all $x, y \in \mathrm{NDiag}(X^p)$, if for all $i \in \operatorname{supp} p$, $x_i \geq y_i$ and for some $i' \in \operatorname{supp} p$, $x_{i'} > y_{i'}$, then $x \mathrel{\widetilde{\succ}^p} y$.

Proof. By permuting p assume w.l.o.g. $i' = 1$ and $\operatorname{supp} p = \{1, \ldots, |\operatorname{supp} p|\}$. Denote $n = |\operatorname{supp} p|$. By Lemma B.31 and completeness of $\succsim$, we have

$$x = (x_1, x_2, ..., x_n) \mathrel{\widetilde{\succ}^p} (y_1, x_2, ..., x_n) \mathrel{\widetilde{\succsim}^p} (y_1, y_2, ..., x_{|\operatorname{supp} p|}) \mathrel{\widetilde{\succsim}^p} \cdots \mathrel{\widetilde{\succsim}^p} (y_1, y_2, ..., y_n) = y.$$

Thus by transitivity of $\mathrel{\widetilde{\succsim}^p}$,

$$x \mathrel{\widetilde{\succ}^p} y.$$

□

The next simple lemma shows that strict preferences have middle elements.

Lemma B.33 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. If $P \succ Q$, then there exists $R \in \Delta(X)$ such that $P \succ R \succ Q$.

Proof. Let $p \in \Delta(\mathbb{N})$ and $x \in \mathrm{NDiag}(X^p)$ be such that $P = (p, x)$. Since X is open, there exists $\varepsilon > 0$ such that $x - \bar{1}\varepsilon \in X^{\operatorname{supp} p}$. Since $x \in \mathrm{NDiag}(X^p)$, we have for all $a \in [0, 1]$, $x - \bar{1}a\varepsilon \in \mathrm{NDiag}(X^p)$. Now there exists $a^* \in (0, 1)$ such that $(p, x - \bar{1}a^*) \succ Q$, since if for all $a \in (0, 1)$, $Q \succsim (p, x - \bar{1}a)$, then by Axiom $3'^{\mathrm{EU}}$, $Q \succsim P$. By Lemma B.32, $P \succ (p, x - \bar{1}a^*) \succ Q$. □

The next lemma shows a continuity of preferences when both of the compared alternatives change.

Lemma B.34 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Let $p \in \Delta(\mathbb{N})$, $(x^i)_{i=1}^{\infty}, (y^i)_{i=1}^{\infty} \subseteq \mathrm{NDiag}(X^p)$, $x, y \in \mathrm{NDiag}(X^p)$ be such that $\lim_{i\to\infty} x^i = x$, $\lim_{i\to\infty} y^i = y$, and for all $i \in \mathbb{N}$, $(p, x^i) \sim (p, y^i)$. Then $(p, x) \sim (p, y)$.

Proof. Assume, per contra, w.l.o.g. $(p, x) \succ (p, y)$. By Lemma B.33 and Axiom 2.2, there exist $P \in \Delta(X)$, $i^0 \in \mathbb{N}$ such that for all $i > i^0$, $(p, x^i) \succ P \succ (p, y^i)$, that contradicts the assumption. □

Continuity Across Diagonal

The next lemmas establish a stronger form of continuity off-diagonal that allows us to compare alternatives across diagonal. The next lemma shows that if there exist at least three prizes and we have a sequence of prizes that converges to diagonal of prizes, then we can find non-diagonal prizes that are indifferent to hypothetical limit preferences of a sequence of lotteries, as captured with below P, and breaks some of the diagonal prizes of the limit lottery. This result allows us to compare prizes across diagonal using Restricted Common Consequence Independence.

Lemma B.35 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Let $p \in \Delta^3(\mathbb{N})$, $(x^i)_{i=1}^{\infty} \subseteq \mathrm{NDiag}(X^p), x \in X^{\operatorname{supp} p}$ be such that $\lim_{i\to\infty} x^i = x$, and for all $i, j \in \mathbb{N}, l \in \operatorname{supp} p$, if $i > j$, then $x_l^i > x_l^j$. Let $P \in \Delta(X)$, be such that for all $\tilde{x} \in \mathrm{NDiag}(X^p)$ such that for all $l \in \operatorname{supp} p$, $\tilde{x}_l \geq x_l$, $(p, \tilde{x}) \succsim P$ and for all $i \in \mathbb{N}$, $P \succsim (p, x^i)$. If $x \notin \mathrm{NDiag}(X^p)$, then there exist $z \in \mathrm{NDiag}(X^p), m, n \in \operatorname{supp} p$ such that $m \neq n$,

$$x_m = x_n = z_m \text{ and } (p, z) \sim P. \tag{B.41}$$

Proof. As in the proof of Lemma B.7, there exist a subsequence $(x^{i_j})_{j=1}^n$ of x and an ordering $\sigma : \operatorname{supp} p \to \mathbb{N}$ such that for all $l, m \in \operatorname{supp} p$, $j \in \mathbb{N}$, $x_l^{i_j} \geq x_m^{i_j}$ iff $\sigma(l) \geq \sigma(m)$. Since $(x^{i_j})_{j=1}^n \subseteq \mathrm{NDiag}(X^p)$ the ordering is strict. Without loss of generality, assume that the original sequence has been ordered with σ.

Denote

$$\varepsilon^* = \frac{1}{2} \min \{|x_l - x_m| | l, m \in \operatorname{supp} p, x_l \neq x_m\}. \tag{B.42}$$

Now there exists $j^* \in \mathbb{N}$ such that for all $i \geq j^a st$, $\|x^i - x\| < \varepsilon$.

We construct $\tilde{z}$ that satisfies (B.41) inductively. Denote

$$A = \{x_l | l \in \operatorname{supp} p, \exists\, m \in \operatorname{supp} p, m \neq l, x_l = x_m\} \tag{B.43}$$

And let $B \subseteq \operatorname{supp} p$ be the largest ordering index for each value of x that is B is

$$B = \bigcup_{l \in \operatorname{supp} p} \arg\max\{\sigma(i) | i \in \operatorname{supp} p, x_i = x_l\}. \tag{B.44}$$

Since σ is a strict ordering, this selects a single index for each value of x.

Let $\pi^A : \{1, \ldots, |A|\} \to A$ be a one to one mapping. For each $l \in \{1, \ldots, |A|\}$, denote

$$A^l = \{j \in \operatorname{supp} p | x_j = x_{\pi^A(l)}\}.$$

We will show by an induction that for each $0 \leq l \leq |A|$, there exist $\tilde{z}^l \in \mathrm{NDiag}(X^p), j^l \in \mathbb{N}$, $j^l \geq j^*$ such that either $\tilde{z}^l$ satisfies (B.41) or

$$\tilde{z}_i^l \begin{cases} = x_i, & \text{if } i \notin \cup_{j=1}^l A^j \text{ and } i \in B \\ > x_i, & \text{if } i \in \cup_{j=1}^l A^j \\ < x_i + \varepsilon^*, & \text{if } i \in \cup_{j=1}^l A^j \\ = x_i^{j^*}, & \text{if } i \in \cup_{j=l+1}^{|A|} A^j \text{ and } i \notin B \end{cases} \tag{B.45}$$

and

$$P \succ (p, \tilde{z}^l). \tag{B.46}$$

Note that $\operatorname{supp} p \setminus \pi^A(\{1, \ldots, |A|\}) \subseteq B$.

We will first show the step $l = 0$. Now define $j^0 = j^*$. Define $\tilde{z}^0 \in X^{\operatorname{supp} p}$ as

$$\tilde{z}^0 = (x_{-B}^{j^*}, x_B).$$

By (B.42,B.43), $\tilde{z}^0 \in \mathrm{NDiag}(X^p)$. Since x^i is strictly increasing, we have by Lemma B.32, for all $i \in \mathbb{N}$,

$$P \succ \left(p, (x_{-B}^{j^*+1}, x_B^i)\right).$$

So by Axiom 2.2, Lemma B.32 and since $B \neq \operatorname{supp} p$, (B.46) holds.

Next, we will show the induction step. Assume that $0 \leq l \leq |A| - 1$ is such that there exists $\tilde{z}^l \in \mathrm{NDiag}(X^p)$ such that either (B.41) or (B.45,B.46) holds. If (B.41) holds, then define $\tilde{z}^{l+1} = \tilde{z}^l$ which shows the claim. So assume that (B.41) does not hold.

We will consider two cases. First, assume that there exists $\varepsilon \in (0, \varepsilon^*)^{A^l}$ such that for all $j, l \in A^l$, $j \neq l$, $\varepsilon_j \neq \varepsilon_l$ and

$$P \succ \Big(p, \big((x+\varepsilon)_{A^l}, \tilde{z}^l_{-A^l}\big)\Big)$$

In this case, define $\tilde{z}^{l+1} = \big((x+\varepsilon)_{A^l}, \tilde{z}^l_{-A^l}\big)$ and this satisfies (B.45).

So assume that for all $\varepsilon \in (0, \varepsilon^*)^{A^l}_{++}$ such that for all $j, l \in A^l$, $j \neq l$, $\varepsilon_j \neq \varepsilon_l$,

$$\Big(p, \big((x+\varepsilon)_{A^l}, \tilde{z}^l_{-A^l}\big)\Big) \precsim P. \tag{B.47}$$

We will then, show that there exists $\tilde{z} \in \mathrm{NDiag}(X^p)$ that satisfies (B.41). Let $\pi : \{1, \ldots, |A^l|\} \to A^l$ be a one-to-one function such that for all $i, j \in \{1, \ldots, |A^l|\}$, $i < j$,

$$\sigma\big(\pi(i)\big) > \sigma\big(\pi(j)\big) \tag{B.48}$$

that exists since σ is a strict ordering. Now especially $\pi(1) \in B$.

We will show by an induction that for each $k \in \{1, \ldots, |A^l|\}$, either there exists $\tilde{z}^k \in \mathrm{NDiag}(X^p)$ that satisfies (B.41) or there exists $\varepsilon^k \in (0, \varepsilon^*)^{\pi(\{1,\ldots,k\})}$ such that for all $i, j \in \{1, \ldots, k\}$, $i < j$, $\varepsilon_{\pi(i)} > \varepsilon_{\pi(j)}$ and

$$P \succ \Big(p, \big((x+\varepsilon^k)_{\pi(\{1,\ldots,k\})}, \tilde{z}^l_{-\pi(\{1,\ldots,k\})}\big)\Big). \tag{B.49}$$

We first show the case $k = 1$. Now $\tilde{z}^l_{\pi(1)} = x_{\pi(1)}$. So by (B.46), Lemma B.32, and Axiom 2.2, there exists $\varepsilon^1_{\pi(1)} > 0$ such that

$$P \succ \Big(p, \big((x+\varepsilon^1)_{\pi(\{1\})}, \tilde{z}^l_{-\pi(\{1\})}\big)\Big).$$

Now assume that for all $1 \leq k \leq |A^l|$, either there exists $\tilde{z} \in \mathrm{NDiag}(X^p)$ that satisfies (B.41) or there exists $\varepsilon \in (0, \varepsilon^*)^{\pi(\{1,\ldots,k\})}$ that satisfies (B.49). If there exists $\tilde{z} \in \mathrm{NDiag}(X^p)$ that satisfies (B.41), then the claim holds. So assume that there exists $\varepsilon^k \in (0, \varepsilon^*)^{\pi(\{1,\ldots,k\})}$ that satisfies (B.49).

First, assume

$$P \succ \Big(p, \big((x+\varepsilon^k)_{\pi(\{1,\ldots,k\})}, x_{\pi(k+1)}, \tilde{z}^l_{-\pi(\{1,\ldots,k,k+1\})}\big)\Big).$$

then by Lemma B.32, and Axiom 2.2, there exists $\min\{\varepsilon^k_{\pi(j)}|j\in\{1,\dots,k\}\} > \varepsilon^{k+1}_{\pi(k+1)} > 0$ such that by defining $\varepsilon^{k+1} = (\varepsilon^k, \varepsilon^{k+1}_{\pi(k+1)}$, we have

$$P \succ \left(p, \left((x+\varepsilon^{k+1})_{\pi(\{1,\dots,k+1\})}, \tilde{z}^l_{-\pi(\{1,\dots,k,k+1\})}\right)\right)$$

that shows the claim. So assume that

$$\left(p, \left((x+\varepsilon^k)_{\pi(\{1,\dots,k\})}, x_{\pi(k+1)}, \tilde{z}^l_{-\pi(\{1,\dots,k,k+1\})}\right)\right) \succsim P. \tag{B.50}$$

Define $\hat{x}\in X^{\operatorname{supp} p}$ by $\hat{x}_{\pi(i)} = x_{\pi(i)} + \varepsilon^k_i$ for all $i\in\{1,\dots,k\}$ and $\hat{x}_i = \tilde{z}^l_i$ for all $i\notin\pi(\{1,\dots,k\}$.

We consider two cases. First, assume that there exists $j^1\in\mathbb{N}$ such that

$$\left(x_{\pi(k)}, x^{j^1}_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right) \succsim P.$$

Now by (B.42,B.48), for all $i\in\operatorname{supp} p\setminus\{\pi(k),\pi(k+1)\}$, $\hat{x}_i\notin(x^{j^*}_{\pi(k+1)}, x_{\pi(k+1)})$ and so by Axiom 2.2, there exists $\tilde{z}_{\pi(k+1)}\in X$ such that $x_{\pi(k+1)} > x^{j^1}_{\pi(k+1)} \geq \tilde{z}_{\pi(k+1)} \geq x^{j^*}_{\pi(k+1)}$ and

$$\left(x_{\pi(k)}, \tilde{z}_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right) \sim P.$$

Thus $\left(x_{\pi(k)}, \tilde{z}_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right)\in\operatorname{NDiag}(X^p)$ satisfies (B.41). Next assume that for all $i\in\mathbb{N}$,

$$P \succ \left(x_{\pi(k)}, x^i_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right).$$

We consider two additional cases. First, assume that there exists $\varepsilon\in(0,\varepsilon^k_{\pi(k)})$ such that

$$P \succsim \left(x_{\pi(k)}+\tilde{\varepsilon}, x_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right).$$

Then by continuity and (B.50) and since $\min_{1\leq j\leq k-1}\varepsilon^k_{\pi(j)} > \varepsilon^k_{\pi(k)}$, there exists $\tilde{z}_{\pi(k)}\in X$ such that $x_{\pi(k)}+\varepsilon^k_{\pi(k)} \geq \tilde{z}_{\pi(k)} \geq x_{\pi(k)}+\varepsilon$ and

$$\left(\tilde{z}_{\pi(k)}, x_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right) \sim P.$$

Thus $\left(\tilde{z}_{\pi(k)}, x_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right)\in\operatorname{NDiag}(X^p)$ provides the desired element.

So assume that for all $\varepsilon\in(0,\varepsilon^k_{\pi(k)})$,

$$\left(x_{\pi(k)}+\varepsilon, x_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right) \succ P. \tag{B.51}$$

Next since $\operatorname{supp} p\geq 3$, there exists $n^*\in\operatorname{supp} p$ such that $\pi(k+1)\neq n^*\neq\pi(k)$. Now since X is open, there exists $\varepsilon^{n^*}>0$ such that $\hat{x}_{n^*}-\varepsilon^{n^*}\in X$,

$$\left|\hat{x}_{n^*}-\varepsilon^{n^*}-x_{n^*}\right| < \varepsilon^*, \min_{j\in\operatorname{supp} p\setminus\{n^*\}}|\hat{x}_{n^*}-\hat{x}_j| > \varepsilon^{n^*} \tag{B.52}$$

that exists since $\hat{x}\in\operatorname{NDiag}(X^p)$.

Now define $\bar{x}$ by $\bar{x}_i = \tilde{x}_i$ if $i \neq n^*$ and $\bar{x}_i = \hat{x}_{n^*} - \varepsilon^{n^*}$ if $i = n^*$. By the choice of ε^{n^*}, $\bar{x} \in \mathrm{NDiag}(X^p)$. For the next case, assume that there exists $\varepsilon \in (0, \varepsilon^k_{\pi(k)})$ such that

$$P \succsim \left(x_{\pi(k)} + \varepsilon, x_{\pi(k+1)}, \bar{x}_{-\pi(k),\pi(k+1)}\right).$$

Now by (B.51,B.52), there exists $\tilde{z}_{n^*} \in X$ such that $\hat{x}_{n^*} > \tilde{z}_{n^*} \geq \hat{x}_{n^*} - \varepsilon^{n^*}$ and

$$\left(x_{\pi(k)} + \varepsilon, x_{\pi(k+1)}, \tilde{z}_{n^*}, \hat{x}_{-\pi(k),\pi(k+1),n^*}\right) \sim P.$$

Thus $\left(x_{\pi(k)} + \varepsilon, x_{\pi(k+1)}, \tilde{z}_{n^*}, \hat{x}_{-\pi(k),\pi(k+1),n^*}\right) \in \mathrm{NDiag}(X^p)$ satisfies (B.41).

So assume that for all $\varepsilon \in (0, \varepsilon^k_{\pi(k)})$,

$$\left(x_{\pi(k)} + \varepsilon, x_{\pi(k+1)}, \bar{x}_{-\pi(k),\pi(k+1)}\right) \succ P. \tag{B.53}$$

Now we have by (B.51,B.53), the induction assumption (B.49), and Lemma B.32, for all $\varepsilon \in (0, \varepsilon^k_{\pi(k)})$,

$$\left(x_{\pi(k)} + \varepsilon, x_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right) \succ P \succ \left(x_{\pi(k)} + \varepsilon, x^{j^*}_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\right)$$

and

$$\left(x_{\pi(k)} + \varepsilon, x_{\pi(k+1)}, \bar{x}_{-\pi(k),\pi(k+1)}\right) \succ P \succ \left(x_{\pi(k)} + \varepsilon, x^{j^*}_{\pi(k+1)}, \bar{x}_{-\pi(k),\pi(k+1)}\right).$$

Thus for each $\varepsilon \in (\varepsilon^k_{\pi(k)}, 0)$, by (B.48,B.52) and Axiom 2.2, there exist $\hat{x}_{\pi(k+1)}(\varepsilon), \bar{x}_{\pi(k+1)}(\varepsilon) \in X$ such that

$$\left(x_{\pi(k)} + \varepsilon, \hat{x}_{\pi(k+1)}(\varepsilon), \hat{x}_{-\pi(k),\pi(k+1)}\right) \sim P$$

and

$$\left(x_{\pi(k)} + \varepsilon, \bar{x}_{\pi(k+1)}(\varepsilon), \bar{x}_{-\pi(k),\pi(k+1)}\right) \sim P.$$

Now there exists $(\varepsilon^j)_{j=1}^{\infty} \subseteq (0, \varepsilon^k_{\pi(k)})$ such that $\varepsilon^j \to 0$ as $j \to \infty$ and $\lim_{j\to\infty} \tilde{x}_{\pi(k+1)}(\varepsilon^j)$, $\lim_{j\to\infty} \bar{x}_{\pi(k+1)}(\varepsilon^j)$ exists in $\mathbb{R} \cup \{\infty, -\infty\}$. Since by above these are bounded, the limits are between $x_{\pi(k+1)}$ and $x^{j^*}_{\pi(k+1)}$.

First, assume that $\lim_{j\to\infty} \bar{x}_{\pi(k+1)}(\varepsilon_j) < x_{\pi(k+1)}$. Now $\lim_{j\to\infty} \bar{x}_{\pi(k+1)}(\varepsilon_j) > x^{j^*}_{\pi(k+1)}$. So by (B.48,B.52), for all $j \in \mathbb{N}$,

$$\left(x_{\pi(k)} + \varepsilon_j, \bar{x}_{\pi(k+1)}(\varepsilon_j), \bar{x}_{-\pi(k),\pi(k+1)}\right), \left(x_{\pi(k)}, \lim_{j\to\infty} \bar{x}_{\pi(k+1)}(\varepsilon_j), \bar{x}_{-\pi(k),\pi(k+1)}\right) \in \mathrm{NDiag}(X^p).$$

So Axiom 2.2

$$\left(x_{\pi(k)}, \lim_{j\to\infty} \bar{x}_{\pi(k+1)}(\varepsilon_j), \bar{x}_{-\pi(k),\pi(k+1)}\right) \sim P$$

and satisfies (B.41). Symmetrically if $\lim_{j\to\infty} \hat{x}_{\pi(k+1)}(\varepsilon_j) < x_{\pi(k+1)}$, then

$$\Big(x_{\pi(k)}, \lim_{j\to\infty} \hat{x}_{\pi(k+1)}(\varepsilon_j), \hat{x}_{-\pi(k),\pi(k+1)}\Big)$$

satisfies (B.41). So assume that

$$\lim_{j\to\infty} \bar{x}_{\pi(k+1)}(\varepsilon_j) = \lim_{j\to\infty} \hat{x}_{\pi(k+1)}(\varepsilon_j) = x_{\pi(k+1)}.$$

Now for all $j \in \mathbb{N}$,

$$\Big(x_{\pi(k)} + \varepsilon, \hat{x}_{\pi(k+1)}(\varepsilon_j), \hat{x}_{-\pi(k),\pi(k+1)}\Big), \Big(x_{\pi(k)} + \varepsilon_j, \bar{x}_{\pi(k+1)}(\varepsilon_j), \bar{x}_{-\pi(k),\pi(k+1)}\Big) \in \mathrm{NDiag}(X^p).$$

Thus for all $j \in \mathbb{N}$, by Axiom 2.10,

$$\Big(x_{\pi(k)} + \varepsilon^k_{\pi(k)}, \hat{x}_{\pi(k+1)}(\varepsilon_j), \hat{x}_{-\pi(k),\pi(k+1)}\Big) \sim \Big(x_{\pi(k)} + \varepsilon^k_{\pi(k)}, \tilde{x}_{\pi(k+1)}(\varepsilon_j), \tilde{x}_{-\pi(k),\pi(k+1)}\Big).$$

Thus by Lemma B.34,

$$\Big(x_{\pi(k)} + \varepsilon^k_{\pi(k)}, x_{\pi(k+1)}, \hat{x}_{-\pi(k),\pi(k+1)}\Big) \sim \Big(x_{\pi(k)} + \varepsilon^k_{\pi(k)}, x_{\pi(k+1)}, \tilde{x}_{-\pi(k),\pi(k+1)}\Big)$$

which contradicts Lemma B.32. This shows the induction claim.

Thus by induction on A^l, either there exists $\tilde{z}^{|A^l|} \in \mathrm{NDiag}(X^p)$ that satisfies (B.41) or there exists $\varepsilon^{|A^l|} \in (0, \varepsilon^*)^{\pi(\{1,\dots,k\})}$ such that for all $i, j \in \{1, \dots, |A^l|\}$, $i < j$, $\varepsilon_{\pi(i)} > \varepsilon_{\pi(j)}$ and this satisfies (B.49). We show that there exists $\tilde{z}^{|A^l|} \in \mathrm{NDiag}(X^p)$ that satisfies (B.41). Assume, per contra, that there exists $\varepsilon^{|A^l|} \in (0, \varepsilon^*)^{A^l}$ that satisfies (B.49). Then this contradicts (B.47). Thus there exists $\tilde{z}^{|A^l|} \in \mathrm{NDiag}(X^p)$ that satisfies (B.41). Thus we define $\tilde{z}^{l+1} = \tilde{z}^{|A^l|}$ that shows the induction claim for the induction on A.

Thus by induction on A, there exists $\tilde{z}^{|A|} \in \mathrm{NDiag}(X^p)$ that satisfies either (B.41) or (B.45,B.46). We show that $\tilde{z}^{|A|}$ satisfies (B.41). Assume, per contra, $\tilde{z}^{|A|}$ satisfies (B.45,B.46). Since $A \neq \varnothing$ by assumption, we have for all $j \in \operatorname{supp} p$, $\tilde{z}^{|A|}_j \geq x_j$ and $P \succ (p, \tilde{z}^{|A|})$ that contradicts the assumption. Thus $\tilde{z}^{|A|} \in \mathrm{NDiag}(X^p)$ satisfies (B.41). This shows the claim.

□

With the above lemma, we can establish a stronger form of continuity that allows us to compare alternatives across diagonal. This states that if we have two non-diagonal sequences converging to the same limit prizes, then they are eventually close to each other. Here, the limit prizes might be on the diagonal.

Lemma B.36 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Then for all lotteries $P \succ Q$ and co-probabilities $p \in \Delta^3(\mathbb{N})$, if $(x^i)_{i=1}^\infty, (y^i)_{i=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ are sequences such that $\lim_{i\to\infty} x^i = \lim_{i\to\infty} y^i$ and if for all $i \in \mathbb{N}$,

$$\sum_{l\in\operatorname{supp} p} p_l x_l^i \succsim P, \text{ then there exists } i^0 \in \mathbb{N} \text{ such that for all } i > i^0, \sum_{l\in\operatorname{supp} p} p_l y_l^i \succsim Q,$$

and if for all $i \in \mathbb{N}$,

$$Q \succsim \sum_{l\in\operatorname{supp} p} p_l x_l^i, \text{ then there exists } i^0 \in \mathbb{N} \text{ such that for all } i > i^0, P \succsim \sum_{l\in\operatorname{supp} p} p_l y_l^i.$$

Proof. We will show the claim for lower contour sets. The upper contour sets follow symmetrically. Let $p \in \Delta(\mathbb{N})$.

Let $x \in X^{\operatorname{supp} p}$. Denote

$$A(x) = \{x_i | i \in \operatorname{supp} p\}$$

and denote the number of joined indices by

$$\theta(x) = \sum_{y\in A(x)} \left(|\{i \in \operatorname{supp} p | x_i = y\}| - 1\right).$$

We will show by induction that for all $0 \le n \le |\operatorname{supp} p|$, the following holds: If $x \in X^{\operatorname{supp} p}, P \succ Q, (x^i)_{i=1}^\infty, (y^i)_{i=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ are such that $\theta(x) = n$, $\lim_{i\to\infty} x^i = \lim_{i\to\infty} y^i = x$, and for all $i \in \mathbb{N}$, $Q \succsim \sum_{l\in\operatorname{supp} p} p_l x_l^i$, then there exists $i^0 \in \mathbb{N}$ such that for all $i > i^0$, $P \succsim \sum_{l\in\operatorname{supp} p} p_l y_l^i$.

We will start with the case $n = 0$. By assumption for all $i \in \mathbb{N}$, $Q \succsim (p, x^i)$. Since $\theta(x) = 0$, $x \in \mathrm{NDiag}(X^p)$, so we have by Axiom 2.2, $P \succ Q \succsim (p, x)$. Since X is open, by Axiom 2.2, there exists $\varepsilon > 0$ such that $P \succ (p, x + \bar{1}\varepsilon)$. Since $y^i \to x$ as $i \to \infty$, there exists $i^0 \in \mathbb{N}$ such that for all $i > i^0$ and $l \in \operatorname{supp} p$, $|y_l^i - x_l| < \varepsilon$. So by Lemma B.32, $P \succ (p, x + \bar{1}\varepsilon) \succsim (p, y^i)$.

Assume that the claim holds for all $0 \le n \le |\operatorname{supp} p| - 1$. We will show that it holds for $n + 1$. Let $x \in X^{\operatorname{supp} p}$ be such that $\theta(x) = n + 1$ and $P \succ Q, (x^i)_{i=1}^\infty, (y^i)_{i=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ are such that $\lim_{i\to\infty} x^i = \lim_{i\to\infty} y^i = x$ and for all $i \in \mathbb{N}$, $Q \succsim \sum_{l\in\operatorname{supp} p} p_l x_l^i$. Assume, per contra, for all $i \in \mathbb{N}$, $(p, y^i) \succ P$.

We first replace $(x^i)_{i=1}^\infty$, with a strictly increasing sequence that converges to x. First, by taking a subsequence for $i \in \operatorname{supp} p$, there exists a subsequence $(x^{i_j})_{j=1}^\infty$ such that each coordinate sequence for $l \in \operatorname{supp} p$, $(x_l^{i_j})_{j=1}^\infty$ is monotone. Second, for each $l \in \operatorname{supp} p$, if the set $\{j \in \mathbb{N} | x_l^{i_j} \ne x_l^{i_{j+1}}\}$ is finite, then the set $\{j \in \mathbb{N} | x_l^{i_j} = x_l\}$ is infinite by taking the last change of index. Thus there exists

a subsequence $(x^{i_{j_k}})_{k=1}^{\infty}$ such that each coordinate sequence for $l \in \operatorname{supp} p$, $(x_l^{i_j})_{j=1}^{\infty}$ is either strictly monotone or constant.

Denote the collection of coordinates $l \in \operatorname{supp} p$ for which $(x_l^{i_{j_k}})_{k=1}^{\infty}$ is decreasing or constant by

$$D = \{l \in \operatorname{supp} p | \forall\, k \in \mathbb{N}, x_l^{i_{j_k}} \geq x_l\}.$$

Since X is open, for each $l \in D$, there exists a strictly increasing sequence $(y_l^k)_{k=1}^{\infty}$ such that $y_l^k \to x_l$ as $k \to \infty$ and $(y_D^k, x_{-D}^{i_{j_k}}) \in \operatorname{NDiag}(X^p)$. Define $(\tilde{x}^k)_{k=1}^{\infty} = (y_D^k, x_{-D}^{i_{j_k}})_{k=1}^{\infty}$. $(\tilde{x}^k)_{k=1}^{\infty}$ is strictly increasing and each coordinate is a subsequence of the original sequence or converging to x and so we have $\tilde{x}^k \to x$ as $k \to \infty$. By the definition of D, we have for each $l \in \operatorname{supp} p$ and $k \in \mathbb{N}$, $x_l^{i_{j_k}} \geq \tilde{x}_l^k$. Thus by Lemma B.32 for each $k \in \mathbb{N}$, $P \succsim (p, x^{i_{j_k}}) \succsim (p, \tilde{x}^k)$.

Next, we replace P, with a tight upper limit. Since X is open, there exists $\varepsilon^x > 0$ such that $x + \bar{1}\varepsilon^x, x - \bar{1}\varepsilon^x \in X^{\operatorname{supp} p}$. There exists $i^0 \in \mathbb{N}$ such that for all $i \geq i^0$, $\left\|\tilde{x}^{i^0} - x\right\| < \varepsilon^x$. Since $\tilde{x}^k$ is strictly increasing sequence, $\tilde{x}^{i^0} + \bar{1}\varepsilon^x \in \operatorname{NDiag}(X^p)$. Now we have for all $i \geq i^0, l \in \operatorname{supp} p$,

$$\tilde{x}_l^{i^0} + \varepsilon^x \geq -|\tilde{x}_l^{i^0} - \tilde{x}_l^i| + \tilde{x}_l^i + \varepsilon^x > \tilde{x}_l^i.$$

Thus by Lemma B.32, for all $i \in \mathbb{N}$,

$$(p, \tilde{x}^{i^0} + \bar{1}\varepsilon^x) \succsim (p, \tilde{x}^i).$$

Now by Axiom 2.2, for each $i > i^0$, the set

$$\{\varepsilon \in [0, \varepsilon^x] | (p, \tilde{x}^{i^0} + \bar{1}\varepsilon) \succsim (p, \tilde{x}^i)\}$$

is closed. Thus the value

$$\varepsilon^* = \inf \bigcap_{i=i^0}^{\infty} \{\varepsilon \in [0, \varepsilon^x] | (p, \tilde{x}^{i^0} + \bar{1}\varepsilon) \succsim (p, \tilde{x}^i)\}$$

exists as an infimum over a compact set since intersections maintain closedness of the set. Additionally, $\varepsilon^* > 0$. Denote $\tilde{P} = (p, \tilde{x}^{i^0} + \bar{1}\varepsilon^*)$. Next we show that if $R \in \Delta(X)$ is such that for all $i > i^0$, $R \succsim (p, \tilde{x}^i)$, then

$$R \succsim \tilde{P}. \tag{B.54}$$

Now for each $\varepsilon \in [0, \varepsilon^*)$, there exists $i^\varepsilon \in \mathbb{N}$ such that

$$R \succsim (p, \tilde{x}^{i^\varepsilon}) \succ (p, \tilde{x}^{i^0} + \bar{1}\varepsilon).$$

Thus by considering an increasing sequence $\varepsilon \to \varepsilon^*$, we have

$$R \succsim (p, \tilde{x}^{i^0} + \bar{1}\varepsilon^*) = \tilde{P}.$$

Additionally, as a contraposition of (B.54), we have that if $R \in \Delta(X)$ is such that $\tilde{P} \succ R$, then there exists $i^1 \in \mathbb{N}$ such that $(p, \tilde{x}^{i^1}) \succ R$ and so since $\tilde{x}^i$ is strictly increasing for all $i > i^1$,

$$(p, \tilde{x}^i) \succ R. \tag{B.55}$$

First, by Equation (B.54),

$$Q \succ P \succsim \tilde{P}. \tag{B.56}$$

Now if $z \in \mathrm{NDiag}(X^p)$ is such that for all $l \in \operatorname{supp} p$, $z_l \geq x_l$, then by Lemma B.32, for all $i \in \mathbb{N}$,

$$(p, z) \succsim (p, \tilde{x}^i)$$

and so by (B.54)

$$(p, z) \succsim \tilde{P}.$$

Thus by Lemma B.35 and since $\theta(x) > 0$, there exist $z^* \in \mathrm{NDiag}(X^p), i^*, j^* \in \operatorname{supp} p, i^* \neq j^*$ such that

$$x_{i^*} = x_{j^*} = z^*_{i^*} \text{ and } (p, z^*) \sim \tilde{P}. \tag{B.57}$$

Since $|\operatorname{supp} p| \geq 2$ and X is open by Axiom 2.2 and Lemma B.32, there exists $z^\dagger \in \mathrm{NDiag}(X^p)$ such that $z^\dagger_{i^*} = x_{i^*}$ and $Q \succ (p, z^\dagger) \succ \tilde{P}$. By Axiom 2.2, there exists $\varepsilon^\dagger > 0$ such that

$$\varepsilon^\dagger < \frac{1}{2} \min_{j,i \in \operatorname{supp} p, i \neq j} \left| z^\dagger_i - z^\dagger_j \right|, \varepsilon^\dagger < \frac{1}{2} \min\{x_i - x_j | i, j \in \operatorname{supp} p, x_i \neq x_j\}, z^\dagger_{i^*} + \varepsilon^\dagger, z^\dagger_{i^*} - \varepsilon^\dagger \in X \tag{B.58}$$

and

$$Q \succ (p, (z^\dagger_{i^*} + \varepsilon^z, z^\dagger_{-i^*})), (p, (z^\dagger_{i^*} - \varepsilon^z, z^\dagger_{-i^*})) \succ \tilde{P}. \tag{B.59}$$

Since X is open and $x_{i^*} + \varepsilon^\dagger \in X$, there exists $i^\dagger > i^0$ such that for all $i > i^\dagger, l \in \operatorname{supp} p$, $|\tilde{x}^i_l - x_l|, |y^i_l - x_l| < \frac{1}{2}\varepsilon^\dagger$ and $x^i_{i^*} + \varepsilon^\dagger, y^i_{i^*} + \varepsilon^\dagger \in X$. So by (B.58), for all $i > i^\dagger$,

$$(x^i_{i^*} + \varepsilon^\dagger, x^i_{-i^*}), (y^i_{i^*} + \varepsilon^\dagger, y^i_{-i^*}) \in \mathrm{NDiag}(X^p).$$

By (B.59) and Lemma B.32, we have for all $i > i^\dagger$,

$$(p, (z^\dagger_{i^*} - \varepsilon^z, z^\dagger_{-i^*})) \succ \tilde{P} \succsim (p, x^i) \succsim (p, (x_{i^*} - \varepsilon^\dagger, x^i_{-i^*})).$$

Thus by Axiom 2.10, for all $i > i^\dagger$,

$$Q \succ (p, (z^\dagger_{i^*} + \varepsilon^z, z^\dagger_{-i^*})) \succ (p, (x_{i^*} + \varepsilon^\dagger, x^i_{-i^*})).$$

Now $(x_{i^*} + \varepsilon^\dagger, x^i_{-i^*})_{i=1}^\infty, (x_{i^*} + \varepsilon^\dagger, y^i_{-i^*})_{i=i^\dagger+1}^\infty \subseteq \mathrm{NDiag}(X^p)$ and

$$\lim_{i \to \infty} (x_{i^*} + \varepsilon^\dagger, x^i_{-i^*}) = \lim_{i \to \infty} (x_{i^*} + \varepsilon^\dagger, y^i_{-i^*}) = (x_{i^*} + \varepsilon^\dagger, x_{-i^*})$$

where by (B.57,B.58),

$$\theta(x_{i^*} + \varepsilon^\dagger, x_{-i^*}) = \theta(x) - 1 = l.$$

Thus by the induction assumption, there exists $i^y > i^\dagger$ such that for all $i > i^y$,

$$Q \succsim (p, (x_{i^*} + \varepsilon^\dagger, y^i_{-i^*})).$$

So by Lemma B.32, for all $i > i^y$,

$$Q \succsim (p, y^i)$$

that shows the induction claim. This proves the claim by induction. □

Continuous and Separable Extension Across Diagonal

Here, we establish that the preferences conditional on a co-probability can be extended continuously and separably to the diagonal of prizes. The main part here is in establishing transitivity for the extended preferences. First, we establish monotonicity for these conditional preferences.

Lemma B.37 Let $\succsim$ satisfy Axioms 2.1 and 2.10. For all $x, y \in X^{\operatorname{supp} p}$, if for all $i \in \operatorname{supp} p$, $x_i > y_i$, then $x \succ^p y$.

Proof. Let $(x^j)_{j=1}^\infty, (y^j)_{j=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ be such that $\lim_{j\to\infty} x^j = x$ and $\lim_{j\to\infty} y^j = y$. Now there exists j^0 such that for all $j \geq j^0$,

$$\left\|y^j - y\right\|, \left\|x^j - x\right\| < \frac{1}{2} \min_{i \in \operatorname{supp} p} |x_i - y_i|.$$

Thus for all $j \geq j^0$ and for all $i \in \operatorname{supp} p$, we have

$$\begin{aligned} x^j_i &= x^j_i - x_i + x_i - y_i + y_i - y^j_i + y^j_i \geq -\left|x^j_i - x_i\right| + \min_{k \in \operatorname{supp} p} |x_k - y_k| - \left|y_i - y^j_i\right| + y^j_i \\ &\geq -\left\|x^j - x\right\| - \left\|y^j - y\right\| + \min_{k \in \operatorname{supp} p} |x_k - y_k| + y^j_i > y^j_i. \end{aligned}$$

Thus for all $j \geq j^0$, by Lemma B.32,

$$x^j_i \mathrel{\tilde{\succ}}^p y^j_i.$$

Thus for all $j \geq j^0$, $(y^j, x^j) \notin \succsim^p$. Since $(x^j)_{j=1}^\infty, (y^j)_{j=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ were arbitrary sequences such that $\lim_{j\to\infty} x^j = x$ and $\lim_{j\to\infty} y^j = y$, we have that

$$(y, x) \notin \mathrm{Cl}_{X^{\operatorname{supp} p} \times X^{\operatorname{supp} p}}(\tilde{\succsim}^p) = \succsim^p .$$

But by above, $x \succsim^p y$ and so $x \succ^p y$. □

The next result established that the conditional preferences are continuous when focusing only on the non-diagonal.

Lemma B.38 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10 and $p = (p_1, \dots, p_n) \in \Delta^3(\mathbb{N})$ be a co-probability. For all lotteries P, the sets

$$\{(x_i)_{i=1}^n | \sum_{i=1}^n p_i x_i \succsim P, \forall\, i \neq j, x_i \neq x_j\} \text{ and } \{(x_i)_{i=1}^n | P \succsim \sum_{i=1}^n p_i x_i, \forall\, i \neq j, x_i \neq x_j\}$$

are closed in the Euclidean relative topology in the set $\{(x_i)_{i=1}^n | \forall\, i \neq j, x_i \neq x_j\}$.[1]

Proof. Let P be a lottery. We show that the set

$$B \coloneqq \{(x_i)_{i=1}^n | \sum_{i=1}^n p_i x_i \succsim P, \forall\, i \neq j, x_i \neq x_j\}$$

is closed in $\mathrm{NDiag}(X^p) \coloneqq \{(x_i)_{i=1}^n | \forall\, i \neq j, x_i \neq x_j\}$. The other set follows symmetrically. We will do it by showing that $\mathrm{Cl}_{X^{\operatorname{supp} p}} B \cap \mathrm{NDiag}(X^p) = B$. Since $X^{\operatorname{supp} p}$ is a metric space, this is equivalent to showing that for all convergent sequences $(x^j)_{j=1}^\infty \subseteq B$ such that $\lim_{j\to\infty} x^j \in \mathrm{NDiag}(X^p)$, we have $\lim_{j\to\infty} x^j \in B$.

Let $(x^j)_{j=1}^\infty \subseteq B$ and $x \in \mathrm{NDiag}(X^p)$ be such that $\lim_{j\to\infty} x^j = x$. We show that $x \in B$. Assume, per contra, $P \succ x$. We consider sequences $(x^{j,*})_{j=1}^\infty, (x_*^j)_{j=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ such that for all $j \in \mathbb{N}$ and $i \in \operatorname{supp} p$,

$$x_i^j + \frac{1}{j} \geq x_i^{j,*} > x_i^j \text{ and } x_i > x_{i,*}^j \geq x_i^j - \frac{1}{j}$$

that exists since X is open. Thus by Lemma B.37, for all $j \in \mathbb{N}$,

$$x^{j,*} \succ^p x^j \succ^p P \succ^p x \succ^p x_*^j.$$

Additionally for all $i \in \operatorname{supp} p$,

$$x_i = \lim_{j\to\infty} x_i^j + \frac{1}{j} \geq \lim_{j\to\infty} x_i^{j,*} \lim_{j\to\infty} x_i^j = x_i$$

and

$$x_i = \lim_{j\to\infty} x_i^j \geq \lim_{j\to\infty} x_{i,*}^j \lim_{j\to\infty} x_i^j - \frac{1}{j} = x_i.$$

So $\lim_{j\to\infty} x^{j,*} = \lim_{j\to\infty} x_*^j = x$. This contradicts Lemma B.36. Thus $x \in B$ which shows the claim since $(x^j)_{j=1}^\infty \subseteq B$ was an arbitrary convergent sequence in the set $\mathrm{NDiag}(X^p)$. □

The next lemma shows that $\succsim^p$ is complete.

Lemma B.39 Let $\succsim$ satisfy Axiom 2.1 and $p \in \Delta^3(\mathbb{N})$. For all $x, y \in X^{\operatorname{supp} p}$, $x \succsim^p y$ or $y \succsim^p x$.

[1] This relative topology is such that a subset E is closed if there exists a closed subset F of X^n (i.e. of $\mathbb{R}^n$) such that $E = F \cap \{(x_i)_{i=1}^n | \forall\, i \neq j, x_i \neq x_j\}$.

Proof. Now there exist $(x_i)_{i=1}^{\infty}, (y_i)_{i=1}^{\infty} \subseteq \text{NDiag}(X^p)$ such that $x_i \to x$ and $y_i \to y$ as $i \to \infty$. Then there exist subsequences $(x_{i_j})_{j=1}^{\infty}, (y_{i_j})_{j=1}^{\infty}$ such that for all j, $x_{i_j} \widetilde{\succsim}^{py}_{i_j}$ and so $x \succsim^p y$ or there exist subsequences $(x_{i_k})_{k=1}^{\infty}, (y_{i_k})_{k=1}^{\infty}$ such that for all k, $y_{i_k} \widetilde{\succsim}^{px}_{i_k}$ and so $y \succsim^p x$. □

The next lemma shows that $\succsim^p$ is continuous.

Lemma B.40 Let $\succsim$ be an order and $p \in \Delta^3(\mathbb{N})$. For all $x \in X^{\operatorname{supp} p}$, $\{z \in X^{\operatorname{supp} p} | x \succsim^p z\}$ and $\{z \in X^{\operatorname{supp} p} | z \succsim^p x\}$ are closed

Proof. Let $z \in X^{\operatorname{supp} p}$. We show that

$$B \coloneqq \{x \in X^{\operatorname{supp} p} | x \succsim^p z\}$$

is closed. The other set follows symmetrically. Let $(x^i)_{i=1}^{\infty} \subseteq B$ and x be such that $x^i \to x$ as $i \to \infty$. Now by the definition of $\succsim^p$, for all $i \in \mathbb{N}$, there exist $(x^{i,j})_{j=1}^{\infty}, (z^{i,j})_{j=1}^{\infty} \subseteq \text{NDiag}(X^p)$ such that for all $j \in \mathbb{N}$, $x^{i,j} \widetilde{\succsim}^p z^{i,j}$ and $x^{i,j} \to x^i$ and $z^{i,j} \to z$ as $j \to \infty$. For all $i \in \mathbb{N}$, choose $j_i \in \mathbb{N}$ such that

$$\left\| x^{i,j_i} - x^i \right\| < \frac{1}{i} \text{ and } \left\| z^{i,j_i} - z \right\| < \frac{1}{i}. \tag{B.60}$$

Now consider the sequences $(x^{i,j_i})_{i=1}^{\infty}$ and $(z^{i,j_i})_{i=1}^{\infty}$. We have by (B.60) $z^{i,j_i} \to z$ as $i \to \infty$ and

$$\left\| x^{i,j_i} - x \right\| = \left\| x^{i,j_i} - x^i + x^i - x \right\| \leq \left\| x^{i,j_i} - x^i \right\| + \left\| x^i - x \right\| < \frac{1}{i} + \left\| x^i - x \right\|.$$

Thus since $x^i \to x$ as $i \to \infty$, we have $x^{i,j_i} \to x$ as $i \to \infty$. Additionally, we have for all i, $x^{i,j_i} \widetilde{\succsim}^p z^{i,j_i}$. Thus by the definition of $\succsim^p$, we have

$$\lim_{i \to \infty} x^{i,j_i} \succsim^p \lim_{i \to \infty} z^{i,j_i} \Rightarrow x \succsim^p z.$$

And so $x \in B$. □

The next lemma establishes monotonicity for $\succsim^p$.

Lemma B.41 Let $\succsim$ satisfy Axioms 2.1 and 2.10 and $p \in \Delta^3(\mathbb{N})$. For all $x, y \in X^{\operatorname{supp} p}$, if for all $i \in \operatorname{supp} p$, $x_i \geq y_i$, then $x \succsim^p y$.

Proof. Let $(x^j)_{j=1}^{\infty} \subseteq X^{\operatorname{supp} p}$ be such that for all $j \in \mathbb{N}$ and $i \in \operatorname{supp} p$,

$$x_i^j + \frac{1}{j} \geq x_i^{j,*} > x_i^j$$

that exists since X is open. Now for all $i \in \operatorname{supp} p$, $x_i^j = x_i + \frac{1}{j} > x_i \geq y_i$. So by Lemma B.37, for all j, $x^j \succ^p y$. Thus since $\lim_{j \to \infty} x^j = x$, by Lemma B.40, $x \succsim^p y$. □

The next lemma shows that off-diagonal the conditional preferences and the conditional extended preferences are the same.

Lemma B.42 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10 and $p \in \Delta^3(\mathbb{N})$. For all $x, y \in \mathrm{NDiag}(X^p)$,

$$x \mathrel{\widetilde{\succ}^p} y \iff x \succ^p y$$

Proof. First, we show that $x \mathrel{\widetilde{\succ}^p} y \Longrightarrow x \succ^p y$. By applying Lemma B.38 twice and since X is open, there exist $\tilde{x}, \tilde{y} \in \mathrm{NDiag}(X^p)$ such that $\tilde{x} \mathrel{\widetilde{\succ}^p} \tilde{y}$ and for all $i \in \operatorname{supp} p$, $x_i > \tilde{x}_i$ and $\tilde{y}_i > y_i$. Now let $(x^j)_{j=1}^\infty, (y^j)_{j=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ be such that $\lim_{j\to\infty} x^j = x$ and $\lim_{j\to\infty} y^j = y$. Now there exists j^0 such that for all $j \geq j^0$,

$$\left\|x^j - x\right\| < \min_{i \in \operatorname{supp} p} |x_i - \tilde{x}_i| \text{ and } \left\|y^j - y\right\| < \min_{i \in \operatorname{supp} p} |y_i - \tilde{y}_i|.$$

Thus for all $j \geq j^0$ and for all $i \in \operatorname{supp} p$, we have

$$\begin{aligned} x_i^j &= x_i^j - x_i + x_i - \tilde{x}_i + \tilde{x}_i \geq -\left|x_i^j - x_i\right| + \min_{k \in \operatorname{supp} p} |x_k - \tilde{x}_k| + \tilde{x}_i \\ &\geq -\left\|x^j - x\right\| + \min_{k \in \operatorname{supp} p} |x_k - \tilde{x}_k| + \tilde{x}_i > \tilde{x}_i \end{aligned}$$

and

$$\begin{aligned} y_i^j &= y_i^j - y_i + y_i - \tilde{y}_i + \tilde{y}_i \leq \left|y_i^j - y_i\right| - |y_i - \tilde{y}_i| + \tilde{y}_i \\ &\leq \left\|y^j - y\right\| - \min_{k \in \operatorname{supp} p} |y_k - \tilde{y}_k| + \tilde{y}_i < \tilde{y}_i. \end{aligned}$$

Thus for all $j \geq j^0$, by Lemma B.32,

$$x^j \mathrel{\widetilde{\succ}^p} \tilde{x} \mathrel{\widetilde{\succ}^p} \tilde{y} \mathrel{\widetilde{\succ}^p} y^j.$$

Thus for all $j \geq j^0$, $(y^j, x^j) \notin \mathrel{\widetilde{\succsim}^p}$. Since $(x^j)_{j=1}^\infty, (y^j)_{j=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ were arbitrary sequences such that $\lim_{j\to\infty} x^j = x$ and $\lim_{j\to\infty} y^j = y$, we have that

$$(y, x) \notin \mathrm{Cl}_{X^{\operatorname{supp} p} \times X^{\operatorname{supp} p}}(\mathrel{\widetilde{\succsim}^p}) = \succsim^p .$$

But $x \succsim^p y$ and so $x \succ^p y$.

Next, we show that $x \succ^p y \Rightarrow x \mathrel{\widetilde{\succ}^p} y$.

$$\mathrm{Cl}_{X^{\operatorname{supp} p} \times X^{\operatorname{supp} p}}(\mathrel{\widetilde{\succsim}^p}) \supseteq \mathrel{\widetilde{\succsim}^p}$$

and so by Lemma B.39,

$$\succ^p \cap \mathrm{NDiag}(X^p) = \complement\, \mathrm{Cl}_{X^{\operatorname{supp} p} \times X^{\operatorname{supp} p}}(\mathrel{\widetilde{\succsim}^p}) \cap \mathrm{NDiag}(X^p) \subseteq \complement\, \mathrel{\widetilde{\succsim}^p} \cap \mathrm{NDiag}(X^p) = \succ^p .$$

□

The next lemma shows that $\succsim^p$ satisfies transitivity for the strict part.

Lemma B.43 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10 and $p \in \Delta^3(\mathbb{N})$. For all $x, y, z \in X^{\operatorname{supp} p}$, if $x \succ^p y \succsim^p z$, then $x \succ^p z$.

Proof. Assume, per contra, $z \succsim^p x$. Now there exist sequences $(x^j)_{j=1}^\infty, (z^j)_{j=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ such that $\lim_{j\to\infty} x^j = x$ and $\lim_{j\to\infty} z^j = z$ and for all $j \in \mathbb{N}$, $z^j \succsim^p x^j$. By applying Lemma B.40 twice, there exist $\tilde{x}, \tilde{y} \in \mathrm{NDiag}(X^p)$ such that for all $i \in \operatorname{supp} p$, $x_i > \tilde{x}_i$, $\tilde{y}_i > y_i$ and $x \succ^p \tilde{x} \succ^p \tilde{y} \succ^p y$. Now let $(x^j)_{j=1}^\infty, (z^{j,*})_{j=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ be such that $\lim_{j\to\infty} x^j = x$ and $\lim_{j\to\infty} z^{j,*} = z$ and for all j, $z^{j,*} \mathrel{\tilde{\succsim}}^p x^j$. Now there exists j^0 such that for all $j \geq j^0$,

$$\left\| x^j - x \right\| < \min_{i \in \operatorname{supp} p} |x_i - \tilde{x}_i|.$$

Thus for all $j \geq j^0$ and for all $i \in \operatorname{supp} p$, we have

$$\begin{aligned} x_i^j &= x_i^j - x_i + x_i - \tilde{x}_i + \tilde{x}_i \geq -\left| x_i^j - x_i \right| + \min_{k \in \operatorname{supp} p} |x_k - \tilde{x}_k| + \tilde{x}_i \\ &\geq -\left\| x^j - x \right\| + \min_{k \in \operatorname{supp} p} |x_k - \tilde{x}_k| + \tilde{x}_i > \tilde{x}_i. \end{aligned}$$

Thus, by Lemmas B.32 and B.42, for all $j \geq j^0$,

$$z^{j,*} \mathrel{\tilde{\succsim}}^p x^j \mathrel{\tilde{\succ}}^p \tilde{x} \mathrel{\tilde{\succ}}^p \tilde{y}. \tag{B.61}$$

Let $(y^j)_{j=1}^\infty, (z_*^j)_{j=1}^\infty \subseteq \mathrm{NDiag}(X^p)$ be such that $\lim_{j\to\infty} y^j = y$ and $\lim_{j\to\infty} z_*^j = z$ and for all $j \in \mathbb{N}$, $y^j \mathrel{\tilde{\succsim}}^p z_*^j$. Now there exists $j^1 > j^0$ such that for all $j \geq j^1$,

$$\left\| y^j - y \right\| < \min_{i \in \operatorname{supp} p} |y_i - \tilde{y}_i|.$$

Thus for all $j \geq j^1$ and for all $i \in \operatorname{supp} p$, we have

$$\begin{aligned} y_i^j &= y_i^j - y_i + y_i - \tilde{y}_i + \tilde{y}_i \leq \left| y_i^j - y_i \right| - |y_i - \tilde{y}_i| + \tilde{y}_i \\ &\leq \left\| y^j - y \right\| - \min_{k \in \operatorname{supp} p} |y_k - \tilde{y}_k| + \tilde{y}_i < \tilde{y}_i. \end{aligned}$$

Thus by Lemmas B.32 and B.42 and (B.61), for all $j \geq j^1$,

$$z^{j,*} \mathrel{\tilde{\succsim}}^p x^j \mathrel{\tilde{\succ}}^p \tilde{x} \mathrel{\tilde{\succ}}^p \tilde{y} \mathrel{\tilde{\succ}}^p y^j \mathrel{\tilde{\succsim}}^p z_*^j.$$

So by Axiom 2.1, for all $j \geq j^1$,

$$z^{j,*} \succ^p \tilde{x} \mathrel{\tilde{\succ}}^p \tilde{y} \mathrel{\tilde{\succ}}^p z_*^j$$

and $\lim_{j\to\infty} z^{j,*} = \lim_{j\to\infty} z_*^j = z$. But this contradicts Lemma B.36. Thus $x \succ^p z$. □

The next lemma shows full transitivity for $\succsim^p$.

Lemma B.44 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10 and $p \in \Delta^3(\mathbb{N})$. For all $x, y, z \in X^{\operatorname{supp} p}$, if $x \succsim^p y \succsim^p z$, then $x \succsim^p z$.

Proof. Let $x, y, z \in X^{\operatorname{supp} p}$ be such that $x \succsim^p y$ and $y \succsim^p z$. Assume, per contra, that $z \succ^p x$. Now

$$z \succ^p x \succsim^p y \succsim^p z.$$

By applying Lemma B.43 twice, we have

$$z \succ^p z$$

which contradicts Lemma B.39. □

The next lemma shows that $\succsim^p$ satisfies Restricted Common Consequence Independence that allows us to connect preferences conditional on different co-probabilities to each other and cancel common prizes with common probabilities.

Lemma B.45 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10 and $p \in \Delta^3(\mathbb{N})$. Let $i^* \notin \operatorname{supp} p$. Then for all $x, y \in X^{\operatorname{supp} p}$,$z_{i^*} \in X$, and $\alpha \in (0,1)$

$$x \succsim^p y \iff (z_{i^*}, x_{-i^*}) \succsim^{\alpha p + (1-\alpha)\delta_{i^*}} (z_{i^*}, y_{-i^*}).$$

Proof. We show the equivalent claim, by Lemma B.39, of

$$y \succ^p x \iff (z_{i^*}, y) \succ^{\alpha p + (1-\alpha)\delta_{i^*}} (z_{i^*}, x).$$

Denote $p\alpha\delta \coloneqq \alpha p + (1-\alpha)\delta_{i^*}$. We first show that $(z_{i^*}, y_{-i^*}) \succ^{p\alpha\delta} (z_{i^*}, x_{-i^*}) \Rightarrow y \succ^p x$.

By applying Lemma B.40 twice, there exist $x^*, y_* \in \operatorname{NDiag}(X^{p\alpha\delta})$ such that for all $i^* \in \operatorname{supp} p\alpha\delta$, $x^*_{i^*} > (z_{i^*}, x_{-i^*})_{i^*}$, and $(z_{i^*}, y_{-i^*})_{i^*} > y_{*,i^*}$ and $(z_{i^*}, y_{-i^*}) \succ^{p\alpha\delta} x^*$ and $y_* \succ^{p\alpha\delta} x^*$. Now by Lemma B.41,

$$(z_{i^*}, y_{*,-i^*}) \succsim^{p\alpha\delta} y_* \succ^{p\alpha\delta} x^* \succsim^{p\alpha\delta} (z_{i^*}, x^*_{-i^*}).$$

Thus by Lemmas B.42 and B.44 and Axiom 2.10 since $(z_{i^*}, y_{*,-i^*}), (z_{i^*}, x^*_{-i^*}) \in \operatorname{NDiag}(X^{p\alpha\delta})$,

$$y_{*,-i^*} \succ^p x^*_{-i^*}.$$

Now by Lemma B.41,

$$y \succsim^p y_{*,-i^*} \succ^p x^*_{-i^*} \succsim^p x.$$

Finally, by Lemma B.44, $y \succ^p x$.

Next, we show that $y \succ^p x \Rightarrow (z_{i^*}, y_{-i^*}) \succ^{p\alpha\delta} (z_{i^*}, x_{-i^*})$. By applying Lemma B.40 twice, there exist $x^*, y_* \in \operatorname{NDiag}(X^p)$ such that for all $i^* \in \operatorname{supp} p$, $x^*_{i^*} \neq z_{i^*} \neq y_{*,i^*}$, $x^*_{i^*} > x_{i^*}$, and $y_{i^*} > y_{*,i^*}$ and $y \succ^p x^*$ and $y_* \succ^p x^*$. Now by Lemmas B.42 and B.44 and Axiom 2.10, since $y_*, x^* \in \operatorname{NDiag}(X^p)$, and for all $j \in \operatorname{supp} p$, $x^*_j \neq z_{i^*} \neq y_{*,j}$, we have

$$(z_{i^*}, y_{*,-i^*}) \succ^{p\alpha\delta} (z_{i^*}, x^*_{-i^*}).$$

Now by Lemma B.41,

$$(z_{i^*}, y) \succsim^{p\alpha\delta} (z_{i^*}, y_{*,-i^*}) \succ^{p\alpha\delta} (z_{i^*}, x^*_{-i^*}) \succsim^{p\alpha\delta} (z_{i^*}, x).$$

Finally, by Lemma B.44, $(z, y_{-i^*}) \succ^p (z, x_{-i^*})$. □

As a corollary of the previous lemma, we get that $\succsim^p$ is separable.

Lemma B.46 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10 and $p \in \Delta^3(\mathbb{N})$. For all $x, y \in X^{\operatorname{supp} p}$, $z, z' \in X$, $i \in \operatorname{supp} p$,

$$(z, x_{-i}) \succsim^p (z, y_{-i}) \Rightarrow (z', x_{-i}) \succsim^p (z', y_{-i}).$$

Proof. Follows from applying Lemma B.45 twice. □

By above, $\succsim^p$ is separable and continuous and hence it has an additive representation.

Lemma B.47 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10 and $p \in \Delta(\mathbb{N})$. Then there exists $(u_i)_{i \in \operatorname{supp} p}$ such that for each $i \in \operatorname{supp} p$, $u_i : X \to \mathbb{R}$ is continuous, strictly increasing, and for all $x, y \in X^{\operatorname{supp} p}$,

$$x \succsim^p y \iff \sum_{i \in \operatorname{supp} p} u_i(x_i) \geq u_i(y_i).$$

Proof. If $\operatorname{supp} p \geq 3$, follows from Lemmas B.32, B.39, B.40, B.44, B.46, and Wakker (1988).

Next assume that $|\operatorname{supp} p| \in 1, 2$. Now denote $n = |\operatorname{supp} p|$, $\tilde{p} = 1/2 p + 1/4 \delta_{n+1} + 1/4 \delta_{n+2}$. Denote $i_1 = n + 1$ and $i_2 = n + 2$. By above, there exists $(u_i)_{i \in \operatorname{supp} \tilde{p}}$ that is continuous, strictly increasing, and additive representation for $\succsim^{\tilde{p}}$. Now for all $x, y \in X^{\operatorname{supp} p}$, let $z \in X$. By Lemma B.45,

$$\begin{aligned}
x \succsim^p y \quad &\iff \quad (z_{i_1}, z_{i_2}, x_{-(i_1,i_2)}) \succsim^{\tilde{p}} (z_{i_1}, z_{i_2}, y_{-(i_1,i_2)}) \\
&\iff \quad \sum_{i \in \operatorname{supp} p} u_i(x_i) + u_{i_1}(z) + u_{i_2}(z) \geq \sum_{i \in \operatorname{supp} p} u_i(x_i) + u_{i_1}(z) + u_{i_2}(z) \\
&\iff \quad \sum_{i \in \operatorname{supp} p} u_i(x_i) \geq \sum_{i \in \operatorname{supp} p} u_i(x_i)
\end{aligned}$$

□

B.2.2 Power Weighted Expected Utility for Preferences Conditional on a Co-probability

In this subsection, we show that the previous additive representation for $\succsim^p$ is a Power Weighted Expected Utility. This follows in several steps. First, we show that each additive representation uses the same utilities with different scaling. Second, we show that if two co-probabilities have two probabilities with the same relative likelihood, then the ratio of utility scalings must be the same in

the additive representations for these co-probabilities. This property will characterize power weighted probabilities by splitting probabilities while maintaining their ratios.

First, we show that binary lotteries with equal probabilities have an additive representation with the same utility.

Lemma B.48 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Then there exists $u : X \to \mathbb{R}$ continuous, strictly increasing such that for all $x, y \in X^{(1/2,1/2)}$,

$$x \succsim^{(1/2,1/2)} y \iff u(x_1) + u(x_2) \geq u(x_1) + u(x_2).$$

Proof. By the definition of $\succsim^{(1/2,1/2)}$, we have for all $x_1, x_2, y_1, y_2 \in X$,

$$(x_1, x_2) \succsim^{(1/2,1/2)} (y_1, y_2) \iff (x_2, x_1) \succsim^{(1/2,1/2)} (y_2, y_1).$$

Thus by Lemma B.47, for all $x_1, x_2, y_1, y_2 \in X$,

$$u_1(x_1) + u_2(x_2) \geq u_1(y_1) + u_2(y_2) \iff u_2(x_1) + u_1(x_2) \geq u_2(y_1) + u_1(y_2).$$

Thus by the uniqueness of additive representation for $\succsim^{(1/2,1/2)}$, there exist $\alpha \in \mathbb{R}_{++}$ and $\beta_1, \beta_2 \in \mathbb{R}$ such that

$$u_1 = \alpha u_2 + \beta_1 \text{ and } u_2 = \alpha u_1 + \beta_2.$$

Thus

$$u_1 = \alpha^2 u_1 + \alpha\beta_2 + \beta_1.$$

Hence

$$\alpha^2 = 1 \text{ and } \alpha > 0 \Rightarrow \alpha = 1.$$

Thus finally for all $x_1, x_2, y_1, y_2 \in X$,

$$(x_1, x_2) \succsim^{(1/2,1/2)} (y_1, y_2) \iff u_1(x_1) + u_1(x_2) + \beta_2 \geq u_1(y_1) + u_1(y_2) + \beta_2$$
$$\iff u_1(x_1) + u_1(x_2) \geq u_1(y_1) + u_1(y_2).$$

□

The next lemma shows that all different co-probabilities use a common utility up to different scalings. This utility is the above binary lottery utility with equal probabilities and follows from considering a co-probability $\frac{1}{2}p + \frac{1}{2}p$.

Lemma B.49 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Then there exists $u : X \to \mathbb{R}$ continuous, strictly increasing and for all co-probabilities p, there exists $\alpha^p \in \mathbb{R}_{++}^{\operatorname{supp} p}$ such that for all $x, y \in X^{\operatorname{supp} p}$,

$$x \succsim^p y \iff \sum_{i \in \operatorname{supp} p} \alpha_i^p u(x_i) \geq \alpha_i^p u(y_i).$$

Proof. Let u be the additive representation for $\succsim^{(1/2,1/2)}$ from Lemma B.48. Denote $\tilde{p} = 1/2p + 1/2p$. By Lemma B.47, there exists $(u_i)_{i \in \operatorname{supp} \tilde{p}}$ that gives an additive representation for $\succsim^{\tilde{p}}$. Let $i \in \operatorname{supp} p$. Let $x, y, z, k, c \in X$. By Lemma B.45,

$$\begin{aligned}
u(x) + u(y) \geq u(z) + u(k) \qquad &\iff \qquad (x, y) \succsim^{(1/2,1/2)} (z, k) \\
&\iff \qquad (x_i, y_{|\operatorname{supp} p|+i}, c_{-(i,|\operatorname{supp} p|+i)}) \succsim^{1/2p+1/2p)} (z_i, k_{|\operatorname{supp} p|+i}, c_{-(i,|\operatorname{supp} p|+i)}) \\
&\iff \qquad u_i(x) + u_{|\operatorname{supp} p|+i}(y) + \sum_{j \in -(i,|\operatorname{supp} p|+i)} u_j(c) \geq u_i(z) + u_{|\operatorname{supp} p|+i}(k) + \sum_{j \in -(i,|\operatorname{supp} p|+i)} u_j(c) \\
&\iff \qquad u_i(x) + u_{|\operatorname{supp} p|+i}(y) \geq u_i(z) + u_{|\operatorname{supp} p|+i}(k)
\end{aligned}$$

Thus $(u_i, u_{|\operatorname{supp} p|+i})$ give an additive representation for $\succsim^{(1/2,1/2)}$. So by the uniqueness of the additive representation, there exist $\alpha_i^p \in \mathbb{R}_{++}, \beta_i, \beta_{|\operatorname{supp} p|+i} \in \mathbb{R}$ such that $u_i = \alpha_i^p u + \beta_i$ and $u_{|\operatorname{supp} p|+i} = \alpha_i^p u + \beta_{|\operatorname{supp} p|+i}$.

The above holds for all $i \in \operatorname{supp} p$. Let $c \in X$. So we have for all $x, y \in X^{\operatorname{supp} p}$, by Lemma B.45,

$$\begin{aligned}
x \succsim^p y \qquad &\iff \qquad (x, c_{-(1,\ldots,|\operatorname{supp} p|)}) \succsim^{1/2p+1/2p} (y, c_{-(1,\ldots,|\operatorname{supp} p|)}) \\
&\iff \qquad \sum_{i \in \operatorname{supp} p} \alpha_i^p u(x_i) + \sum_{i \in \operatorname{supp} p} \alpha_i^p u(c) + \beta_i + \beta_{|\operatorname{supp} p|+i} \geq \sum_{i \in \operatorname{supp} p} \alpha_i^p u(y_i) + \sum_{i \in \operatorname{supp} p} \alpha_i^p u(c) + \beta_i + \beta_{|\operatorname{supp} p|+i} \\
&\iff \qquad \sum_{i \in \operatorname{supp} p} \alpha_i^p u(x_i) \geq \sum_{i \in \operatorname{supp} p} \alpha_i^p u(y_i).
\end{aligned}$$

□

Next, we define the probability weighting for binary lotteries as a function f of the likelihood ratio of the binary lottery. We will show that this function f is a power function.

Definition For a likelihood ratio $\alpha \in (0, \infty)$, denote by $f(\alpha) \in \mathbb{R}$ such that $\succsim^{(\frac{1}{1+\alpha}, \frac{\alpha}{1+\alpha})}$ is represented by $u(x_1) + f(\alpha)u(x_2)$.

First, we establish that $f(1) = 1$.

Lemma B.50 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Now $f(1) = 1$.

Proof. By Lemma B.48. □

Second, we establish that f is reciprocal symmetric around 1.

Lemma B.51 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Then for all $\alpha > 0$,

$$f(\alpha) = f\Big(\frac{1}{\alpha}\Big)^{-1}$$

Proof. For all $x, y, z, k \in X$, we have

$$\begin{aligned} &u(x) + f(\alpha)u(y) \geq u(z) + f(\alpha)u(k) \iff (x,y) \succsim^{(\frac{1}{1+\alpha}, \frac{\alpha}{1+\alpha})} (z,k) \\ &\iff (x,y) \succsim^{\left(\frac{1/\alpha}{1/\alpha+1}, \frac{1}{1/\alpha+1}\right)} (z,k) \iff (y,x) \succsim^{\left(\frac{1/\alpha}{1/\alpha+1}, \frac{1}{1/\alpha+1}\right)} (k,z) \\ &\iff u(y) + f\Big(\frac{1}{\alpha}\Big)u(x) \geq u(k) + f\Big(\frac{1}{\alpha}\Big)u(z) \iff f\Big(\frac{1}{\alpha}\Big)u(x) + u(y) \geq f\Big(\frac{1}{\alpha}\Big)u(z) + u(k). \end{aligned}$$

Since this holds for all $x, y, z, k \in X$, $(f(\frac{1}{\alpha})u, u)$ is another additive representation for $\succsim^{(\frac{1}{1+\alpha}, \frac{\alpha}{1+\alpha})}$. Thus by the uniqueness of the additive representation, there exists $\gamma > 0$ and $\beta_1, \beta_2 \in \mathbb{R}$ such that

$$u = \gamma f(\frac{1}{\alpha})u + \beta_1 \text{ and } f(\alpha)u = \gamma u + \beta_2.$$

Thus

$$1 = \gamma f\Big(\frac{1}{\alpha}\Big) \text{ and } f(\alpha) = \gamma \Longrightarrow 1 = f(\alpha) f\Big(\frac{1}{\alpha}\Big) \Longrightarrow f(\alpha) = f\Big(\frac{1}{\alpha}\Big)^{-1}.$$

□

Third, we show f is positive.

Lemma B.52 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Then for all $\alpha > 0$,

$$f(\alpha) > 0.$$

Proof. Follows from applying Lemma B.32 to $\succsim^{(\frac{1}{1+\alpha}, \frac{\alpha}{1+\alpha})}$ since u is strictly increasing. □

The next, lemma shows for two probabilities of a co-probability, the ratio of utility scalings for these probabilities only depends on the ratio of the probabilities and is given by the function f. This is the main lemma establishing that f is a power function.

Lemma B.53 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. Let p be a co-probability and u and $(\alpha_i^p)_{i \in \operatorname{supp} p}$ an additive representation for $\succsim^p$. Then for all $i, j \in \operatorname{supp} p$, $i \neq j$,

$$\frac{\alpha_j^p}{\alpha_i^p} = f\Big(\frac{p_j}{p_i}\Big).$$

Proof. Let $i, j \in \operatorname{supp} p$, $i \neq j$. Denote

$$\tilde{p} = \left(\frac{1}{1 + \frac{p_j}{p_i}}, \frac{\frac{p_j}{p_i}}{1 + \frac{p_j}{p_i}}\right) = \left(\frac{p_i}{p_i + p_j}, \frac{p_j}{p_i + p_j}\right).$$

Let $x, y, z, k, c \in X$. By Lemma B.45 and the definition of f,

$$\begin{aligned}
& u(x) + f\Big(\frac{p_j}{p_i}\Big)u(y) \geq u(z) + f\Big(\frac{p_j}{p_i}\Big)u(k) \iff (x,y) \succsim^{\tilde{p}} (z,k) \\
\iff & (x_i, y_j, c_{-(i,j)}) \succsim^{1/2p+1/2p)} (z_i, k_j, c_{-(i,j)}) \\
\iff & \alpha_i^p u(x) + \alpha_j^p u(y) + \sum_{l \in -(i,j)} \alpha_l^p u(c) \geq \alpha_i^p u(z) + \alpha_j^p u(k) + \sum_{l \in -(i,j)} \alpha_l^p u(c) \\
\iff & \alpha_i^p u(x) + \alpha_j^p u(y) \geq \alpha_i^p u(z) + \alpha_j^p u(k)
\end{aligned}$$

Thus $(\alpha_i^p u, \alpha_j^p u)$ give an additive representation for $\succsim^{\tilde{p}}$. So by the uniqueness of the additive representation, there exist $\alpha \in \mathbb{R}_{++}, \beta_i, \beta_j \in \mathbb{R}$ such that

$$u = \alpha\alpha_i^p u + \beta_i \text{ and } f\Big(\frac{p_j}{p_i}\Big)u = \alpha\alpha_j^p u + \beta_j.$$

Thus

$$1 = \alpha\alpha_i^p \text{ and } f\Big(\frac{p_j}{p_i}\Big) = \alpha\alpha_j^p.$$

Hence

$$\alpha = \frac{1}{\alpha_i^p} \text{ and } f\Big(\frac{p_j}{p_i}\Big) = \frac{\alpha_j^p}{\alpha_i^p}.$$

□

Next, by using the above lemma, for any ratio $x > 1$, we can clone this ratio n times to establish that $f(x^{1/n}) = f(x)^{1/n}$.

Lemma B.54 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. For all $x > 1$, $n \in \mathbb{N}$

$$f(x^{1/n}) = f(x)^{1/n}.$$

Proof. If $n = 1$, this holds by identity. So assume $n > 1$. Consider the non-normalized co-probability $p \coloneqq (1, x^{\frac{1}{n}}, x^{\frac{2}{n}}, \ldots, x^{\frac{n-1}{n}}, x)$. Let u and $(\alpha_i^p)_{i \in \operatorname{supp} p}$ an additive representation for $\succsim^p$. Now we have by Lemma B.53,

$$f(x) = \frac{a_{n+1}}{a_1} = \prod_{i=1}^{n} \frac{a_{i+1}}{a_i} = \prod_{i=1}^{n} f\Big(\frac{x^{\frac{i+1}{n}}}{x^{\frac{i}{n}}}\Big) = \prod_{i=1}^{n} f(x^{\frac{1}{n}}) = f(x^{\frac{1}{n}})^n$$

And so

$$f(x^{\frac{1}{n}}) = f(x)^{\frac{1}{n}}.$$

□

Next, by using the above lemma, for any ratio $x > 1$, we can split it again into n pieces but join the first m pieces to get that $f(x^{m/n}) = f(x)^{m/n}$.

Lemma B.55 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. For all $x > 1$, $n, m \in \mathbb{N}$, $m \leq n$,

$$f(x^{m/n}) = f(x)^{m/n}.$$

Proof. Consider the non-normalized co-probability $p := (1, x^{\frac{m}{n}}, x^{\frac{m+1}{n}}, \ldots, x^{\frac{n-1}{n}}, x)$. Let u and $(\alpha_i^p)_{i \in \operatorname{supp} p}$ an additive representation for $\succsim^p$. Now we have by Lemmas B.53 and B.54,

$$\begin{aligned} f(x) = \frac{a_{m-n+1}}{a_1} = \prod_{i=1}^{n-m+1} \frac{a_{i+1}}{a_i} = f(x^{\frac{m}{n}}) \prod_{i=m}^{n-1} f\Big(\frac{x^{\frac{i+1}{n}}}{x^{\frac{i}{n}}}\Big) = f(x^{\frac{m}{n}}) \prod_{i=m}^{n-1} f(x^{\frac{1}{n}}) \\ = f(x^{\frac{m}{n}}) f(x^{\frac{1}{n}})^{n-m} = f(x^{\frac{m}{n}}) f(x)^{\frac{n-m}{n}} \end{aligned}$$

And so

$$f(x^{\frac{m}{n}}) = f(x) f(x)^{-\frac{n-m}{n}} = f(x)^{\frac{m}{n}}$$

□

This shows that f is a power function on $\exp(\mathbb{Q}_{++})$.

Lemma B.56 Let $\succsim$ satisfy Axioms 2.1, 2.2, and 2.10. There exists $\lambda \in \mathbb{R}$ such that for all $x \in \exp(\mathbb{Q}_{++})$,

$$f(x) = x^{\lambda}.$$

Proof. By Lemma B.52, $f > 0$, so the function $\ln \circ f : \mathbb{R}_{++} \to \mathbb{R}$ is well-defined. Thus the function $\ln \circ f \circ \exp : \mathbb{R}_{++} \to \mathbb{R}$ is also well-defined. By the properties of exp and ln, we have for all $x \in \mathbb{R}_{++}, c \in \mathbb{Q}_{++} \cap (0, 1]$,

$$\begin{aligned} \ln \circ f \circ \exp(cx) = \ln\Big(f\big(\exp(cx)\big)\Big) = \ln\Big(f\big(\exp(x)^c\big)\Big) \\ \overset{\text{Lemma B.55}}{=} \ln\Big(f\big(\exp(x)\big)^c\Big) = c \ln\Big(f\big(\exp(x)\big)\Big) = c \ln \circ f \circ \exp(x). \end{aligned}$$

First, for all $x \in \mathbb{Q}_{++} \cap (0, 1]$, we have

$$\ln \circ f \circ \exp(x) = \ln \circ f \circ \exp(x \cdot 1) = x \ln \circ f \circ \exp(1).$$

Second, for all $x \in \mathbb{Q}_{++} \cap (1, \infty)$, we have since $1/x \in \mathbb{Q}_{++} \cap (0, 1)$

$$\begin{aligned} \ln \circ f \circ \exp(1) = \ln \circ f \circ \exp\Big(x \cdot \frac{1}{x}\Big) = \frac{1}{x} \ln \circ f \circ \exp(x) \\ \Longrightarrow \ln \circ f \circ \exp(x) = x \ln \circ f \circ \exp(1). \end{aligned}$$

By defining $\lambda := \ln \circ f \circ \exp(1)$, we have for all $x \in \mathbb{Q}_{++}$,

$$\ln \circ f \circ \exp(x) = \lambda x.$$

Hence for all $x \in \mathbb{Q}_{++}$,

$$f \circ \exp(x) = \exp \circ \ln \circ f \circ \exp(x) = \exp(\lambda x) = \exp(x)^{\lambda}.$$

Now we have for all $x \in \exp(\mathbb{Q}_{++})$,

$$f(x) = x^{\lambda}.$$

□

Finally, since $\exp(\mathbb{Q}_{++})$ is a dense set in $\mathbb{R}_{++}$, this shows the claim if f is either monotonic or continuous. This is the only place where we need continuity of probabilities,Axiom 2.3, to get that $\succsim^p$ has a power expected utility representation. First, we establish a simple lemma that we can take the limits of probabilities on both sides of preferences.

Lemma B.57 Let $\succsim$ satisfy Axioms 2.1 and 2.3. Let $p, q, r, s \in \Delta(\mathbb{N})$ and $x, y \in \mathrm{NDiag}(X^{p \cup q})$. If for all $\alpha \in (0, 1]$,

$$\sum_{i \in \operatorname{supp} p\alpha q} (p\alpha q)_i x_i \succsim \sum_{i \in \operatorname{supp} r\alpha s} (r\alpha s)_i y_i,$$

then

$$\sum_{i \in \operatorname{supp} q} q_i x_i \succsim \sum_{i \in \operatorname{supp} s} s_i y_i.$$

Proof. Assume, per contra, $\sum_{i \in \operatorname{supp} s} s_i y_i \succ \sum_{i \in \operatorname{supp} q} q_i x_i$. By Axiom 2.3, there exists $R \in \Delta(X)$ such that

$$\sum_{i \in \operatorname{supp} s} s_i y_i \succ R \succ \sum_{i \in \operatorname{supp} q} q_i x_i.$$

By Axiom 2.3, there exist $\alpha^x, \alpha^y \in (0, 1)$ such that for all $\alpha \in [0, \alpha^x)$ and $\tilde{\alpha} \in [0, \alpha^y)$,

$$\sum_{i \in \operatorname{supp} r\alpha s} (r\alpha s)_i y_i \succ R \succ \sum_{i \in \operatorname{supp} p\tilde{\alpha} q} (p\tilde{\alpha} q)_i x_i$$

which is a contradiction.

□

With continuity in probabilities, f is a power function. First, we establish this for $x > 1$.

Lemma B.58 Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.10. There exists $\lambda \in \mathbb{R}$ such that for all $x > 1$,

$$f(x) = x^{\lambda}.$$

Proof. By Lemma B.56, there exists $\lambda \in \mathbb{R}$ such that for all $x \in \exp(\mathbb{Q}_{++})$, $f(x) = x^{\lambda}$. We show that this holds for all $x > 1$. Assume, per contra, that there exists $\alpha > 1$ such that $f(\alpha) \neq \alpha^{\lambda}$. Assume w.l.o.g. that $f(\alpha) > \alpha^{\lambda}$. Let $p \in (0, 1)$ be such that $\frac{1-p}{p} = \alpha$ that is $p = \frac{1}{1+\alpha}$. Now $\exp(\mathbb{Q}_{++})$ is a dense set in $(1, \infty)$. We consider two cases. First, if $\lambda \leq 0$, then there exists a sequence $(\alpha_i)_{i=1}^{\infty} \subseteq \exp(\mathbb{Q}_{++})$

such that $\lim_{i\to\infty}\alpha_i=\alpha$ and for all i, $\alpha_i\leq\alpha_{i+1}$ so that $\alpha_i^\lambda\geq\alpha_{i+1}\lambda$. Second, if $\lambda\geq 0$, then there exists a sequence $(\alpha_i)_{i=1}^\infty\subseteq\exp(\mathbb{Q}_{++})$ such that $\lim_{i\to\infty}\alpha_i=\alpha$ and for all i, $\alpha_i\geq\alpha_{i+1}$ so that $\alpha_i^\lambda\geq\alpha_{i+1}\lambda$. So we assume that for all i, $\alpha_i^\lambda\geq\alpha_{i+1}\lambda$. Note that by Lemma B.56, for all i, $f(\alpha_i)=\alpha_i^\lambda$. Now denote for all $1\leq i$, $p^i=\frac{1}{1+\alpha_i}$. Since the power function to the power of $\lambda\in\mathbb{R}$ is continuous in $(1,\infty)$, there exists i^0 such that $f(\alpha)>\alpha_{i^0}^\lambda$.

Let $\tilde{x}^*>y^*>y_*>\tilde{x}_*$. First, if $u(\tilde{x}_*)+f(\alpha_{i^0})u(\tilde{x}^*)\geq u(y_*)+f(\alpha_{i^0})u(y^*)$, then by the continuity and strict monotonicity of u, there exists $x^*\in[\tilde{x}^*,y^*)$ such that $u(\tilde{x}_*)+f(\alpha_{i^0})u(x^*)\geq u(y_*)+f(\alpha_{i^0})u(y^*)$. Second, if $u(\tilde{x}_*)+f(\alpha_{i^0})u(\tilde{x}^*)\leq u(y_*)+f(\alpha_{i^0})u(y^*)$, then by the continuity and strict monotonicity of u, there exists $x_*\in(y_*,\tilde{x}_*]$ such that $u(x_*)+f(\alpha_{i^0})u(\tilde{x}^*)\geq u(y_*)+f(\alpha_{i^0})u(y^*)$. So we have $x^*>y^*>y_*>x_*$ such that $u(x_*)+f(\alpha_{i^0})u(x^*)=u(y_*)+f(\alpha_{i^0})u(y^*)$. Now since $\alpha_i^\lambda\geq\alpha_{i+1}^\lambda$ for all i, and by the monotonicity of u, we have $u(x^*)>u(y^*)$, so for all $i>i^0$, since $f(\alpha_i)-f(\alpha_{i^0})\leq 0$,

$$\begin{aligned}u(x_*)+f(\alpha_i)u(x^*)&=u(x_*)+f(\alpha_{i^0})u(x^*)+u(x^*)\big(f(\alpha_i)-f(\alpha_{i^0})\big)\\ &\leq u(y_*)+f(\alpha_{i^0})u(y^*)+u(y^*)\big(f(\alpha_i)-f(\alpha_{i^0})\big)=u(y_*)+f(\alpha_i)u(y^*).\end{aligned}$$

Thus by the definition of f, we have for all $i>i^0$,

$$p_iy_*+(1-p_i)y^*\gtrsim p_ix_*+(1-p_i)x^*.$$

Now since $\lim_{i\to\infty}p_i=p$, we have by Lemma B.57,

$$py_*+(1-p)y^*\gtrsim px_*+(1-p)x^*.$$

But we have since $f(\alpha_{i^0})<f(\alpha)$,

$$\begin{aligned}u(x_*)+f(\alpha)u(x^*)&=u(x_*)+f(\alpha_{i^0})u(x^*)+u(x^*)\big(f(\alpha)-f(\alpha_{i^0})\big)\\ &>u(y_*)+f(\alpha_{i^0})u(y^*)+u(y^*)\big(f(\alpha)-f(\alpha_{i^0})\big)=u(y_*)+f(\alpha)u(y^*).\end{aligned}$$

So by the definition of f, since $\alpha=\frac{1}{1-p}$,

$$px_*+(1-p)x^*\succ py_*+(1-p)y^*.$$

which is a contradiction. □

Second, we establish that f is a power function for all $x\in\mathbb{R}_{++}$.

Lemma B.59 Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.10. There exists $\lambda\in\mathbb{R}$ such that for all $x\in\mathbb{R}_{++}$,

$$f(x)=x^\lambda.$$

Proof. From Lemma B.59, there exists $\lambda \in \mathbb{R}$ such that for all $x > 1$,

$$f(x) = x^{\lambda}.$$

Now, we have

$$f(1) = 1 = 1^{\lambda}.$$

Finally, for all $x \in (0,1)$, we have by Lemma B.51 and since $\frac{1}{x} > 1$,

$$f(x) = f\Big(\frac{1}{x}\Big)^{-1} = \bigg(\Big(\frac{1}{x}\Big)^{\lambda}\bigg)^{-1} = x^{\lambda}.$$

□

Lemma B.60 Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.10. There exists $\lambda \in \mathbb{R}_{++}$ such that for all $x \in \mathbb{R}_{++}$,

$$f(x) = x^{\lambda}.$$

Proof. From Lemma B.59, there exists $\lambda \in \mathbb{R}$ such that for all $x \in \mathbb{R}_{++}$,

$$f(x) = x^{\lambda}.$$

Assume, per contra, that $\lambda \leq 0$. Let $\tilde{x}^* > y^* > y_* > \tilde{x}_*$. Denote for all $i \in \mathbb{N}$, $p_i = \frac{1}{i}$.

First, if $u(\tilde{x}_*) + u(\tilde{x}^*) \geq u(y_*) + u(y^*)$, then by the continuity and strict monotonicity of u, there exists $x^* \in [\tilde{x}^*, y^*)$ such that $u(\tilde{x}_*) + u(x^*) \geq u(y_*) + u(y^*)$. Second, if $u(\tilde{x}_*) + u(\tilde{x}^*) \leq u(y_*) + u(y^*)$, then by the continuity and strict monotonicity of u, there exists $x_* \in (y_*, \tilde{x}_*]$ such that $u(x_*) + u(\tilde{x}^*) \geq u(y_*) + u(y^*)$. So we have $x^* > y^* > y_* > x_*$ such that $u(x_*) + u(x^*) = u(y_*) + u(y^*)$. Now we have for all $i \geq 2$, since f is a weakly decreasing function and so for all $i \geq 2$, $f\Big(\frac{1-p_i}{p_i}\Big) \leq f(1) = 1$ and since $u(x^*) > u(y^*)$,

$$\begin{aligned} u(x_*) + f\Big(\frac{1-p_i}{p_i}\Big)u(x^*) &= u(x_*) + u(x^*) + u(x^*)\bigg(f\Big(\frac{1-p_i}{p_i}\Big) - 1\bigg) \\ \leq u(y_*) + u(y^*) + u(y^*)\bigg(f\Big(\frac{1-p_i}{p_i}\Big) - 1\bigg) &= u(y_*) + f\Big(\frac{1-p_i}{p_i}\Big) - 1\Big)u(y^*). \end{aligned}$$

By the definition of f, we have for all $i \geq 2$,

$$p_i y_* + (1-p_i)y^* \succsim p_i x_* + (1-p_i)x^*$$

Now since $\lim_{i\to\infty} p_i = 1$, we have by Lemma B.57,

$$y^* \succsim x^*$$

which contradicts Axiom 2.1.

□

This establishes that $\succsim^p$ has a power expected utility representation with a common utility and a common degree of power for all co-probabilities.

Lemma B.61 Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.10. Then there exist $u : X \to \mathbb{R}$ continuous, strictly increasing and $\lambda > 0$ such that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, $x, y \in \mathrm{NDiag}(X^p)$

$$x \succsim^p y \iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) \geq p_i^\lambda u(y_i).$$

Proof. Follows from Lemmas B.49, B.53, and B.60 □

Finally, under restricted FOSD, we have $\lambda = 1$ by splitting a probability into two equal pieces.

Lemma B.62 Let $\succsim$ satisfy Axioms 2.1 to 2.3, 2.5, and 2.10. For all $x \in \mathbb{R}_{++}$,

$$f(x) = x.$$

Proof. Consider a co-probability $p = (1/8, 1/8, 1/4, 1/2)$ and $x^* \in X$. Since X is open, there exists $\varepsilon^* > 0$ such that $x^* + \varepsilon^*, x^* - \varepsilon^* \in X$. Let $0 < \varepsilon^1 < \varepsilon^2 < \varepsilon^*$. By Axiom 2.5, we have for all $0 < \varepsilon^3 < \varepsilon^*$ such that $\varepsilon^2 + \varepsilon^3 < \varepsilon^*$, $\varepsilon^2 \neq \varepsilon^1 + \varepsilon^3$, and $\varepsilon^1 \neq \varepsilon^2 - \varepsilon^3$

$$\big(p, (x^* + \varepsilon^1 + \varepsilon^3, x^* + \varepsilon^1, x^* + \varepsilon^2, x^*)\big) \succsim \big(p, (x^* + \varepsilon^2 - \varepsilon^3, x^* + \varepsilon^2, x^* + \varepsilon^1, x^*)\big)$$

and

$$\big(p, (x^* + \varepsilon^1 - \varepsilon^3, x^* + \varepsilon^1, x^* + \varepsilon^2, x^*)\big) \precsim \big(p, (x^* + \varepsilon^2 + \varepsilon^3, x^* + \varepsilon^2, x^* + \varepsilon^1, x^*)\big).$$

Thus by Lemma B.61, we have

$$\begin{aligned} &p_1^\lambda u(x^* + \varepsilon^1 + \varepsilon^3) + p_2^\lambda u(x^* + \varepsilon^1) + p_3^\lambda u(x^* + \varepsilon^2) + p_4^\lambda u(x^*) \\ &\geq p_1^\lambda u(x^* + \varepsilon^2 - \varepsilon^3) + p_2^\lambda u(x^* + \varepsilon^2) + p_3^\lambda u(x^* + \varepsilon^1) + p_4^\lambda u(x^*) \end{aligned}$$

and

$$\begin{aligned} &p_1^\lambda u(x^* + \varepsilon^1 - \varepsilon^3) + p_2^\lambda u(x^* + \varepsilon^1) + p_3^\lambda u(x^* + \varepsilon^2) + p_4^\lambda u(x^*) \\ &\leq p_1^\lambda u(x^* + \varepsilon^2 + \varepsilon^3) + p_2^\lambda u(x^* + \varepsilon^2) + p_3^\lambda u(x^* + \varepsilon^1) + p_4^\lambda u(x^*). \end{aligned}$$

So by taking $\varepsilon^3 \to 0$, we have by the continuity of u since $2(1/8)^\lambda = 2^{-3\lambda+1}$ and $(1/4)^\lambda = 2^{-2\lambda}$

$$2^{-3\lambda+1} u(x^* + \varepsilon^1) + 2^{-2\lambda} u(x^* + \varepsilon^2) = 2^{-3\lambda+1} u(x^* + \varepsilon^2) + 2^{-2\lambda} u(x^* + \varepsilon^1).$$

Thus we have

$$(2^{-3\lambda+1} - 2^{-2\lambda})\big(u(x^* + \varepsilon^1) - u(x^* + \varepsilon^2)\big) = 0.$$

Since $\varepsilon^1 \neq \varepsilon^2$ and u is strictly increasing, we have

$$2^{-3\lambda+1} = 2^{-2\lambda}.$$

Since $\mathbb{R} \ni x \mapsto 2^x$ is a strictly increasing function, we have

$$-3\lambda + 1 = -2\lambda \Rightarrow \lambda = 1.$$

□

B.3 Variational Expected Utility

Next, we move on to showing Variational Expected Utility representation by considering how different power weighted probability representations for different co-probabilities relate to each other and show Theorem 2.5. Here, we add the axiom Axiom 2.12 and build on Appendix B.2. This proof follows in two parts. First, we show that locally for a common preference range different power weighted expected utilities for different co-probabilities are the same up to different levels. Especially the intensities of the different representations are the same. Second, we extend this local property to a global property by a transfinite induction.

B.3.1 Local Power Weighted Expected Utility Representation

This proof follows in two steps. First, we show that if we have two co-probabilities that share a common probability and a common preference range where the two co-probabilities share a common prize with the common probability at the boundaries of the preference range, then we can use this common probability and the common prizes to measure the length of the preference interval using Axiom 2.12 in terms of both of the co-probabilities. This gives that the intensities of both power expected utility representations must be the same. That is the representations can differ only by a level in this preference interval.

Lemma B.63 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exists a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i \in \operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12.

Assume that $p, q \in \Delta^4(\mathbb{N}), i^* \in \operatorname{supp} p, j^* \in \operatorname{supp} q$, and $p_{i^*} = q_{j^*}$. If there exist $x^1, x^2, x^3 \in \operatorname{NDiag}(X^p)$ and $y^1, y^2, y^3 \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \sim (q, y^1)$, $(p, x^2) \sim (q, y^2)$, $(p, x^3) \sim (q, y^3)$,

$(p,x^3) \succsim (p,x^2) \succsim (p,x^1)$, $x^1_{i^*} = y^1_{j^*}$, and $x^3_{i^*} = y^3_{j^*}$, then for all $x \in \mathrm{NDiag}(X^p), y \in \mathrm{NDiag}(X^q)$ such that $(p,x^3) \succsim (p,x), (q,y) \succsim (p,x^1)$,

$$(p,x) \succsim (q,y) \iff \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^2) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^2)$$

and

$$(q,y) \succsim (p,x) \iff \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^2) \geq \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^2).$$

Additionally, for all $x \in \mathrm{NDiag}(X^p), y \in \mathrm{NDiag}(X^q)$ such that $(p,x^3) \succsim (p,x) \sim (q,y) \succsim (p,x^1)$, there exist $\tilde{x} \in \mathrm{NDiag}(X^p), \tilde{y} \in \mathrm{NDiag}(X^q)$ such that $\tilde{x}_{i^*} = \tilde{y}_{j^*}$, $(p,\tilde{x}) \sim (p,x)$ and $(q,\tilde{y}) \sim (q,y)$.

Proof. By permuting the indices for q, we can assume that $i^* = j^*$. Second, by permuting the indices for p and q, we can assume that $i^* = j^* = 1$. First, we show that we can assume that $x^3_1 > x^1_1$: Since X is open there exist $\tilde{x}^3_1, \tilde{x}^1_1 \in X$ such that $\tilde{x}^3_1 > \max\{x^3_1, x^1_1\} \geq \min\{x^3_1, x^1_1\} > \tilde{x}^1_1$ and by Axiom 2.12,

$$\big(p,(\tilde{x}^3_i, x^3_{-i})\big) \sim \big(p,(\tilde{x}^3_i, y^3_{-i})\big) \succsim (p,x^3) \succsim (p,x^1) \succsim \big(p,(\tilde{x}^1_i, x^1_{-i})\big) \sim \big(p,(\tilde{x}^1_i, y^1_{-i})\big).$$

So now we could prove the claim for this larger interval instead. So assume that $x^3_1 > x^1_1$.

Assume w.l.o.g. that i as above is 1. Since u is continuous and $|\operatorname{supp} p| \geq 4$, there exist $\tilde{x}^3 \in \mathrm{NDiag}(X^p), \tilde{y}^3 \in \mathrm{NDiag}(X^q)$ such that $\tilde{x}^3 = \tilde{x}^3 = x^3_1 = y^3_1$, for all $i \in \operatorname{supp} p$, $\tilde{x}^3_i \neq x^1_1$, for all $i \in \operatorname{supp} q$, $\tilde{y}^3_i \neq x^3_1$,

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u(\tilde{x}^3_i) = \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x^3_i), \text{ and } \sum_{i\in\operatorname{supp} q} q_i^\lambda u(\tilde{y}^3_i) = \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y^3_i). \tag{B.62}$$

Now by the additive representations for $\succsim^p$ and $\succsim^q$,

$$(p,\tilde{x}^3) \sim (p,x^3) \sim (q,y^1) \sim (q,\tilde{y}^3). \tag{B.63}$$

Let $n^0 \in \mathbb{N}$ be the smallest non-negative integer such that

$$n^0 \geq \frac{\sum_{i\in\operatorname{supp} p\setminus\{1\}} p_i^\lambda u(x^3_i) - \sum_{i\in\operatorname{supp} p\setminus\{1\}} p_i^\lambda u(x^1_i)}{\big(u(x^3_1) - u(x^1_1)\big)p_1^\lambda}.$$

We define by induction for each $1 \leq n \leq n^0$, $\varphi(p,n) \in \mathrm{NDiag}(X^p), \varphi(q,n) \in \mathrm{NDiag}(X^q)$ such that $\varphi(p,n)_1 = \varphi(q,n)_1 = x^1_1$, for all $i \in \operatorname{supp} p, j \in \operatorname{supp} q$, $\varphi(p,n)_i \neq x^3_1$, $\varphi(q,n)_j \neq x^3_1$, $\big(p,\varphi(p,n)\big) \sim \big(q,\varphi(q,n)\big)$ for all $1 \leq n \leq n^0$, $\big(p,(x^3_1, \varphi(p,n)_{-1})\big) \sim \big(p,\varphi(p,n+1)\big)$. By Axiom 2.12, we especially have then for all $1 \leq n \leq n^0$, $\big(q,(x^3_1, \varphi(q,n)_{-1})\big) \sim \big(q,\varphi(q,n+1)\big)$.

First, since u is continuous and $|\operatorname{supp} p| \geq 4$, there exists $\varphi(p,1) \in \mathrm{NDiag}(X^p), \varphi(q,1) \in \mathrm{NDiag}(X^q)$ such that $\varphi(q,1) = \varphi(p,1)_1 = x^1$, for all $i \in \operatorname{supp} p$, $\varphi(p,1)_i \neq x^3_1$, for all $i \in \operatorname{supp} q$, $\varphi(q,1)_i \neq x^3_1$,

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u(\varphi(p,1)_i) = \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x^1_i), \text{ and } \sum_{i\in\operatorname{supp} q} q_i^\lambda u(\varphi(q,1)_i) = \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y^1_i). \tag{B.64}$$

Now by the additive representations for $\succsim^p$ and $\succsim^q$,

$$\big(p,\varphi(p,1)\big)\sim(p,x^1)\sim(q,y^1)\sim\big(q,\varphi(q,1)\big).$$

Now if $n^0{=}0$, we are done. So assume that $n^0{>}0$. Let $1{<}n{\leq}n^0$ assume that $\varphi(p,n{-}1),\varphi(q,n{-}1)$ as above have been defined. Now by the additive representations for $\succsim^p$, we have for all $1<m<n$,

$$\begin{aligned}\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big(\varphi(p,m)_i\big)&=\sum_{i\in\operatorname{supp}p\setminus\{1\}}p_i^\lambda u\big(\varphi(p,m-1)_i\big)+p_1^\lambda u(x_1^3)\\&=\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big(\varphi(p,m-1)_i\big)+p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big).\end{aligned}\tag{B.65}$$

Thus by doing the recursion to first step,

$$\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big(\varphi(p,n-1)_i\big)=\sum_{i\in\operatorname{supp}p}p_i^\lambda u(x_i^1)+(n-2)p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big).$$

Now by the choice of n^0 and since $n^0>0$,

$$(n^0-1)\big(u(x_1^3)-u(x_1^1)\big)p_1^\lambda<\sum_{i\in\operatorname{supp}p\setminus\{1\}}p_i^\lambda u(x_i^3)-\sum_{i\in\operatorname{supp}p\setminus\{1\}}p_i^\lambda u(x_i^1).$$

Hence we have

$$\begin{aligned}\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big((x_1^3,\varphi(p,n-1)_{-1})_i\big)&=\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big(\varphi(p,n-1)_i\big)+p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big)\\&\overset{(B.64)}{=}\sum_{i\in\operatorname{supp}p}p_i^\lambda u(x_i^1)+(n-2+1)p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big)\\&\leq\sum_{i\in\operatorname{supp}p}p_i^\lambda u(x_i^1)+(n^0-1)p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big)\\&<\sum_{i\in\operatorname{supp}p}p_i^\lambda u(x_i^1)+\sum_{i\in\operatorname{supp}p\setminus\{1\}}p_i^\lambda u(x_i^3)-\sum_{i\in\operatorname{supp}p\setminus\{1\}}p_i^\lambda u(x_i^1)\\&=\sum_{i\in\operatorname{supp}p\setminus\{1\}}p_i^\lambda u(x_i^3)+p_1^\lambda u(x_1^1)=\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big((x_1^1,x_{-1}^3)_i\big).\end{aligned}\tag{B.66}$$

And we have $\varphi(p,n-1)_1=x_1^1$ and

$$\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big((x_1^3,\varphi(p,n-1)_{-1})_i\big)>\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big(\varphi(p,n-1)_i\big).$$

Now since X is a connected set as an interval u is continuous, and $|\operatorname{supp}p|\geq4$, there exists $\varphi(p,n)\in\operatorname{NDiag}(X^p)$ such that $\varphi(p,n)_1=x^1$, for all $i\in\operatorname{supp}p$, $\varphi(p,n)_i\neq x_1^3$, and

$$\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big((x_1^3,\varphi(p,n-1)_{-1})_i\big)=\sum_{i\in\operatorname{supp}p}p_i^\lambda u\big(\varphi(p,n)_i\big).$$

Thus by the additive representation for $\succsim^p$, $(x_1^3,\varphi(p,n-1)_{-1})\sim\varphi(p,n)$.

Second, by the induction assumption, additive representations, and Axiom 2.12, we have

$$\varphi(q,n{-}1){\sim}\varphi(p,n{-}1){\Rightarrow}\big(x_1^3,\varphi(q,n{-}1)_{-1}\big){\sim}\big(x_1^3,\varphi(p,n{-}1)_{-1}\big)\overset{(B.62,B.66)}{\prec}(x_1^1,\tilde{x}_{-1}^3)\overset{(B.63)}{\sim}(x_1^1,\tilde{x}_{-1}^3).\tag{B.67}$$

And by Lemma B.31, $(x_1^3, \varphi(q, n-1)_{-1}) \succ \varphi(q, n-1)$. Thus by the additive representation,

$$\sum_{i \in \operatorname{supp} q} q_i^\lambda u\big((x_1^1, \tilde{y}_{-1}^3)_i\big) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u\big((x_1^3, \varphi(q, n-1)_{-1})_i\big) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u\big(\varphi(q, n-1)_i\big).$$

Now since X is a connected set as an interval u is continuous, and $|\operatorname{supp} q| \geq 4$, there exists $\varphi(q, n) \in \operatorname{NDiag}(X^q)$ such that $\varphi(q, n)_1 = x^1$, for all $i \in \operatorname{supp} q$, $\varphi(q, n)_i \neq x_1^3$, and

$$\sum_{i \in \operatorname{supp} q} q_i^\lambda u\big((x_1^3, \varphi(q, n-1)_{-1})_i\big) = \sum_{i \in \operatorname{supp} q} q_i^\lambda u\big(\varphi(q, n)_i\big).$$

Thus by the additive representation and (B.67),

$$\varphi(q, n) \sim (x_1^3, \varphi(q, n-1)_{-1}) \sim (x_1^3, \varphi(p, n-1)_{-1}) \sim \varphi(p, n).$$

Next by the definition of n^0, we have

$$n^0\big(u(x_1^3) - u(x_1^1)\big)p_1^\lambda \geq \sum_{i \in \operatorname{supp} p \setminus \{1\}} p_i^\lambda u(x_i^3) - \sum_{i \in \operatorname{supp} p \setminus \{1\}} p_i^\lambda u(x_i^1).$$

and so

$$\begin{aligned}
&\sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^3, \varphi(p, n^0)_{-1})_i\big) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big(\varphi(p, n^0)_i\big) + p_1^\lambda\big(u(x_1^3) - u(x_1^1)\big) \\
&\overset{(B.64, B.65)}{=} \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^1) + (n^0 - 1 + 1)p_1^\lambda\big(u(x_1^3) - u(x_1^1)\big) \\
&\geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^1) + \sum_{i \in \operatorname{supp} p \setminus \{1\}} p_i^\lambda u(x_i^3) - \sum_{i \in \operatorname{supp} p \setminus \{1\}} p_i^\lambda u(x_i^1) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^1, x_{-1}^3)_i\big).
\end{aligned}$$

Thus there exists $a_1 \in [x_1^1, x_1^3]$ such that

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^3, \varphi(p, n^0)_{-1})_i\big) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((a_1, x_{-1}^3)_i\big).$$

Next, replace $\tilde{x}^3, \tilde{y}^3$ by $\hat{x}^3 \in \operatorname{NDiag}(X^p), \hat{y}^3 \in \operatorname{NDiag}(X^q)$ such that $\hat{x}^3 = \hat{x}^3 = x_1^3 = y_1^3$, for all $i \in \operatorname{supp} p$, $a_1 \neq \tilde{x}_i^3 \neq x_1^1$, for all $i \in \operatorname{supp} q$, $a_1 \neq \hat{y}_i^3 \neq x_1^3$,

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u(\hat{x}_i^3) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^3), \text{ and } \sum_{i \in \operatorname{supp} q} q_i^\lambda u(\hat{y}_i^3) = \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i^3) \tag{B.68}$$

that exist as above.

Now we have by the additive representation and Axiom 2.12, $\hat{x}^3 \sim x^3 \sim y^3 \sim \hat{y}^3$ and

$$(q, (a_1, \hat{y}_{-1}^3)) \sim (p, (a_1, \hat{x}_{-1}^3)) \sim (x_1^3, \varphi(p, n^0)) \sim (x_1^3, \varphi(q, n^0)). \tag{B.69}$$

Now we have symmetrically to (B.65) for q since $q_1 = p_1$ for all $1 < n \leq n^0$,

$$\begin{aligned}
&\sum_{i \in \operatorname{supp} q} q_i^\lambda u\big(\varphi(q, n)_i\big) = \sum_{i \in \operatorname{supp} q} q_i^\lambda u\big(\varphi(q, 1)_i\big) + (n^0 - 1)q_1^\lambda\big(u(x_1^3) - u(x_1^1)\big) \qquad \text{(B.70)} \\
&\overset{(B.64)}{=} \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_1^1) + (n-1)p_1^\lambda\big(u(x_1^3) - u(x_1^1)\big).
\end{aligned}$$

Additionally

$$\sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^3) \overset{\text{(B.68)}}{=} \sum_{i\in\operatorname{supp} q} q_i^\lambda u\big((a_1,\hat{y}^3_{-1})_i\big) + q_1^\lambda\big(u(x_1^3)-u(a_1)\big) \tag{B.71}$$
$$\overset{\text{(B.69)}}{=} \sum_{i\in\operatorname{supp} q} q_i^\lambda u\big(\varphi(q,n^0)_i\big) + q_1^\lambda\big(u(x_1^3)-u(a_1)\big)$$
$$\overset{\text{(B.70)}}{=} \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_1^1) + (n^0-1)p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big) + p_1^\lambda\big(u(x_1^3)-u(a_1)\big).$$

Additionally, symmetrically for p, we have

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^3) = \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_1^1) + (n^0-1)p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big) + p_1^\lambda\big(u(x_1^3)-u(a_1)\big). \tag{B.72}$$

Finally, let $x\in \mathrm{NDiag}(X^p), y\in \mathrm{NDiag}(X^q)$ be such that $(p,x^3)\succsim(p,x),(q,y)\succsim(p,x^1),$. We will first show that

$$(p,x)\succsim(q,y) \iff \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^1) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^1).$$

Now by the additive representations, we have

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^3) \geq \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) \geq \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^1)$$

and

$$\sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^3) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^1).$$

Now since $u(a_1)\geq u(x_1^1)$, by (B.71,B.72), there exist $0\leq i^x, i^y\leq n^0+1$ such that

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^1) + (i^x+1)p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big) > \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) \geq \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^1) + i^x p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big)$$

and

$$\sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^1) + (i^y+1)p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big) > \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^1) + i^y p_1^\lambda\big(u(x_1^3)-u(x_1^1)\big).$$

For the case $i^x = n^0+1$, define $\varphi(p,n^0+1)=\hat{x}^3$, and $\varphi(q,n^0+1)=\hat{y}^3$. For the case $i^x=n^0$, define $\varphi(p,n^0)=(a_1,\hat{x}^3_{-1})$ and $\varphi(q,n^0)=(a_1,\hat{y}^3_{-1})$.

First, assume that $i^x\neq i^y$. In this case, we have

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^1) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^1) \overset{i^x\neq i^y}{\iff} i^x > i^y$$
$$\iff \varphi(p,i^x)\succ\varphi(p,i^y)\sim\varphi(q,i^y) \iff \varphi(p,i^x)\succ(q,y) \overset{i^x\neq i^y}{\iff} (p,x)\succ(q,y),$$

where the first equivalency follows from the definition of i^x, i^y and the assumption that $i^x\neq i^y$, the second equivalency follows from the additive representation for $\succsim^p$, (B.65), and the definition of φ with extensions to n^0 and n^0+1, the third equivalency follows from the definition of i^y, and the last one from the assumption that $i^x\neq i^y$ and the definition of i^x.

Next, assume that $i^x = i^y$. If $i^x = n^0 + 1$, then $(p, x) \sim (p, x^3) \sim (q, y^3) \sim (q, y)$. So assume that $i^x < n^0 + 1$. Now by the definitions of i^x, i^y, φ, and continuity of utility, there exist $z_1^x, z_1^y \in [x_1^1, x_1^3]$ such that

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) = \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((z_1^x, \varphi(p, i^y)_{-1})_i\big) = \sum_{i\in\operatorname{supp} p} p_i^\lambda u(\varphi(p, i^y)_i) + p_1^\lambda\Big(u(z_1^x) - u(\varphi(p, i^y)_1)\Big)$$

and since $p_1 = q_1$ and $\varphi(p, i^y)_1 = \varphi(q, i^y)_1$

$$\sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) = \sum_{i\in\operatorname{supp} q} q_i^\lambda u\big((z_1^y, \varphi(q, i^y)_{-1})_i\big) = \sum_{i\in\operatorname{supp} q} q_i^\lambda u(\varphi(q, i^y)_i) + p_1^\lambda\Big(u(z_1^x) - u(\varphi(p, i^y)_1)\Big).$$

Now, since u is continuous and $|\operatorname{supp} p| \geq 4$, there exist $\tilde\varphi(p, i^y) \in \operatorname{NDiag}(X^p), \tilde\varphi(q, i^y) \in \operatorname{NDiag}(X^q)$ such that $\tilde\varphi(q, i^y) = \tilde\varphi(p, i^y)_1 = x^1$, for all $z \in \{x_1^1, x_1^3, z_1^x, z_1^y\}$, for all $i \in \operatorname{supp} p$, $\varphi(p, i^y)_i \neq z$, for all $j \in \operatorname{supp} q$, $\varphi(q, i^y)_j \neq z$,

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u\big(\tilde\varphi(p, i^y)_i\big) = \sum_{i\in\operatorname{supp} p} p_i^\lambda u(\varphi(p, i^y)_i), \text{ and } \sum_{i\in\operatorname{supp} q} q_i^\lambda u\big(\tilde\varphi(q, i^y)_i\big) = \sum_{i\in\operatorname{supp} q} q_i^\lambda u(\varphi(q, i^y)_i).$$

Now by the additive representations for $\succsim^p$ and $\succsim^q$,

$$(p, \tilde\varphi(p, i^y)) \sim (p, \varphi(p, i^y)) \sim (q, \varphi(q, i^y)) \sim (q, \tilde\varphi(q, i^y))$$

and so by Axiom 2.12,

$$\Big(q, (z_1^y, \tilde\varphi(q, i^y))\Big) \sim \Big(q, (z_1^y, \tilde\varphi(q, i^y))\Big). \tag{B.73}$$

Thus we have by (B.65,B.70),

$$\begin{aligned}
&\sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^1) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^1)\\
\iff &\sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^1) + (i^y - 1)q_1^\lambda\big(u(x_1^3) - u(x_1^1)\big) + p_1^\lambda\Big(u(z_1^x) - u(\varphi(p, i^y)_1)\Big) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i^1)\\
&\qquad \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^1) + (i^y - 1)q_1^\lambda\big(u(x_1^3) - u(x_1^1)\big) + p_1^\lambda\Big(u(z_1^x) - u(\varphi(p, i^y)_1)\Big) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i^1)\\
\iff &\qquad p_1^\lambda\Big(u(z_1^x) - u(\varphi(p, i^y)_1)\Big) \geq p_1^\lambda\Big(u(z_1^y) - u(\varphi(p, i^y)_1)\Big)\\
\iff &\qquad u(z_1^x) \geq u(z_1^y)\\
\iff &\qquad \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((z_1^x, \tilde\varphi(p, i^y)_{-1})_i\big) \geq \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((z_1^y, \tilde\varphi(p, i^y)_{-1})_i\big)\\
\iff &\qquad \Big(p, (z_1^x, \tilde\varphi(p, i^y)_{-1})\Big) \succsim \Big(p, (z_1^y, \tilde\varphi(p, i^y)_{-1})\Big) \overset{(B.73)}{\sim} \Big(q, (z_1^y, \tilde\varphi(q, i^y)_{-1})\Big)\\
\iff &\qquad (p, x) \succsim (q, y),
\end{aligned}$$

where the second to last equivalence follows from the additive representation for $\succsim^p$.

Thus we have for all $x \in \mathrm{NDiag}(X^p), y \in \mathrm{NDiag}(X^q)$ such that $(p, x^3) \succsim (p, x), (q, y) \succsim (p, x^1)$,

$$(p,x) \succsim (q,y) \iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^1) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i^1)$$

and symmetrically, we have

$$(q,y) \succsim (p,x) \iff \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i^1) \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^1).$$

Now since $(p, x^2) \sim (q, y^2)$, we have

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^2) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^1) = \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i^2) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i^1)$$

and so we can add it to the above equivalencies to get the claim.

Finally, the additional claim follows from the definition of φ, Axiom 2.12, continuity of utility, and the main claim.

□

The second step of the proof shows that we can use continuity to create common probabilities and common prizes for any two co-probabilities. That is if we have two co-probabilities with a common preference range, then by continuity and above, the intensities of the representations must be the same in this interval.

Lemma B.64 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exists a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i \in \operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12.

Assume that $p, q \in \Delta^4(\mathbb{N}), x \in \mathrm{NDiag}(X^p), y \in \mathrm{NDiag}(X^q)$ are such that $(p, x) \sim (q, y)$. Then there exist $x^1, x^2 \in \mathrm{NDiag}(X^p), y^1, y^2 \in \mathrm{NDiag}(X^q)$ such that

$$(p, x^2) \sim (q, y^2) \succ (p, x) \sim (q, y) \succ (p, x^1) \sim (q, y^1)$$

and for all $\tilde{x} \in \mathrm{NDiag}(X^p), \tilde{y} \in \mathrm{NDiag}(X^q)$ such that $(p, x^2) \succsim (p, \tilde{x}), (q, \tilde{y}) \succsim (p, x^1)$,

$$(p,\tilde{x}) \succsim (q,\tilde{y}) \iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(\tilde{x}_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u(\tilde{y}_i) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i)$$

and

$$(q,\tilde{y}) \succsim (p,\tilde{x}) \iff \sum_{i \in \operatorname{supp} q} q_i^\lambda u(\tilde{y}_i) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u(\tilde{x}_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$

Proof. Let $c \in X, i^p \in \operatorname{supp} p, i^q \in \operatorname{supp} q$. Denote $c_{i^p} = c$. Since X is open, there exist $x_{i^p}^1, x_{i^p}^2, x_{i^p}^3, x_{i^p}^4 \in X$ such that for all $i \in \operatorname{supp} p, j \in \{1, 2, 3, 4\}$, $x_i \neq x_{i^p}^j$ and

$$x_{i^p}^4 > x_{i^p}^3 > x_{i^p} > x_{i^p}^2 > x_{i^p}^1.$$

By Lemma B.31,

$$\big(p,(x^4_{i^p},x_{-i^p})\big)\succ\big(p,(x^3_{i^p},x_{-i^p})\big)\succ(p,x)\succ\big(p,(x^2_{i^p},x_{-i^p})\big)\succ\big(p,(x^1_{i^p},x_{-i^p})\big). \tag{B.74}$$

Now by Axiom 2.3, there exists $\alpha^p\in(0,1)$ such that for all $\alpha\in(0,\alpha^p)$,

$$\alpha\big(p,(c,x_{-i^p})\big)+(1-\alpha)\big(p,(x^4_{i^p},x_{-i^p})\big)\succ\big(p,(x^3_{i^p},x_{-i^p})\big) \tag{B.75}$$

and

$$\big(p,(x^2_{i^p},x_{-i^p})\big)\succ\alpha\big(p,(c,x_{-i^p})\big)+(1-\alpha)\big(p,(x^1_{i^p},x_{-i^p})\big). \tag{B.76}$$

Secondly, let $i^q\in\operatorname{supp}q$. Denote $c_{i^q}=c$. Now since X is open and $(p,x)\sim(q,y)$ by (B.74) and Axiom 2.2, there exist $y^1_{i^q},y^2_{i^q}\in X$ such that for all $i\in\operatorname{supp}q$, $y^2_{i^q}\neq y_i\neq y^1_{i^q}$, $y^2_{i^q}>y_{i^q}>y^1_{i^q}$ and

$$\big(p,(x^3_{i^p},x_{-i^p})\big)\succ\big(q,(y^4_{i^q},y_{-i^q})\succ\big(q,(y^3_{i^q},y_{-i^q})\big)\succ(q,y) \tag{B.77}$$

and

$$(q,y)\succ\big(q,(y^2_{i^q},y_{-i^q})\big)\succ\big(q,(y^1_{i^q},y_{-i^q})\big)\succ\big(p,(x^2_{i^p},x_{-i^p})\big). \tag{B.78}$$

Now by Axiom 2.3, there exists $\alpha^q\in(0,1)$ such that for all $\alpha\in(0,\alpha^q)$,

$$\big(q,(y^4_{i^q},y_{-i^q})\big)\succ\alpha\Big(q,\big(c_{i^q},y_{-i^q}\big)\Big)+(1-\alpha)\big(q,(y^3_{i^q},y_{-i^q})\big)\succ(q,y) \tag{B.79}$$

and

$$(q,y)\succ\alpha\Big(q,\big(c_{i^q},y_{-i^q}\big)\Big)+(1-\alpha)\big(q,(y^2_{i^q},y_{-i^q})\big)\succ\big(q,(y^1_{i^q},y_{-i^q})\big). \tag{B.80}$$

Now there exist $\hat\alpha^p\in(0,\alpha^p),\hat\alpha^q\in(0,\alpha^q)$ such that $\hat\alpha^p p_{i^p}=\hat\alpha^q q_{i^q}$.

Denote $i^{p^*}=\max\{i\in\operatorname{supp}p\}+1$ and $\hat p=\big((p_i)_i\in\operatorname{supp}p\setminus\{i^p\},(1-\hat\alpha^p)p_{i^p},(\hat\alpha^p p_{i^p})_{i^{p^*}}\big)$ where $(\hat\alpha^p p_{i^p})_{i^{p^*}}$ denotes $\operatorname{pr}_{i^{p^*}}\hat p=\hat\alpha^p p_{i^p}$. Similarly denote $i^{q^*}=\max\{i\in\operatorname{supp}q\}+1$ and $\hat q=\big((q_i)_i\in\operatorname{supp}q\setminus\{i^q\},(1-\hat\alpha^q)q_{i^q},(\hat\alpha^q q_{i^q})_{i^{q^*}}\big)$. Now using this notation for all $a\in X$, $\tilde x\in\operatorname{NDiag}(X^p)$ such that for all $i\in\operatorname{supp}p$, $\tilde x_i\neq a$ and

$$\hat\alpha^p\big(p,(a_{i^p},\tilde x_{-i^p})\big)+(1-\hat\alpha^p)(p,\tilde x)=\big(\hat p,(a_{i^{p*}},\tilde x_{-i^{p^*}})\big)$$

and for all $a\in X$, $\tilde y\in\operatorname{NDiag}(X^q)$ such that for all $i\in\operatorname{supp}q$, $\tilde y_i\neq a$ and

$$\hat\alpha^q\big(q,(a_{i^q},\tilde y_{-i^q})\big)+(1-\hat\alpha^q)(q,\tilde y)=\big(\hat q,(a_{i^{q*}},\tilde y_{-i^{q^*}})\big).$$

So by (B.75,B.77,B.79),

$$\big(p,(x^4_{i^p},x_{-i^p})\big)\succ\big(\hat p,(c_{i^{p*}},x^4_{i^p},x_{-(i^{p^*},i^p)})\big)\succ\big(q,(y^4_{i^q},y_{-i^q})\big)\succ\big(\hat q,(c_{i^{q*}},y^3_{i^q},y_{-(i^{q^*},i^q)})\big)\succ(p,x) \tag{B.81}$$

and by (B.76,B.78,B.80),

$$(p,x)\succ\big(\hat q,(c_{i^{q*}},y^2_{i^q},y_{-(i^{q^*},i^q)})\big)\succ\big(q,(y^4_{i^q},y_{-i^q})\big)\succ\big(\hat p,(c_{i^{p*}},x^1_{i^p},x_{-(i^{p^*},i^p)})\big)\succ\big(p,(x^1_{i^p},x_{-i^p})\big). \tag{B.82}$$

Let $i^{p,2} \in \operatorname{supp} p, i^p \neq i^{p,2}$ and $i^{q,2} \in \operatorname{supp} q, i^q \neq i^{q,2}$. Since $|\operatorname{supp} p|, |\operatorname{supp} q| \geq 4$ by Axiom 2.2 and (B.81,B.82), there exist $\hat{x}^1, \hat{x}^2$ such that

$$(c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^1_{-(i^{p^*}, i^{p,2})}), (c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^2_{-(i^{p^*}, i^{p,2})}) \in \mathrm{NDiag}(X^{\hat{p}}),$$
$$\big(\hat{p}, (c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^2_{-(i^{p^*}, i^{p,2})})\big) \sim \big(\hat{q}, (c_{i^{q*}}, y^2_{i^q}, y_{-(i^{q^*}, i^q)})\big), \tag{B.83}$$

and

$$\big(\hat{p}, (c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^1_{-(i^{p^*}, i^{p,2})})\big) \sim \big(\hat{q}, (c_{i^{q*}}, y^1_{i^q}, y_{-(i^{q^*}, i^q)})\big). \tag{B.84}$$

Additionally, similarly by (B.81,B.82), there exist $\tilde{x}^1, \tilde{x}^2, \tilde{y}^1, \tilde{y}^2$ such that

$$(x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}}), (x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}}) \in \mathrm{NDiag}(X^p), (y_{i^{q,2}}, \tilde{y}^1_{-i^{q,2}}), (y_{i^{q,2}}, \tilde{y}^1_{-i^{q,2}}) \in \mathrm{NDiag}(X^q),$$
$$\big(p, (x_{i^{p,2}}, \tilde{x}^2_{-i^{p,2}})\big) \sim \big(q, (y_{i^{q,2}}, \tilde{y}^2_{-i^{q,2}})\big) \sim \big(\hat{q}, (c_{i^{q*}}, y^2_{i^q}, y_{-(i^{q^*}, i^q)})\big), \tag{B.85}$$

and

$$\big(p, (x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}})\big) \sim \big(q, (y_{i^{q,2}}, \tilde{y}^1_{-i^{q,2}})\big) \sim \big(\hat{q}, (c_{i^{q*}}, y^1_{i^q}, y_{-(i^{q^*}, i^q)})\big). \tag{B.86}$$

Now especially, we have

$$(x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}})_{i^{p,2}} = (c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^1_{-(i^{p^*}, i^{p,2})})_{i^{p,2}} = (x_{i^{p,2}}, \tilde{x}^2_{-i^{p,2}})_{i^{p,2}} = (c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^2_{-(i^{p^*}, i^{p,2})})_{i^{p,2}}$$

with $\hat{p}_{i^{p,2}} = p_{i^{p,2}}$,

$$(c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^1_{-(i^{p^*}, i^{p,2})})_{i^{p^*}} = (c_{i^{q*}}, y^1_{i^q}, y_{-(i^{q^*}, i^q)})_{i^{q^*}} = (c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^2_{-(i^{p^*}, i^{p,2})})_{i^{p^*}} = (c_{i^{q*}}, y^2_{i^q}, y_{-(i^{q^*}, i^q)})_{i^{q^*}}$$

with $\hat{p}^*_{i^{p^*}} = \hat{\alpha}^p p_{i^p} = \hat{\alpha}^q q_{i^q} = \hat{q}^*_{i^{q^*}}$, and

$$(y_{i^{q,2}}, \tilde{y}^1_{-i^{q,2}})_{i^{q,2}} = (c_{i^{q*}}, y^1_{i^q}, y_{-(i^{q^*}, i^q)})_{i^{q,2}} = (y_{i^{q,2}}, \tilde{y}^2_{-i^{q,2}})_{i^{q,2}} = (c_{i^{q*}}, y^2_{i^q}, y_{-(i^{q^*}, i^q)})_{i^{q,2}}$$

with $\hat{q}_{i^{q,2}} = q_{i^{q,2}}$.

By Lemma B.63, for all $z^p \in \mathrm{NDiag}(X^p), z^{\hat{p}} \in \mathrm{NDiag}(X^{\hat{p}}), z^q \in \mathrm{NDiag}(X^{\hat{q}}), z^q \in \mathrm{NDiag}(X^q)$ such that

$$\big(p, (x_{i^{p,2}}, \tilde{x}^2_{-i^{p,2}})\big) \succsim (p, z^p), (\hat{p}, z^{\hat{p}}), (\hat{q}, z^{\hat{q}}), (q, z^q) \succsim \big(p, (x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}})\big),$$

we have

$$z^p \succsim z^{\hat{p}} \iff$$
$$\sum_{i \in \operatorname{supp} p} p^\lambda_i u(z^p_i) - \sum_{i \in \operatorname{supp} p} p^\lambda_i u\big((x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}})_i\big) \geq \sum_{i \in \operatorname{supp} \hat{p}} \hat{p}^\lambda_i u(z^{\hat{p}}_i) - \sum_{i \in \operatorname{supp} \hat{p}} \hat{p}^\lambda_i u\big((c_{i^{p*}}, x_{i^{p,2}}, \hat{x}^1_{-(i^{p^*}, i^{p,2})})_i\big)$$

with

$$z^{\hat{p}} \succsim z^p \iff$$

$$\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u(z_i^{\hat{p}})-\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u\big((c_{i^{p*}},x_{i^{p,2}},\hat{x}^1_{-(i^{p^*},i^{p,2})})_i\big)\geq\sum_{i\in\operatorname{supp}p}p_i^{\lambda}u(z_i^{p})-\sum_{i\in\operatorname{supp}p}p_i^{\lambda}u\big((x_{i^{p,2}},\tilde{x}^1_{-i^{p,2}})_i\big),$$

$$z^{\hat{p}}\succsim z^{\hat{q}}\iff$$

$$\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u(z_i^{\hat{p}})-\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u\big((c_{i^{p*}},x_{i^{p,2}},\hat{x}^1_{-(i^{p^*},i^{p,2})})_i\big)\geq\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u(z_i^{\hat{q}})-\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u\big((c_{i^{q*}},y_{i^{q,2}},\hat{y}^1_{-(i^{q^*},i^{q,2})})_i\big)$$

with

$$z^{\hat{q}}\succsim z^{\hat{p}}\iff$$

$$\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u(z_i^{\hat{q}})-\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u\big((c_{i^{q*}},y_{i^{q,2}},\hat{y}^1_{-(i^{q^*},i^{q,2})})_i\big)\geq\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u(z_i^{\hat{p}})-\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u\big((c_{i^{p*}},x_{i^{p,2}},\hat{x}^1_{-(i^{p^*},i^{p,2})})_i\big)$$

and

$$z^{\hat{q}}\succsim z^{q}\iff$$

$$\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u(z_i^{\hat{q}})-\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u\big((c_{i^{q*}},x_{i^{q,2}},\hat{x}^1_{-(i^{q^*},i^{q,2})})_i\big)\geq\sum_{i\in\operatorname{supp}q}q_i^{\lambda}u(z_i^{q})-\sum_{i\in\operatorname{supp}q}q_i^{\lambda}u\big((y_{i^{q,2}},\tilde{y}^1_{-i^{q,2}})_i\big)$$

with

$$z^{q}\succsim z^{\hat{q}}\iff$$

$$\sum_{i\in\operatorname{supp}q}q_i^{\lambda}u(z_i^{q})-\sum_{i\in\operatorname{supp}q}q_i^{\lambda}u\big((y_{i^{q,2}},\tilde{y}^1_{-i^{q,2}})_i\big)\geq\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u(z_i^{\hat{q}})-\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u\big((c_{i^{q*}},x_{i^{q,2}},\hat{x}^1_{-(i^{q^*},i^{q,2})})_i\big).$$

Finally, let $z^p\in\operatorname{NDiag}(X^p)$, $z^q\in\operatorname{NDiag}(X^q)$ be such that

$$\big(p,(x_{i^{p,2}},\tilde{x}^2_{-i^{p,2}})\big)\succsim(p,z^p),(q,z^q)\succsim\big(p,(x_{i^{p,2}},\tilde{x}^1_{-i^{p,2}})\big).$$

Now by (B.83-B.86), there exist $z^{\hat{p},p},z^{\hat{p},q}\in\operatorname{NDiag}(X^{\hat{p}})$ and $z^{\hat{q},p},z^{\hat{q},q}\in\operatorname{NDiag}(X^{\hat{q}})$ such that

$$(p,z^p)\sim(\hat{p},z^{\hat{p},p})\text{ and }(p,z^q)\sim(\hat{q},z^{\hat{q},q}).\tag{B.87}$$

Now by above, we have

$$\sum_{i\in\operatorname{supp}p}p_i^{\lambda}u(z_i^{p})-\sum_{i\in\operatorname{supp}p}p_i^{\lambda}u\big((x_{i^{p,2}},\tilde{x}^1_{-i^{p,2}})_i\big)=\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u(z_i^{\hat{p},p})-\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u\big((c_{i^{p*}},x_{i^{p,2}},\hat{x}^1_{-(i^{p^*},i^{p,2})})_i\big)\tag{B.88}$$

and

$$\sum_{i\in\operatorname{supp}q}q_i^{\lambda}u(z_i^{q})-\sum_{i\in\operatorname{supp}q}q_i^{\lambda}u\big((y_{i^{q,2}},\tilde{y}^1_{-i^{q,2}})_i\big)=\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u(z_i^{\hat{q},q})-\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u\big((c_{i^{q*}},x_{i^{q,2}},\hat{x}^1_{-(i^{q^*},i^{q,2})})_i\big).\tag{B.89}$$

So

$$(p,z^p)\succsim(q,z^q)\qquad\overset{(\text{B.87})}{\iff}\qquad(\hat{p},z^{\hat{p},p})\succsim(\hat{p},z^{\hat{p},p})$$

$$\iff\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u(z_i^{\hat{p},p})-\sum_{i\in\operatorname{supp}\hat{p}}\hat{p}_i^{\lambda}u\big((c_{i^{p*}},x_{i^{p,2}},\hat{x}^1_{-(i^{p^*},i^{p,2})})_i\big)$$

$$\geq\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u(z_i^{\hat{q},q})-\sum_{i\in\operatorname{supp}\hat{q}}\hat{q}_i^{\lambda}u\big((c_{i^{q*}},y_{i^{q,2}},\hat{y}^1_{-(i^{q^*},i^{q,2})})_i\big)$$

$$\overset{\text{(B.88,B.89)}}{\iff} \sum_{i\in\operatorname{supp} p} p_i^\lambda u(z_i^p) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}})_i\big) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(z_i^q) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u\big((y_{i^{q,2}}, \tilde{y}^1_{-i^{q,2}})_i\big)$$

and symmetrically

$$(q, z^q) \succsim (p, z^p) \iff$$

$$\sum_{i\in\operatorname{supp} q} q_i^\lambda u(z_i^q) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u\big((y_{i^{q,2}}, \tilde{y}^1_{-i^{q,2}})_i\big) \geq \sum_{i\in\operatorname{supp} p} p_i^\lambda u(z_i^p) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}})_i\big).$$

Finally, since $(p, x) \sim (q, y)$, we have

$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_{i^{p,2}}, \tilde{x}^1_{-i^{p,2}})_i\big) = \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) - \sum_{i\in\operatorname{supp} q} q_i^\lambda u\big((y_{i^{q,2}}, \tilde{y}^1_{-i^{q,2}})_i\big).$$

So we can subtract it from the both sides in above equivalencies for the final claim. □

B.3.2 Transfinite Induction

Next steps extend the local representations to a global representation using transfinite induction. We will do induction on connected preference intervals starting from a uniform lottery with 4 prizes. This is done in several steps. First, we show that there is a unique way to extend a representation using Axiom 2.12. Second, we show the first step and establish the representation for a preference interval that contains a uniform lottery with 4 prizes. Third, we show that any chain of extensions has an upper bound extension. Fourth, we show that any non maximal-representation can be extended further. These steps will establish the desired representation.

Before this, we show that for a given co-probability, if a representation for $\succsim^p$ is locally power expected utility representation, then it is globally a power expected utility representation for $\succsim^p$ for a connected preference interval.

Lemma B.65 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i\in\operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12. Let $p \in \Delta^4(\mathbb{N})$. Assume that there exists a function I such that for all $x, y \in \operatorname{NDiag}(X^p)$, if $(p, x), (p, y) \in \operatorname{dom} I$, then

$$(p, x) \succsim (p, y) \iff I(p, x) \geq I(p, y),$$

for all $x \in \operatorname{NDiag}(X^p)$, if $(p, x) \in \operatorname{dom} I$, then there exist $x^1, x^2 \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, x) \succ (p, x^2)$ and for all $z \in \operatorname{NDiag}(X^p)$, if $(p, z) \in \operatorname{dom} I$ and $(p, x^1) \succ (p, z) \succ (p, x^2)$, then

$$I(p, z) - I(p, x) = \sum_{i\in\operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i), \tag{B.90}$$

and there exist $P, Q \in \Delta(X)$ such that if $x \in \mathrm{NDiag}(X^p)$ is such that $P \succ (p,x) \succ Q$, then

$$(p,x) \in \operatorname{dom} I. \tag{B.91}$$

If $x \in \mathrm{NDiag}(X^p)$ is such that $P \succ (p,x) \succ Q$, then for all $z \in \mathrm{NDiag}(X^p)$ such that $P \succ (p,z) \succ Q$,

$$I(p,z) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + I(p,x).$$

Proof. Without loss of generality assume that $1 \in \operatorname{supp} p$ by permuting the indices. Let $P, Q \in \Delta(X)$ be as above and $x^0 \in \mathrm{NDiag}(X^p)$ such that $P \succ (p,x^0) \succ Q$. Define $\widetilde{\succsim}^p$ by for all $x, y \in \mathrm{NDiag}(X^p)$ such that $P \succ (p,x), (p,y) \succ Q$,

$$x \mathrel{\widetilde{\succsim}^p} y \iff (p,x) \succsim (p,y).$$

By an assumption, $(p_i^\lambda u)_{i \in \operatorname{supp} p}$ is an additive representation for $\widetilde{\succsim}^p$. Additionally, by (B.91), I is another representation for $\widetilde{\succsim}^p$. Thus by the uniqueness for $\widetilde{\succsim}^p$, there exists a strictly increasing function φ such that for all $x \in \mathrm{NDiag}(X^p)$ such that $P \succ (p,x) \succ Q$,

$$I(p,x) = \varphi\Big(\sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) \Big).$$

We show that φ is differentiable with derivative of 1. Let $x \in \mathrm{NDiag}(X^p)$ be such that $P \succ (p,x) \succ Q$. By (B.90), there exist $x^1, x^2 \in \mathrm{NDiag}(X^p)$ such that $(p,x^1) \succ (p,x) \succ (p,x^2)$ and for all $z \in \mathrm{NDiag}(X^p)$, if $(p,z) \in \operatorname{dom} I$ and $(p,x^1) \succ (p,z) \succ (p,x^2)$, then

$$I(p,z) - I(p,x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i). \tag{B.92}$$

By Axiom 2.2, assume w.l.o.g. that $P \succ (p,x^1) \succ (p,x^2) \succ Q$.

Denote $\theta = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i)$. Let $(\alpha_j)_{i=1}^\infty \subseteq \mathbb{R}$ be such that $\alpha_j \to \theta$ as $j \to \infty$. Since X is open by Axiom 2.2 and the representation, let $\tilde{x}^*, \tilde{x}^\dagger \in X$ be such that $\tilde{x}^* > x_1 > \tilde{x}^\dagger$ and

$$(p,x^1) \succ \big(p,(\tilde{x}^*, x_{-1})\big) \succ \big(p,(\tilde{x}^\dagger, x_{-1})\big) \succ (p,x^2).$$

$$x_1^* = \min\{x_i | i \in \operatorname{supp} p \setminus \{1\}, x_i > x_1\} \cup \{\tilde{x}^2\} \;\&\; x_1^\dagger = \max\{x_i | i \in \operatorname{supp} p \setminus \{1\}, x_i < x_1\} \cup \{\tilde{x}^1\}$$

that exist since $\operatorname{supp} p$ is finite. Denote

$$\beta = p_1^\lambda \min\Big\{ |u(x_1) - u(x_1^*)|, \Big|u(x_1) - u(x_1^\dagger)\Big| \Big\}.$$

Since $\lim_{j \to \infty} \alpha_j = \theta$, there exists j^0 such that for all $j > j^0$,

$$|\alpha_j - \theta| < \beta.$$

Thus for all $j > j^0$, since $u(x_1^*) > u(x_1) > u(x_1^\dagger)$,

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^*, x_{-1})_i\big) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + p_1^\lambda |u(x_1^*) - u(x_1)|$$

$$\geq \theta + \beta > \alpha_j > \theta - \beta \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) - p_1^\lambda \Big| u(x_1) - u(x_1^\dagger) \Big| = \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger, x_{-1})_i\big).$$

Thus by the continuity of u for all $j > j^0$, there exists $x_1^j \in X$, $x_1^* > x_1^j > x_1^\dagger$ such that $\sum_{i \in \operatorname{supp} p} p_j^\lambda u\big((x_1^j, x_{-1})_i\big) = \alpha_j$. Now by the choice of $x_1^*, x_1^\dagger$, $(x_1^j, x_{-1}) \in \operatorname{NDiag}(X^p)$. Thus we have for all $j > j^0$,

$$\alpha_j - \theta = \sum_{i \in \operatorname{supp} p} p_j^\lambda u\big((x_1^j, x_{-1})_i\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) = p_1^\lambda \big(u(x_1^j) - u(x_1)\big).$$

Additionally,

$$P \succ (p, x^1) \succ \big(p, (\tilde{x}^*, x_{-1})\big) \succ \big(p, (x_1^j, x_{-1})\big) \succ \big(p, (\tilde{x}^\dagger, x_{-1})\big) \succ (p, x^2) \succ Q$$

and so $\alpha_j \in \operatorname{dom} \varphi$ and by (B.92)

$$\varphi(\alpha_j) - \varphi(\theta) = I\big(p, (x_1^j, x_{-1})\big) - I(p, x) = p_1^\lambda \big(u(x_1^j) - u(x_1)\big) = \alpha_j - \theta.$$

Thus

$$\lim_{j \to \infty} \frac{\varphi(\alpha_j) - \varphi(\theta)}{\alpha_j - \theta} = \lim_{j \to \infty} \frac{\alpha_j - \theta}{\alpha_j - \theta} = 1.$$

So φ is differentiable with derivative of 1 in its domain. Let $x^0 \in \operatorname{NDiag}(X^p)$ such that $P \succ (p, x^0) \succ Q$ and denote $\theta^0 = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^0)$. Let $y \in \operatorname{NDiag}(X^p)$ be such that $P \succ (p, y) \succ Q$ and denote $\theta^y = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(y_i)$. Now for all $a \in [\theta^0, \theta^y] \cup [\theta^y, \theta^0]$, by the continuity of u, there exists $z \in \operatorname{NDiag}(X^p)$ such that

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) = a.$$

Then by the additive representation for $\succsim^p$,

$$P \succ (p, z) \succ Q$$

and so $a \in \operatorname{dom} \varphi$. Thus $[\theta^0, \theta^y] \cup [\theta^y, \theta^0] \subseteq \operatorname{dom} \varphi$. Since φ is differentiable, we have by the second fundamental theorem of calculus,

$$I(p, y) = \varphi(\theta^y) = \int_{\theta^0}^{\theta^y} \varphi'(z)\, dz + \varphi(\theta^0) = \int_{\theta^0}^{\theta^y} 1\, dz + \varphi(\theta^0) = \theta^y - \theta^0 + \varphi(\theta^0)$$
$$= \sum_{i \in \operatorname{supp} p} p_i^\lambda u(y_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^0) + I(p, x^0).$$

□

The next lemma shows that there is a unique way to do the extension of representation.

Lemma B.66 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i \in \operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12. Let I^0 be such that for all $P \in \operatorname{Dist}(\Delta^{(\frac{1}{4})_{i=1}^4})$,

$$I^0(P) = \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x).$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \Delta^4(X)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $P \in \Delta^4(X)$, if there exist $\tilde{P}, \tilde{Q} \in \operatorname{dom} \widetilde{I}$ such that $\tilde{P} \succsim P \succsim \tilde{Q}$, then $P \in \operatorname{dom} \widetilde{I}$.
3. $\operatorname{dom} I^0 \subseteq \operatorname{dom} \widetilde{I}$ and for all $P \in \operatorname{dom} I^0$, $I^0(P) = \widetilde{I}(P)$.
4. For all $p \in \Delta^4(\mathbb{N}), x \in \operatorname{NDiag}(X^p)$, if $(p, x) \in \operatorname{dom} \widetilde{I}$, then there exist $x^1, x^2 \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, x) \succ (p, x^2)$ and for all $z \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, z) \succ (p, x^2)$ and $(p, z) \in \operatorname{dom} \widetilde{I}$,
$$\widetilde{I}(p, z) - \widetilde{I}(p, x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$
5. for all $P, Q \in \operatorname{dom} \widetilde{I}$, $P \succsim Q$ iff $\widetilde{I}(P) \geq \widetilde{I}(Q)$.

Let $A_1, A_2 \in \mathcal{A}$, $A_1 \subseteq A_2$ and I_1, I_2 be the associated functions. Then $\operatorname{dom} I_1 \subseteq \operatorname{dom} I_2$ and for all $P \in \operatorname{dom} I_1$, $I_1(P) = I_2(P)$.

Proof. We show the claim for all $P \in \Delta(X)$ such that $I_1(P) \geq I_1(P^0)$, since the values below follow symmetrically. Denote $p^0 = (\frac{1}{4})_{i=1}^4$, let $x^0 \in \operatorname{NDiag}(X^{p^0})$. Denote

$$a = \inf\{b \in \mathbb{R}_+ | \exists P \in \operatorname{dom} I_1, I_1(P) - I^0(p^0, x^0) = b, [P \notin \operatorname{dom} I_2 \text{ or } I_1(P) \neq I_2(P)]\}. \tag{B.93}$$

We show that $a = \infty$ that shows the claim. Assume first that $a = 0$. Since X is open, there exist $x_1^1, x_1^2 \in X$ such that $x_1^2 > x_1^0 > x_1^1$, and for all $i \in \operatorname{supp} p^0$,

$$\left|x_1^1 - x_1^0\right|, \left|x_1^2 - x_1^0\right| < \left|x_i^2 - x_1^0\right|. \tag{B.94}$$

Now by Lemma B.31, $(p^0, (x_1^2, x_{-1})) \succ (p^0, x) \succ (p^0, (x_1^1))$. Second, by (B.94) for all $x_1^2 \geq \tilde{x}_1 \geq x_1^1$, $(\tilde{x}_1, x_{-1}) \in \operatorname{NDiag}(X^{p^0})$. Let $P \in \Delta^4(X)$ be such that $(p^0, (x_1^2, x_{-1})) \succsim P \succsim (p^0, (x_1^1, x_{-1}))$. By Axiom 2.2, there exists $x_1^P \in X$ such that $x_1^2 \geq x_1^P \geq x_1^1$ and $(p^0, (x_1^1, x_{-1})) \sim P$. Now by Conditions 2 and 3, $P \in \operatorname{dom} I_1, \operatorname{dom} I_2$. Second, by Conditions 3 and 5

$$I_1(P) = I_1(p^0, (x_1^1, x_{-1})) = I^0(p^0, (x_1^1, x_{-1})) = I_2(p^0, (x_1^1, x_{-1})) = I_2(P).$$

Finally, by Condition 5, if $P \in \operatorname{dom} I_1$ and $P \succ \big(p^0, (x_1^2, x_{-1})\big)$, then $I_1(P) > I_1\big(p^0, (x_1^2, x_{-1})\big)$ and $P \in \operatorname{dom} I_1$ and $\big(p^0, (x_1^1, x_{-1})\big) \succ P$, then $I_1\big(p^0, (x_1^2, x_{-1})\big) > I_1(P)$. Thus

$$a \geq I_1\big(p^0, (x_1^2, x_{-1})\big) > 0.$$

Next assume, per contra, that $a < \infty$. By Condition 1, there exist $p \in \Delta^4(\mathbb{N}), x \in \operatorname{NDiag}(X^p)$ such that $(p, x) \in \operatorname{dom} I_1$ such that $I_1(p, x) = a + I^0(p^0, x^0)$. By permuting the indices of p, assume w.l.o.g. that $1 \in \operatorname{supp} p$.

Now $I_1(p, x) \in A_1$. Since A_1 is open by Condition 1, there exist $Q, R \in \operatorname{dom} I_1$ such that $I_1(Q) \geq I_1(p, x) \geq I_1(R)$. Since $A_2 \supseteq A_1$, there exists $S \in \operatorname{dom} I_2$ such that $I_2(S) > I_1(p, x)$. First, we will show that $S \succsim (p, x)$.

By Lemma B.31 and Axiom 2.2 and since X is open, there exists $x_1^* \in X$ such that $x_1^* < x_1$, for all $\tilde{x}_1 \in [x_1^*, x_1])$, $(\tilde{x}_1, x_{-1}) \in \operatorname{NDiag}(X^p)$, and $(p, x) \succ \big(p, (x_1^*, x_{-1})\big) \succ P^0$. Now for all $\tilde{x}_1 \in [x_1^*, x_1])$, we have by Condition 2, $\big(p, (\tilde{x}_1, x_{-1})\big) \in \operatorname{dom} I_1$, and by Condition 5,

$$I_1\big(p, (\tilde{x}_1, x_{-1})\big) < I_1(p, x) = a.$$

Next, we have for all $\tilde{x}_1 \in [x_1^*, x_1)$, by (B.93), $\big(p, (\tilde{x}_1, x_{-1})\big) \in \operatorname{dom} I_2$ and

$$I_2\big(p, (\tilde{x}_1, x_{-1})\big) = I_1\big(p, (\tilde{x}_1, x_{-1})\big).$$

So especially by Condition 5, for all $\tilde{x}_1 \in [x_1^*, x_1)$,

$$S \succsim \big(p, (\tilde{x}_1, x_{-1})\big).$$

So by Axiom 2.2,

$$S \succsim (p, x).$$

So by Conditions 2 and 5,

$$(p, x) \in \operatorname{dom} I_2 \text{ and } S \succ (p, x).$$

Next from Condition 4 for both I_1 and I_2, there exist $x^1, x^2 \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, x) \succ (p, x^2)$ and for all $z \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, z) \succ (p, x^2)$, if $(p, z) \in \operatorname{dom} I_1$, then

$$I_1(p, z) - I_1(p, x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i)$$

and if $(p, z) \in \operatorname{dom} I_2$, then

$$I_2(p, z) - I_2(p, x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$

By Axiom 2.2 and since X is open, there exist $x_1^2, x_1^1 \in X$ such that $x_1^1 < x_1 < x_1^2$, for all $\tilde{x}_1 \in [x_1^1, x_1^2]$,

$$(\tilde{x}_1, x_{-1}) \in \mathrm{NDiag}(X^p), \tag{B.95}$$

$$(p, x^2), S, R \succ (p, (x_1^2, x_{-1})), \text{ and } (p, (x_1^1, x_{-1})) \succ (p, x^1), P^0.$$

By Condition 2, for all $\tilde{x}_1 \in [x_1^1, x_1^2]$,

$$\big(p, (\tilde{x}_1, x_{-1})\big) \in \mathrm{dom}\, I_1, \mathrm{dom}\, I_2.$$

So by the choice of x^1, x^2, we have for all $\tilde{x}_1 \in [x_1^1, x_1^2]$,

$$I_1\big(p, (\tilde{x}_1, x_{-1})\big) - I_1(p, x) = I_2\big(p, (\tilde{x}_1, x_{-1})\big) - I_2(p, x) = \sum_{i \in \mathrm{supp}\, p} p_i^\lambda u(z_i) - \sum_{i \in \mathrm{supp}\, p} p_i^\lambda u(x_i). \tag{B.96}$$

Especially, we have by (B.93)

$$I_1\big(p, (x_1^1, x_{-1})\big) = I_2\big(p, (x_1^1, x_{-1})\big) \overset{\text{(B.96)}}{\Longrightarrow} I_1(p, x) = I_2(p, x)$$

and so for all $\tilde{x}_1 \in [x_1^1, x_1^2]$,

$$I_1\big(p, (\tilde{x}_1, x_{-1})\big) = I_2\big(p, (\tilde{x}_1, x_{-1})\big). \tag{B.97}$$

Now if $\big(p, (x_1^2, x_{-1})\big) \succsim Q \succsim \big(p, (x_1^1, x_{-1})\big)$, then by Axiom 2.2 and (B.95), there exists $x_1^Q \in [x_1^1, x_2^2]$ such that

$$\big(p, (x_1^Q, x_{-1})\big) \sim Q$$

and so by Conditions 2 and 5, $Q \in \mathrm{dom}\, I_1, \mathrm{dom}\, I_2$ and

$$I_1(Q) = I_1\big(p, (x_1^Q, x_{-1})\big) \overset{\text{(B.97)}}{=} I_2\big(p, (x_1^Q, x_{-1})\big) = I_2(Q).$$

Finally, by Condition 5, for all $Q^1, Q^2 \in \mathrm{dom}\, I_1$ such that

$$Q^2 \succ \big(p, (x_1^2, x_{-1})\big), \big(p, (x_1^1, x_{-1})\big) \succ Q^1,$$

we have

$$I_1(Q^2) > I_1\big(p, (x_1^2, x_{-1})\big) > I_1\big(p, (x_1^1, x_{-1})\big) > I_1(Q^1).$$

But this contradicts the counterassumption that $a < \infty$ by (B.93). Thus $a = \infty$ and so for all $P \in \mathrm{dom}\, I_1$ such that $I_1(P) \geq I_1(P^0)$, we have $P \in \mathrm{dom}\, I_2$ and $I_1(P) = I_2(P)$. The other case of $P \in \mathrm{dom}\, I_1$ and $I_1(P) \leq I_1(P^0)$ goes symmetrically. □

The next lemma shows the first step of induction and shows that on the preference interval given by uniform lotteries with 4 prizes, there exists the desired representation.

Lemma B.67 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i \in \operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12. Let I^0 be such that for all $P \in \operatorname{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$,

$$I^0(P) = \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x). \tag{B.98}$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \Delta^4(X)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $P \in \Delta^4(X)$, if there exist $\tilde{P}, \tilde{Q} \in \operatorname{dom} \widetilde{I}$ such that $\tilde{P} \succsim P \succsim \tilde{Q}$, then $P \in \operatorname{dom} \widetilde{I}$.
3. $\operatorname{dom} I^0 \subseteq \operatorname{dom} \widetilde{I}$ and for all $P \in \operatorname{dom} I^0$, $I^0(P) = \widetilde{I}(P)$.
4. For all $p \in \Delta^4(\mathbb{N}), x \in \operatorname{NDiag}(X^p)$, if $(p, x) \in \operatorname{dom} \widetilde{I}$, then there exist $x^1, x^2 \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, x) \succ (p, x^2)$ and for all $z \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, z) \succ (p, x^2)$ and $(p, z) \in \operatorname{dom} \widetilde{I}$,
$$\widetilde{I}(p, z) - \widetilde{I}(p, x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$
5. for all $P, Q \in \operatorname{dom} \widetilde{I}$, $P \succsim Q$ iff $\widetilde{I}(P) \geq \widetilde{I}(Q)$.

Then $\operatorname{Im} I^0 \in \mathcal{A}$.

Proof. First, since X is open and connected, and u is continuous $\operatorname{Im} I^*$ is an open interval. We define a function I^* on all $P \in \Delta^4(X)$ such that there exists $Q \in \operatorname{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$, $P \sim Q$,

$$I^*(P) = I^0(Q). \tag{B.99}$$

We show that $\operatorname{Im} I^0$ with the function I^* satisfies Conditions 1-5.

Condition 1: By the definition of I^*, $\operatorname{Im} I^* = \operatorname{Im} I^0$.

Condition 2: Let $P \in \Delta^4(X), \tilde{P}, \tilde{Q} \in \operatorname{dom} \widetilde{I}$ be such that $\tilde{P} \succsim P \succsim \tilde{Q}$. Now there exist $\tilde{P}^1, \tilde{Q}^1 \in \operatorname{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$ such that

$$\tilde{P} \sim \tilde{P}^1, \tilde{Q} \sim \tilde{Q}^1, I^*(\tilde{P}) = I^0(\tilde{P}^1), \text{ and } I^*(\tilde{Q}) = I^0(\tilde{Q}^1)$$

Now $\tilde{P}^1 \succsim P \succsim \tilde{Q}^1$. Since X is connected, by Axiom 2.2 and Lemma B.36, there exists $P^1 \in \operatorname{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$ such that $P \sim P^1$. Thus by the definition of I^*, $P \in \operatorname{dom} I^*$.

Condition 3: Follows trivially.

Condition 4: Let $p \in \Delta^4(\mathbb{N}), x \in \operatorname{NDiag}(X^p)$, $(p, x) \in \operatorname{dom} \widetilde{I}$, and $Q \in \operatorname{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$ be such that $(p, x) \sim Q$. By Lemma B.64, there exist $Q^1, Q^2 \in \operatorname{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}}), x^1, x^2 \in \operatorname{NDiag}(X^p)$ such that

$$(p, x^2) \sim Q^2 \succ (p, x) \sim Q \succ (p, x^1) \sim Q^1$$

and for all $\tilde{x} \in \mathrm{NDiag}(X^p), \tilde{y} \in \mathrm{NDiag}(X^{(\frac{1}{4})^4_{i=1}})$, $(p, x^2) \succsim (p, \tilde{x}), \big((\frac{1}{4})^4_{i=1}, \tilde{y}\big) \succsim (p, x^1)$

$$(p, \tilde{x}) \succsim \big((1/4)^4_{i=1}, \tilde{y}\big) \iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(\tilde{x}_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) \geq \sum_{i=1}^4 4^{-\lambda} u(\tilde{y}_i) - \sum_{i=1}^4 4^{-\lambda} u(y_i)$$

and

$$\big((1/4)^4_{i=1}, \tilde{y}\big) \succsim (p, \tilde{x}) \iff \sum_{i=1}^4 4^{-\lambda} u(\tilde{y}_i) - \sum_{i=1}^4 4^{-\lambda} u(y_i) \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u(\tilde{x}_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$

Let $z \in \mathrm{NDiag}(X^p)$ be such that $(p, x^2) \succ (p, z) \succ (p, x^1)$. Now $Q^2 \succ (p, z) \succ Q^1$ and so by Axiom 2.2 and Lemma B.36, there exists $\hat{y} \in \mathrm{NDiag}(X^{(\frac{1}{4})^4_{i=1}})$ such that $(p, z) \sim \big((1/4)^4_{i=1}, \hat{y}\big)$. So by above, we have

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u(z) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) = \sum_{i=1}^4 4^{-\lambda} u(\hat{y}_i) - \sum_{i=1}^4 4^{-\lambda} u(y_i).$$

Now we have by the definition of I^* and (B.98),

$$\sum_{i=1}^4 4^{-\lambda} u(\hat{y}_i) - \sum_{i=1}^4 4^{-\lambda} u(y_i) \overset{\text{(B.98)}}{=} I^0\big((1/4)^4_{i=1}, \hat{y}\big) - I^0(Q) \overset{\text{(B.99)}}{=} I^*(p, z) - I^*(p, x).$$

This shows the condition.

Condition 5: Let $P, Q \in \operatorname{dom} I^*$ and let $\tilde{P}, \tilde{Q} \in \mathrm{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$ be such that $\tilde{P} \sim P$, $\tilde{Q} \sim Q$. Now since I^0 represents $\succsim^{(\frac{1}{4})^4_{i=1}}$, we have

$$P \succsim Q \iff \tilde{P} \succsim \tilde{Q} \iff I^0(\tilde{P}) \geq I^0(\tilde{Q}) \overset{\text{(B.99)}}{\iff} I^*(P) \geq I^*(Q).$$

This shows that $\operatorname{Im} I^0 \in \mathcal{A}$. □

The next lemma shows that any chain of representation has an upper bound where the preference interval is the union of all the intervals in the chain.

Lemma B.68 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i \in \operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12. Let I^0 be such that for all $P \in \mathrm{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$,

$$I^0(P) = \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x).$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \Delta^4(X)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $P \in \Delta^4(X)$, if there exist $\tilde{P}, \tilde{Q} \in \operatorname{dom} \widetilde{I}$ such that $\tilde{P} \succsim P \succsim \tilde{Q}$, then $P \in \operatorname{dom} \widetilde{I}$.
3. $\operatorname{dom} I^0 \subseteq \operatorname{dom} \widetilde{I}$ and for all $P \in \operatorname{dom} I^0$, $I^0(P) = \widetilde{I}(P)$.

4. For all $p \in \Delta^4(\mathbb{N}), x \in \mathrm{NDiag}(X^p)$, if $(p,x) \in \operatorname{dom} \widetilde{I}$, then there exist $x^1, x^2 \in \mathrm{NDiag}(X^p)$ such that $(p,x^1) \succ (p,x) \succ (p,x^2)$ and for all $z \in \mathrm{NDiag}(X^p)$ such that $(p,x^1) \succ (p,z) \succ (p,x^2)$ and $(p,z) \in \operatorname{dom} \widetilde{I}$,
$$\widetilde{I}(p,z) - \widetilde{I}(p,x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$
5. for all $P, Q \in \operatorname{dom} \widetilde{I}$, $P \succsim Q$ iff $\widetilde{I}(P) \geq \widetilde{I}(Q)$.

Let $\mathcal{B} \subseteq \mathcal{A}$, $\mathcal{B} \neq \varnothing$, be such that for all $B_1, B_2 \in \mathcal{B}$, $B_1 \subseteq B_2$ or $B_2 \subseteq B_2$. Then
$$\bigcup_{B \in \mathcal{B}} B \in \mathcal{A}.$$

Proof. Denote $B^* = \cup_{B \in \mathcal{B}} B$. For each $B \in \mathcal{B}$, denote by I^B the associated function. Define I^* by for all $P \in \Delta^4(X)$ such that there exists $B \in \mathcal{B}$ with $P \in \operatorname{dom} I^B$, by
$$I^*(P) = I^B(P). \tag{B.100}$$
By Lemma B.66, I^* is well-defined. We show that B^* with the function I^* satisfies Conditions 1-5.

Condition 1: Let $a \in B^*$. Now there exists $B \in \mathcal{B}$ such that $a \in B$. So by Condition 1 for B, there exists $P \in \operatorname{dom} I^B$ such that $I^B(P) = a$. So by (B.100) $I^*(P) = I^B(P) = a$ and so $a \in \Im I^*$. So $\operatorname{Im} I^* \supseteq B^*$. Additionally, if $a \in \operatorname{Im} I^*$, then by (B.100), there exists $B \in \mathcal{B}$ such that $a \in \operatorname{Im} I^B = B \subseteq B^*$. Thus $\operatorname{Im} I^* \subseteq B^*$ and so $\operatorname{Im} I^* = B^*$.

Conditions 2-3: Follow trivially as below Condition 5.

Condition 4: Let $p \in \Delta^4(\mathbb{N}), x \in \mathrm{NDiag}(X^p), (p,x) \in \operatorname{dom} I^*$. Now there exists $B \in \mathcal{B}$ such that $I^*(p,x) = I^B(p,x)$. Since B is open, there exist $R, S \in \operatorname{dom} I^B$ such that
$$I^B(R) > I^B(p,x) > I^B(S).$$
By Condition 5, $R \succ (p,x) \succ S$. By Condition 4, there exist $x^1, x^2 \in \mathrm{NDiag}(X^p)$ such that $(p,x^1) \succ (p,x) \succ (p,x^2)$ and for all $z \in \mathrm{NDiag}(X^p)$ such that $(p,x^1) \succ (p,z) \succ (p,x^2)$ and $(p,z) \in \operatorname{dom} I^B$,
$$I^B(p,z) - I^B(p,x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i). \tag{B.101}$$
Now by Axiom 2.2, there exist $\tilde{x}^1, \tilde{x}^2 \in \mathrm{NDiag}(X^p)$ such that
$$R, (p,x^1) \succ (p,\tilde{x}^1) \succ (p,x) \succ (p,\tilde{x}^2) \succ (p,x^2), S.$$
So for all $z \in \mathrm{NDiag}(X^p)$ such that $(p,\tilde{x}^1) \succ (p,z) \succ (p,\tilde{x}^2)$, we have by Condition 2, $(p,z) \in \operatorname{dom} I^B$ and so by (B.100,B.101),
$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) = I^B(p,z) - I^B(p,x) = I^*(p,z) - I^*(p,x).$$

Condition 5: Let $P, Q \in \operatorname{dom} I^*$. Now there exist $B_1, B_2 \in \mathcal{B}$ such that $P \in \operatorname{dom} I^{B_1}$ and $Q \in \operatorname{dom} I^{B_2}$ and $B_1 \subseteq B_2$ or $B_2 \subseteq B_1$. Assume w.l.o.g. $B_1 \subseteq B_2$. By Lemma B.66, $P \in \operatorname{dom} I^{B_2}$. So by Condition 5 for B_2 and (B.100), we have

$$P \succsim Q \iff I^{B_2}(P) \geq I^{B_2}(Q) \iff I^*(P) \geq I^*(Q).$$

□

The next lemma shows that if an extension is not maximal and there exists a lottery that is not in the domain of the extension, then it can be extended further.

Lemma B.69 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i \in \operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12. Let I^0 be such that for all $P \in \operatorname{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$,

$$I^0(P) = \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x).$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \Delta^4(X)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $P \in \Delta^4(X)$, if there exist $\tilde{P}, \tilde{Q} \in \operatorname{dom} \widetilde{I}$ such that $\tilde{P} \succsim P \succsim \tilde{Q}$, then $P \in \operatorname{dom} \widetilde{I}$.
3. $\operatorname{dom} I^0 \subseteq \operatorname{dom} \widetilde{I}$ and for all $P \in \operatorname{dom} I^0$, $I^0(P) = \widetilde{I}(P)$.
4. For all $p \in \Delta^4(\mathbb{N}), x \in \operatorname{NDiag}(X^p)$, if $(p, x) \in \operatorname{dom} \widetilde{I}$, then there exist $x^1, x^2 \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, x) \succ (p, x^2)$ and for all $z \in \operatorname{NDiag}(X^p)$ such that $(p, x^1) \succ (p, z) \succ (p, x^2)$ and $(p, z) \in \operatorname{dom} \widetilde{I}$,
$$\widetilde{I}(p, z) - \widetilde{I}(p, x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$
5. for all $P, Q \in \operatorname{dom} \widetilde{I}$, $P \succsim Q$ iff $\widetilde{I}(P) \geq \widetilde{I}(Q)$.

If $B \in \mathcal{A}$ with an associated function I^B and there exists $P \in \Delta^4(X)$ such that $P \notin \operatorname{dom} I^B$, then there exists $B^* \supsetneq B$ such that $B^* \in \mathcal{A}$.

Proof. Let $P^0 \in \operatorname{dom} I^0$.

Denote

$$\alpha^* = \inf\{\alpha \in [0, 1] | \alpha P + (1 - \alpha) P^0 \notin \operatorname{dom} I^B\}. \tag{B.102}$$

First, since X is open, there exist $R^0, S^0 \in \operatorname{dom} I^0$ such that $R^0 \succ P^0 \succ S^0$. By Axiom 2.3, there exists $\alpha^1 \in (0, 1)$ such that for all $\alpha \in [0, \alpha^1]$,

$$R^0 \succ \alpha P + (1 - \alpha) P^0 \succ S^0.$$

By Conditions 2 and 3, for all $\alpha \in [0, \alpha^1)$, $\alpha P + (1-\alpha)P^0 \in \operatorname{dom} I^B$. So $\alpha^* > 0$.

Now as above it is not possible that $\alpha^* P + (1-\alpha^*)P^0 \sim P^0$. So assume w.l.o.g. that $\alpha^* P + (1-\alpha^*)P^0 \succ P^0$.

Let $p \in \Delta^4(\mathbb{N})$, $x^0 \in \operatorname{NDiag}(X^p)$, be such that $\alpha^* P + (1-\alpha^*)P^0 = (p, x^0)$. First, we show that

$$(p, x^0) \notin \operatorname{dom} I^B. \tag{B.103}$$

Assume per contra, that $(p, x^0) \in \operatorname{dom} I^B$. Since B is open, there exists $R \in \operatorname{dom} I^B$ such that $I^B(R) > I^B(p, x^0)$. Now

$$R \succ \alpha^* P + (1-\alpha^*)P^0 \succ P^0,$$

so by Axiom 2.3, there exists $\alpha^1 < \alpha^* < \alpha^2$ such that for all $\alpha \in (\alpha^1, \alpha^2)$,

$$R \succ \alpha P + (1-\alpha)P^0 \succ P^0.$$

By Condition 2, for all $\alpha \in (\alpha^1, \alpha^2)$,

$$\alpha P + (1-\alpha)P^0 \in \operatorname{dom} I^B$$

which contradicts, the choice of α^* in (B.102). So $(p, x^0) \in \operatorname{dom} I^B$.

Assume w.l.o.g. $1 \in \operatorname{supp} p$. Since X is open, there exist $x_1^1, x_1^2 \in X$ such that $x_1^2 > x_1^0 > x_1^1$, for all $x_1 \in [x_1^1, x_1^2]$,

$$(x_1, x_{-1}^0) \in \operatorname{NDiag}(X^p) \text{ and } \big(p, (x_1^1, x_{-1}^0)\big) \succ P^0. \tag{B.104}$$

Now

$$\big(p, (x_1^2, x_{-1}^0)\big) \succ \big(p, (x_1^0, x_{-1}^0)\big) \succ \big(p, (x_1^1, x_{-1}^0)\big).$$

By Axiom 2.3, there exists $\alpha^\dagger \in (0, \alpha^*)$ such that

$$\big(p, (x_1^2, x_{-1}^0)\big) \succ \alpha^\dagger P + (1-\alpha^\dagger)P^0 \succ \big(p, (x_1^1, x_{-1}^0)\big).$$

Now by Axiom 2.2 and (B.104), there exists $x_1^\dagger \in [x_1^1, x_1^2]$ such that

$$\big(p, (x_1^\dagger, x_{-1}^0)\big) \sim \alpha^\dagger P + (1-\alpha^\dagger)P^0.$$

By Condition 2, we have

$$\big(p, (x_1^\dagger, x_{-1}^0)\big) \in \operatorname{dom} I^B.$$

Next define I^* by the following. For all $Q \in \operatorname{dom} I^B$, define $I^*(Q) = I^B(Q)$. Second, for all $Q \notin \operatorname{dom} I^B$ such that $\big(p,(x_1^2,x_{-1}^0)\big) \succ Q \succ \big(p,(x_1^1,x_{-1}^0)\big)$, by Axiom 2.2 and (B.104), there exists $x_1^Q \in (x_1^\dagger, x_1^2)$ such that

$$\big(p,(x_1^Q,x_{-1}^0)\big) \sim Q$$

and define

$$I^*(Q) = I^B\big(p,(x_1^\dagger,x_{-1}^0)\big) + \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1^Q,x_{-1}^0)_i\big) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger,x_{-1}^0)_i\big).$$

First, we show that for all $Q \in \Delta^4(X)$ such that $\big(p,(x_1^2,x_{-1}^0)\big) \succ Q \succ \big(p,(x_1^1,x_{-1}^0)\big)$, and $x_1^Q \in (x_1^1, x_1^2)$ such that

$$\big(p,(x_1^Q,x_{-1}^0)\big) \sim Q,$$

we have

$$I^*(Q) = I^B\big(p,(x_1^\dagger,x_{-1}^0)\big) + \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1^Q,x_{-1}^0)_i\big) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger,x_{-1}^0)_i\big). \tag{B.105}$$

If $Q \notin \operatorname{dom} I^B$, then the claim is trivial. So assume that $Q \in \operatorname{dom} I^B$. By (B.104)

$$\big(p,(x_1^2,x_{-1}^0)\big) \succ Q \succ \big(p,(x_1^1,x_{-1}^0)\big) \succ P^0.$$

Since B is open, there exists $R \in \operatorname{dom} I^B$ such that

$$I^B(R) > I^B(Q) \text{ and } I^B(R) > I^B\big(p,(x_1^\dagger,x_{-1}^0)\big).$$

By Conditions 2 and 4, the condition (B.90) holds and (B.91) holds for $R \succ P^0$ and so for all $z \in \operatorname{NDiag}(X^p)$ such that $R \succ (p,z) \succ P^0$, we have since $R \succ \big(p,(x_1^\dagger,x_{-1}^0)\big) \succ P^0$,

$$I^B(p,z) = \sum_{i\in\operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger,x_{-1}^0)_i\big) + I^B\big(p,(x_1^\dagger,x_{-1}^0)\big).$$

So especially, we have by (5), since $R \succ \big(p,(x_1^Q,x_{-1}^0)\big) \succ P^0$,

$$I^*(Q) = I^B(Q) = I^B\big(p,(x_1^Q,x_{-1}^0)\big) =$$
$$\sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1^Q,x_{-1}^0)_i\big) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger,x_{-1}^0)_i\big) + I^B\big(p,(x_1^\dagger,x_{-1}^0)\big).$$

Next, we show that I^* satisfies Conditions 1-5 for the set $\operatorname{Im} I^*$.

Condition 1: We show that $\operatorname{Im} I^*$ is an open interval. Since u is strictly increasing and continuous function and X is an interval, by Austin (1985) u^{-1} is continuous and u is an open map. By the definition of I^* and (B.105),

$$\operatorname{Im} I^* = \operatorname{Im} I^B \cup \Big\{ I\big(p,(x_1^\dagger,x_{-1}^0)\big) + \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1,x_{-1}^0)_i\big) - \sum_{i\in\operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger,x_{-1}^0)_i\big) \,|\, x_1 \in (x_1^1,x_1^2)\Big\}.$$

Since u is an open map, $\operatorname{Im} I^*$ is a union of two open sets and so open. Additionally, since u is continuous, it is a union of two connected sets with an intersection, and so especially it is connected. Thus $\operatorname{Im} I^*$ is an open interval.

Condition 2: Let $P \in \Delta^4(X), \tilde{P}, \tilde{Q} \in \operatorname{dom} I^*$ such that $\tilde{P} \succsim P \succsim \tilde{Q}$. We consider two cases. First, assume that $P \succsim \alpha^* P + (1-\alpha^*)P^0$. If $\tilde{P} \in \operatorname{dom} I^B$, then by Condition 2 for I^B, $\alpha^* P + (1-\alpha^*)P^0 \in \operatorname{dom} I^B$ that contradicts (B.103). So $\tilde{P} \notin \operatorname{dom} I^B$. By the definition of I^* it must be that $\big(p, (x_1^2, x_{-1}^0)\big) \succ \tilde{P}$. Now

$$\big(p, (x_1^2, x_{-1}^0)\big) \succ \tilde{P} \succsim P \succsim \alpha^* P + (1-\alpha^*)P^0 \succ \big(p, (x_1^2, x_{-1}^0)\big)$$

and so by the definition of I^*, $P \in \operatorname{dom} I^*$. Second, assume that $\alpha^* P + (1-\alpha^*)P^0 \succ P$. Now by the definition of I^*, it must be that $\tilde{Q} \in \operatorname{dom} I^B$. Additionally, by Axiom 2.3, there exists $\alpha^1 \in (0, \alpha^*)$ such that $\alpha^1 P + (1-\alpha^1)P^0 \succ P$. By (B.102), $\alpha^1 P + (1-\alpha^1)P^0 \in \operatorname{dom} I^B$. By Condition 2 for I^B, we have $P \in \operatorname{dom} I^B$.

Condition 3: Follows trivially since I^B satisfies Condition 3.

Condition 4: Let $q \in \Delta^4(\mathbb{N}), x \in \operatorname{NDiag}(X^q)$ be such that $(q, y) \in \operatorname{dom} I^*$. If $(q, y) \in \operatorname{dom} I^B$, then the claim follows since I^B satisfies Condition 4. So assume that $(q, y) \notin \operatorname{dom} I^B$. By the definition of I^*,

$$\big(p, (x_1^2, x_{-1}^0)\big) \succ (q, y) \succ \big(p, (x_1^1, x_{-1}^0)\big).$$

Now by Axiom 2.2 and (B.104), there exists $x_1^y \in (x_1^1, x_1^2)$ such that

$$\big(p, (x_1^y, x_{-1}^0)\big) \sim (q, y).$$

By Lemma B.64, there exist $\tilde{z}^1, \tilde{z}^2 \in \operatorname{NDiag}(X^p), \tilde{y}^1, \tilde{y}^2 \in \operatorname{NDiag}(X^q)$ such that

$$(p, \tilde{z}^2) \sim (q, \tilde{y}^2) \succ \big(p, (x_1^y, x_{-1}^0)\big) \sim (q, y) \succ (p, \tilde{z}^1) \sim (q, \tilde{y}^1)$$

and for all $\tilde{x} \in \operatorname{NDiag}(X^p), \tilde{y} \in \operatorname{NDiag}(X^q)$ such that $(p, \tilde{z}^2) \succsim (p, \tilde{x}), (q, \tilde{y}) \succsim (p, \tilde{z}^1)$,

$$(p, \tilde{x}) \succsim (q, \tilde{y}) \iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(\tilde{x}_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^y, x_{-1}^0)_i\big) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u(\tilde{y}_i) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i)$$

and

$$(q, \tilde{y}) \succsim (p, \tilde{x}) \iff \sum_{i \in \operatorname{supp} q} q_i^\lambda u(\tilde{y}_i) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^y, x_{-1}^0)_i\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$

By Axiom 2.2, there exist $z^1, z^2 \in \operatorname{NDiag}(X^p)$ such that

$$(p, \tilde{z}^2), \big(p, (x_1^2, x_{-1}^0)\big) \succ (p, z^2) \succ (q, y) \succ (p, z^1) \succ \big(p, (x_1^1, x_{-1}^0)\big), (p, \tilde{z}^1). \tag{B.106}$$

By Axiom 2.2 and Lemma B.36, there exist $y^1, y^2 \in \mathrm{NDiag}(X^q)$ such that

$$(p, z^2) \sim (q, y^2) \text{ and } (p, z^1) \sim (q, y^1).$$

Let $\tilde{y} \in \mathrm{NDiag}(X^q)$ be such that $(q, y^2) \succ (q, \tilde{y}) \succ (q, y^1)$ and $(q, \tilde{y}) \in \operatorname{dom} I^*$. By (B.106),

$$\big(p, (x_1^2, x_{-1}^0)\big) \succ (q, y) \succ \big(p, (x_1^1, x_{-1}^0)\big).$$

By Axiom 2.2 and Lemma B.36 and (B.104), there exists $x_1^{\tilde{y}} \in (x_1^1, x_1^2)$ such that

$$\big(p, (x_1^{\tilde{y}}, x_{-1}^0)\big) \sim (q, \tilde{y}).$$

Now we have by above,

$$\sum_{i \in \operatorname{supp} q} q_i^\lambda u(\tilde{y}_i) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^y, x_{-1}^0)_i\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^{\tilde{y}}, x_{-1}^0)_i\big). \tag{B.107}$$

And we have by (B.105) and the definition of I^* or since I^B satisfies Condition 5,

$$\begin{aligned}
&I^*(q, \tilde{y}) - I^*(q, y) = I^*\big(p, (x_1^{\tilde{y}}, x_{-1}^0)\big) - I^*\big(p, (x_1^y, x_{-1}^0)\big) \\
&= \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^{\tilde{y}}, x_{-1}^0)_i\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger, x_{-1}^0)_i\big) + I^B\big(p, (x_1^\dagger, x_{-1}^0)\big) \\
&\qquad\qquad - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^y, x_{-1}^0)_i\big) + \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger, x_{-1}^0)_i\big) - I^B\big(p, (x_1^\dagger, x_{-1}^0)\big) \\
&= \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^{\tilde{y}}, x_{-1}^0)_i\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^y, x_{-1}^0)_i\big) \overset{\text{(B.107)}}{=} \sum_{i \in \operatorname{supp} q} q_i^\lambda u(\tilde{y}_i) - \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i)
\end{aligned}$$

Condition 5: Let $P, Q \in \operatorname{dom} I^*$. First, if $P, Q \in \operatorname{dom} I^B$, then since I^B satisfies Condition 5, we have $P \succsim Q \iff I^*(P) \geq I^*(Q)$. So assume w.l.o.g. that $P \notin \operatorname{dom} I^B$. By the definition of I^*, Axiom 2.2, and (B.104,B.105), there exists $x_1^P \in (x_1^1, x_1^2)$ such that

$$\big(p, (x_1^P, x_{-1}^0)\big) \sim P \tag{B.108}$$

and

$$I^*(P) = I^B\big(p, (x_1^\dagger, x_{-1}^0)\big) + \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^P, x_{-1}^0)_i\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger, x_{-1}^0)_i\big). \tag{B.109}$$

First, if $\big(p, (x_1^\dagger, x_{-1}^0)\big) \succsim P$, then by the definition of I^*,

$$P \succsim \big(p, (x_1^1, x_{-1}^0)\big) \succ P^0$$

and so by Condition 2, $P \in \operatorname{dom} I^B$ which is a contradiction. Thus

$$\big(p, (x_1^P, x_{-1}^0)\big) \sim P \succ \big(p, (x_1^\dagger, x_{-1}^0)\big).$$

Since $(p_i^\lambda u)_{i \in \operatorname{supp} p}$ is an additive representation for $\succsim^p$, we have

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^P, x_{-1}^0)_i\big) > \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger, x_{-1}^0)_i\big)$$

and so

$$I^*(P) > I^*\big(p, (x_1^\dagger, x_{-1}^0)\big).$$

First, if $\big(p, (x_1^\dagger, x_{-1}^0)\big) \succsim Q$, then by Condition 2, $Q \in \operatorname{dom} I^B$ and so by Condition 5,

$$I^*(Q) \leq I^*\big(p, (x_1^\dagger, x_{-1}^0)\big) < I^*(P)$$

and

$$Q \precsim \big(p, (x_1^\dagger, x_{-1}^0)\big) \prec P.$$

Next assume that $Q \succ \big(p, (x_1^\dagger, x_{-1}^0)\big)$. Now by Axiom 2.2, and (B.104), there exists $x_1^Q \in (x_1^1, x_1^2)$ such that

$$\big(p, (x_1^Q, x_{-1}^0)\big) \sim Q. \tag{B.110}$$

By (B.105)

$$I^*(Q) = I^B\big(p, (x_1^\dagger, x_{-1}^0)\big) + \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^Q, x_{-1}^0)_i\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger, x_{-1}^0)_i\big). \tag{B.111}$$

Since $(p_i^\lambda u)_{i \in \operatorname{supp} p}$ is an additive representation for $\succsim^p$, we have

$$\begin{aligned}
P \succsim Q \overset{\text{(B.108,B.110)}}{\iff} & \quad \big(p, (x_1^P, x_{-1}^0)\big) \succsim \big(p, (x_1^Q, x_{-1}^0)\big) \\
\iff & \quad \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^P, x_{-1}^0)_i\big) \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^Q, x_{-1}^0)_i\big) \\
\iff & \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^P, x_{-1}^0)_i\big) + I^B\big(p, (x_1^\dagger, x_{-1}^0)\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger, x_{-1}^0)_i\big) \\
& \qquad \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^Q, x_{-1}^0)_i\big) + I^B\big(p, (x_1^\dagger, x_{-1}^0)\big) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u\big((x_1^\dagger, x_{-1}^0)_i\big) \\
\overset{\text{(B.109,B.111)}}{\iff} & \quad I^*(P) = I^*(Q).
\end{aligned}$$

Thus $\operatorname{Im} I^* \in \mathcal{A}$ and by Condition 5 for I^*, $\operatorname{Im} I^* \supsetneq \operatorname{Im} I^B$. □

By putting the previous lemmas together, we have by Zorn's lemma almost the desired representation without continuity for C.

Lemma B.70 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i \in \operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12. Then there exists $C : \Delta^4(\mathbb{N}) \to \mathbb{R}$ such that for all $p, q \in \Delta^4(\mathbb{N})$, $x \in \operatorname{NDiag}(X^p), y \in \operatorname{NDiag}(X^q)$

$$(p, x) \succsim (q, y) \iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + C(p) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) + C(q).$$

Proof. Denote I^0 be such that for all $P \in \mathrm{Dist}(\Delta^{(\frac{1}{4})^4_{i=1}})$,

$$I^0(P) = \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x).$$

Let $\mathcal{A} \subseteq \{(a_1, a_2) \subseteq \mathbb{R} | a_1 \leq a_2\}$ be such that $A \in \mathcal{A}$ if and only if

1. There exists $\widetilde{I}$ such that $\operatorname{dom} \widetilde{I} \subseteq \Delta^4(X)$ and $\operatorname{Im} \widetilde{I} = A$.
2. For all $P \in \Delta^4(X)$, if there exist $\tilde{P}, \tilde{Q} \in \operatorname{dom} \widetilde{I}$ such that $\tilde{P} \succsim P \succsim \tilde{Q}$, then $P \in \operatorname{dom} \widetilde{I}$.
3. $\operatorname{dom} I^0 \subseteq \operatorname{dom} \widetilde{I}$ and for all $P \in \operatorname{dom} I^0$, $I^0(P) = \widetilde{I}(P)$.
4. For all $p \in \Delta^4(\mathbb{N}), x \in \mathrm{NDiag}(X^p)$, if $(p, x) \in \operatorname{dom} \widetilde{I}$, then there exist $x^1, x^2 \in \mathrm{NDiag}(X^p)$ such that $(p, x^1) \succ (p, x) \succ (p, x^2)$ and for all $z \in \mathrm{NDiag}(X^p)$ such that $(p, x^1) \succ (p, z) \succ (p, x^2)$ and $(p, z) \in \operatorname{dom} \widetilde{I}$,
$$\widetilde{I}(p, z) - \widetilde{I}(p, x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i).$$
5. for all $P, Q \in \operatorname{dom} \widetilde{I}$, $P \succsim Q$ iff $\widetilde{I}(P) \geq \widetilde{I}(Q)$.

We order $\mathcal{A}$ by $\subseteq$. By Lemma B.67, $\mathcal{A} \neq \varnothing$. By Lemma B.68, if $\mathcal{B} \subseteq \mathcal{A}$ is a chain, then there exists $D \in \mathcal{A}$ such that for all $B \in \mathcal{B}$, $B \subseteq D$. Thus by Zorn's lemma, there exists a maximal element $B^* \in \mathcal{A}$. Let I^* be the associated function. By Lemma B.70, $\operatorname{dom} I^* = \Delta^4(X)$. By Condition 5, the only thing left to show is that there exists $C : \Delta^4(\mathbb{N}) \to \mathbb{R}$ such that for all $p \in \Delta^4(\mathbb{N})$ and $x \in \mathrm{NDiag}(X^p)$,

$$I^*(p, x) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + C(p).$$

For each $p \in \Delta^4(\mathbb{N})$, let $x^p \in \mathrm{NDiag}(X^p)$. Define

$$C(p) = I^*(p, x^p) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^p). \tag{B.112}$$

We show that for all $p \in \Delta^4(\mathbb{N})$, $y \in \mathrm{NDiag}(X^p)$

$$I^*(p, y) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(y_i) + C(p).$$

Let $p \in \Delta^4(\mathbb{N}), y \in \mathrm{NDiag}(X^p)$. Assume w.l.o.g. that $(p, x^p) \succsim (p, y)$. Since X is open, there exist $x^1, x^2 \in \mathrm{NDiag}(X^p)$ such that

$$(p, x^2) \succ (p, x^p) \succsim (p, y) \succ (p, x^1).$$

By Condition 4 and Lemma B.65,

$$I^*(p, y) = \sum_{i \in \operatorname{supp} p} p_i^\lambda u(y_i) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i^p) + I^*(p, x^p) \overset{\text{(B.112)}}{=} \sum_{i \in \operatorname{supp} p} p_i^\lambda u(y_i) + C(p).$$

□

Finally, we establish that C is linearly continuous.

Lemma B.71 Assume that for all co-probabilities $p \in \Delta^4(\mathbb{N})$, there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that $\succsim^p$ has an additive representation by $(p_i^\lambda u)_{i \in \operatorname{supp} p}$. Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.12. Then there exists a linearly continuous $C : \Delta(\mathbb{N}) \to \mathbb{R}$ such that for all $p, q \in \Delta(\mathbb{N})$, $x \in \operatorname{NDiag}(X^p), y \in \operatorname{NDiag}(X^q)$

$$(p, x) \succsim (q, y) \iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + C(p) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) + C(q).$$

Proof. By Lemma B.70, there exists $C : \Delta^4(\mathbb{N}) \to \mathbb{R}$ such that for all $p, q \in \Delta^4(\mathbb{N})$, $x \in \operatorname{NDiag}(X^p)$, $y \in \operatorname{NDiag}(X^q)$

$$(p, x) \succsim (q, y) \iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + C(p) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) + C(q).$$

Let $p^0 \in \Delta(\mathbb{N}) \setminus \Delta^4(\mathbb{N})$. First, we show that for all $p, q \in \Delta^4(\mathbb{N})$,

$$\lim_{\alpha \to 1} c(\alpha p^0 + (1-\alpha)p) = \lim_{\alpha \to 1} c(\alpha p^0 + (1-\alpha)q). \tag{B.113}$$

Let $\varepsilon > 0$ and $x^0 \in \operatorname{NDiag}(X^{p^0 \cup p \cup q})$. Since X is open, by Lemma B.31, there exist $x^2, x^1 \in \operatorname{NDiag}(X^{p^0 \cup p})$ such that

$$(p^0, x^2) \succ (p^0, x^0) \succ (p^0, x^1).$$

By Axiom 2.3, there exists $\alpha^0 \in (0, 1)$ such that

$$(\alpha^0 p^0 + (1-\alpha^0)p, x^2) \succ (p^0, x^0) \succ (\alpha^0 p^0 + (1-\alpha^0)p, x^1). \tag{B.114}$$

Denote $r = \alpha^0 p^0 + (1-\alpha^0)p$. Now $|\operatorname{supp} r| \geq |\operatorname{supp} p| \geq 4$. Additionally, by Lemma B.36 and (B.114), there exists $z^0 \in \operatorname{NDiag}(X^r)$ such that

$$(p^0, x^0) \sim (r, z^0).$$

And by Lemma B.36 since X is open and u is continuous, there exist $z^1, z^2 \in \operatorname{NDiag}(X^r)$ such that $(r, z^2) \succ (r, z^0) \succ (r, z^1)$ and

$$\Big| \sum_{i \in \operatorname{supp} r} r_i^\lambda u(z_i^2) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i^0) \Big|, \Big| \sum_{i \in \operatorname{supp} r} r_i^\lambda u(z_i^1) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(z_i^0) \Big| < \frac{\varepsilon}{4}.$$

Since $\lambda > 0$, the mappings

$$[0, 1] \ni \alpha \mapsto \sum_{i \in \operatorname{supp} p^0 \cup \operatorname{supp} p} (\alpha p_i^0 + (1-\alpha)p_i)^\lambda u(x^0)$$

and

$$[0, 1] \ni \alpha \mapsto \sum_{i \in \operatorname{supp} p^0 \cup \operatorname{supp} q} (\alpha p_i^0 + (1-\alpha)q_i)^\lambda u(x^0)$$

are continuous functions as a sum of power functions, there exists $\delta > 0$ such that for all $\alpha \in (\delta, 1)$,

$$\Big| \sum_{i \in \operatorname{supp} p^0 \cup \operatorname{supp} p} (\alpha p_i^0 + (1-\alpha) p_i)^\lambda u(x^0) - \sum_{i \in \operatorname{supp} p^0} (p_i^0)^\lambda u(x^0) \Big| < \frac{\varepsilon}{4} \tag{B.115}$$

and

$$\Big| \sum_{i \in \operatorname{supp} p^0 \cup \operatorname{supp} q} (\alpha p_i^0 + (1-\alpha) q_i)^\lambda u(x^0) - \sum_{i \in \operatorname{supp} p^0} (p_i^0)^\lambda u(x^0) \Big| < \frac{\varepsilon}{4}$$

Now since for all $\alpha \in [0,1]$, $p(\alpha)_1 = p_1$, we have $\sum_{i \in \operatorname{supp} p(\alpha) \setminus \{1\}} p(\alpha)_i^\lambda > 0$. Now since $\lambda > 0$, this is a continuous function as a function of $\alpha \in [0,1]$ and so especially $\min_{\alpha \in [0,1]} \sum_{i \in \operatorname{supp} p(\alpha) \setminus \{1\}} p(\alpha)_i^\lambda > 0$.

By Axiom 2.3, there exists $\alpha^1 \in (\delta, 1)$ such that for all $\alpha \in (\alpha^1, 1)$,

$$(r, z^2) \succ (\alpha p^0 + (1-\alpha) p, x^0), (\alpha p^0 + (1-\alpha) q, x^0) \succ (r, z^1).$$

Let $\alpha, \alpha' \in (\alpha^1, 1)$. Now

$$\Big| \operatorname{supp} \alpha' p^0 + (1-\alpha') p \Big|, \Big| \operatorname{supp} \alpha p^0 + (1-\alpha) p \Big|, \Big| \operatorname{supp} \alpha p^0 + (1-\alpha) q \Big| \geq 4.$$

Denote for all $z \in \operatorname{NDiag}(X^r)$,

$$I(r, z) = \sum_{i \in \operatorname{supp} r} (r_i)^\lambda u(z_i^0) + c(r).$$

First, we have by the representation,

$$I(r, z^2) > \sum_{i \in \operatorname{supp} p^0 \cup p} (\alpha p_i^0 + (1-\alpha) p_i)^\lambda u(x_i^0) + c\big(\alpha p^0 + (1-\alpha) p\big) > I(r, z^1)$$

and so

$$\begin{aligned} \frac{\varepsilon}{4} &> I(r, z^2) - I(r, z) > \sum_{i \in \operatorname{supp} p^0 \cup p} (\alpha p_i^0 + (1-\alpha) p_i)^\lambda u(x_i^0) + c\big(\alpha p^0 + (1-\alpha) p\big) - I(r, z) \\ &> I(r, z^1) - I(r, z) > -\frac{\varepsilon}{4}. \end{aligned} \tag{B.116}$$

So we have

$$\begin{aligned} &\Big| \sum_{i \in \operatorname{supp} p^0} (p_i^0)^\lambda u(x_i^0) + c\big(\alpha p^0 + (1-\alpha) p\big) - I(r, z^0) \Big| \\ &\leq \Big| \sum_{i \in \operatorname{supp} p^0} (p_i^0)^\lambda u(x^0) - \sum_{i \in \operatorname{supp} p^0 \cup p} (\alpha p_i^0 + (1-\alpha) p_i)^\lambda u(x_i^0) \Big| \\ &\qquad + \Big| \sum_{i \in \operatorname{supp} p^0 \cup p} (\alpha p_i^0 + (1-\alpha) p_i)^\lambda u(x_i^0) + c\big(\alpha p^0 + (1-\alpha) p\big) - I(r, z^0) \Big| \\ &\overset{(B.115, B.116)}{<} \frac{\varepsilon}{2}. \end{aligned}$$

And symmetrically

$$\Big| \sum_{i \in \operatorname{supp} p^0} (p_i^0)^\lambda u(x_i^0) + c\big(\alpha p^0 + (1-\alpha) q\big) - I(r, z^0) \Big| < \frac{\varepsilon}{2}$$

and

$$\Big| \sum_{i\in\operatorname{supp} p^0} (p_i^0)^\lambda u(x_i^0) + c\big(\alpha' p^0 + (1-\alpha')p\big) - I(r,z^0)\Big| < \frac{\varepsilon}{2}.$$

So we have first,

$$\begin{aligned}
&\Big|c\big(\alpha p^0 + (1-\alpha)p\big) - c\big(\alpha' p^0 + (1-\alpha')p\big)\Big| \qquad \text{(B.117)}\\
&\leq \Big| \sum_{i\in\operatorname{supp} p^0} (p_i^0)^\lambda u(x_i^0) + c\big(\alpha p^0 + (1-\alpha)p\big) - I(r,z^0)\Big| + \Big| \sum_{i\in\operatorname{supp} p^0} (p_i^0)^\lambda u(x_i^0) + c\big(\alpha' p^0 + (1-\alpha')p\big) - I(r,z^0)\Big|\\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

and second

$$\begin{aligned}
&\Big|c\big(\alpha p^0 + (1-\alpha)p\big) - c\big(\alpha p^0 + (1-\alpha)q\big)\Big| \qquad \text{(B.118)}\\
&\leq \Big| \sum_{i\in\operatorname{supp} p^0} (p_i^0)^\lambda u(x_i^0) + c\big(\alpha p^0 + (1-\alpha)p\big) - I(r,z^0)\Big| + \Big| \sum_{i\in\operatorname{supp} p^0} (p_i^0)^\lambda u(x_i^0) + c\big(\alpha p^0 + (1-\alpha)q\big) - I(r,z^0)\Big|\\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Thus first by (B.117), $\big(c(\alpha p^0 + (1-\alpha)p)\big)_{\alpha\in(0,1)}$ forms a Cauchy sequence as $\alpha \to 1$ so $\lim_{\alpha\to 1} c(\alpha p^0 + (1-\alpha)p)$ exists. Second, by (B.118), $\lim_{\alpha\to 1} c(\alpha p^0 + (1-\alpha)q)$ also exists and $\lim_{\alpha\to 1} c(\alpha p^0 + (1-\alpha)q) = \lim_{\alpha\to 1} c(\alpha p^0 + (1-\alpha)p)$. This shows (B.113).

Let $p^* \in \Delta^4(\mathbb{N})$. For all $p^0 \in \Delta(\mathbb{N}) \setminus \Delta^4(\mathbb{N})$, define

$$c(p^0) = \lim_{\alpha\to 1} c\big(\alpha p^0 + (1-\alpha)p^*\big).$$

We show that for all $p, q \in \Delta(\mathbb{N})$, $x \in \operatorname{NDiag}(X^p)$, $y \in \operatorname{NDiag}(X^q)$

$$(p,x) \gtrsim (q,y) \iff \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) + c(p) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) + c(q).$$

If $p, q \in \Delta^4(\mathbb{N})$, then the claim follows from Lemma B.70. So assume w.l.o.g. that $p \notin \Delta^4(\mathbb{N})$. We will show the equivalent claim that

$$(p,x) \succ (q,y) \Rightarrow \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) + c(p) \geq \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) + c(q) \qquad \text{(B.119)}$$

and

$$(p,x) \sim (q,y) \Rightarrow \sum_{i\in\operatorname{supp} p} p_i^\lambda u(x_i) + c(p) = \sum_{i\in\operatorname{supp} q} q_i^\lambda u(y_i) + c(q).$$

So let $(p,x) \succ (q,y)$ and $i^* \in \operatorname{supp} p$. Since X is open, there exists $x_{i^*}^1 \in X$ such that $x_{i^*} > x_{i^*}^1$, $(x_{i^*}^1, x_{-i^*}) \in \operatorname{NDiag}(X^p)$ and $\big(p, (x_{i^*}^1, x_{-i^*})\big) \succ (q,y)$. By Lemma B.31, $(p,x) \succ \big(p, (x_{i^*}^1, x_{-i^*})\big)$. Let

$z^1 \in \mathrm{NDiag}(X^{p^* \cup p})$, $z^2 \in \mathrm{NDiag}(X^{p^* \cup q})$ be such that for all $i \in \operatorname{supp} p$, $z^1_i = (x^1_{i^*}, x_{-i^*})_i$ and for all $j \in \operatorname{supp} q$, $z^2_i = y_i$. By Axiom 2.3, there exists $R \in \Delta(X)$ such that

$$\big(p, (x^1_{i^*}, x_{-i^*})\big) \succ R \succ (q, y).$$

By Axiom 2.3, there exists $\alpha^* \in (0,1)$ such that for all $\alpha \in (\alpha^*, 1)$,

$$\big(\alpha p + (1-\alpha)p^*, z^1\big)\big) = \alpha\big(p, (x^1_{i^*}, x_{-i^*})\big) + (1-\alpha)(p^*, z^1) \succ R$$

and

$$R \succ \alpha(q, y) + (1-\alpha)(p^*, z^2) = \big(\alpha p + (1-\alpha)p^*, z^2\big)\big).$$

Since $|\operatorname{supp} p^*| \geq 4$ by the representation, we have

$$\sum_{i \in \operatorname{supp} p \cup \operatorname{supp} p^*} (\alpha p_i + (1-\alpha)p^*_i)^\lambda_i u(z^1_i) + C(\alpha p + (1-\alpha)p^*)$$
$$> \sum_{i \in \operatorname{supp} q \cup \operatorname{supp} p^*} (\alpha q_i + (1-\alpha)p^*_i)^\lambda_i u(z^2_i) + C(\alpha q + (1-\alpha)p^*).$$

Now the mappings

$$[0,1] \ni \alpha \mapsto \sum_{i \in \operatorname{supp} p^0 \cup \operatorname{supp} p} (\alpha p_i + (1-\alpha)p*_i)^\lambda u(z^1)$$

and

$$[0,1] \ni \alpha \mapsto \sum_{i \in \operatorname{supp} p^0 \cup \operatorname{supp} q} (\alpha q_i + (1-\alpha)p^*_i)^\lambda u(z^2)$$

are continuous as power functions. So by (B.113), we have by taking $\alpha \to 1$ by the definition of z^1 and z^2,

$$\sum_{i \in \operatorname{supp} p} p^\lambda_i u\big((x^1_{i^*}, x_{-i^*})_i\big) + C(p) \geq \sum_{i \in \operatorname{supp} q} q^\lambda_i u(y_i) + C(q).$$

Finally, since u is strictly increasing and $i^* \in \operatorname{supp} p$, we have

$$\sum_{i \in \operatorname{supp} p} p^\lambda_i u(x_i) + C(p) > \sum_{i \in \operatorname{supp} q} q^\lambda_i u(y_i) + C(q).$$

Next assume that $(p, x) \sim (q, y)$ and $i^* \in \operatorname{supp} p$. Now since X is open, there exists $(x^m_{i^*})_{m=1}^\infty \subseteq X$ such that $x^m_{i^*} \to x_{i^*}$ as $m \to \infty$ and for all $m \in \mathbb{N}$, $x^m_{i^*} > x_{i^*}$ and $(x^m_{i^*}, x_{-i^*}) \in \mathrm{NDiag}(X^p)$. By Lemma B.31, for all $m \in \mathbb{N}$,

$$\big(p, (x^m_{i^*}, x_{-i})\big) \succ (q, y)$$

and by (B.119)

$$\sum_{i \in \operatorname{supp} p} p^\lambda_i u\big((x^m_{i^*}, x_{-i^*})_i\big) + C(p) > \sum_{i \in \operatorname{supp} q} q^\lambda_i u(y_i) + C(q).$$

Now by taking $m \to \infty$, we have by the continuity of u,

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + C(p) \geq \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) + C(q).$$

Symmetrically since X is open, there exists a sequence $(\tilde{x}_{i^*}^m)_{m=1}^\infty \subseteq X$ such that $\tilde{x}_{i^*}^m \to x_{i^*}$ as $m \to \infty$ and for all $m \in \mathbb{N}$, $\tilde{x}_{i^*}^m < x_{i^*}$ and $(\tilde{x}_{i^*}^m, x_{-i^*}) \in \mathrm{NDiag}(X^p)$. Using this sequence, we get

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + C(p) \leq \sum_{i \in \operatorname{supp} q} q_i^\lambda u(y_i) + C(q).$$

This shows the representation claim.

Finally, we show that c is linearly continuous. Let $p, q \in \Delta(\mathbb{N})$, $\varepsilon > 0$, $x^0 \in \mathrm{NDiag}(X^{p \cup q})$. First, since the mapping

$$[0,1] \ni \alpha \mapsto \sum_{i \in \operatorname{supp} p^0 \cup \operatorname{supp} p} (\alpha p_i + (1-\alpha) q_i)^\lambda u(x^0)$$

is continuous, there exists $\delta > 0$ such that for all $\alpha \in (\delta, 1]$,

$$\Big| \sum_{i \in \operatorname{supp} p^0 \cup \operatorname{supp} p} (\alpha p_i + (1-\alpha) q_i)^\lambda u(x^0) - \sum_{i \in \operatorname{supp} p^0} (p_i^0)^\lambda u(x^0) \Big| < \frac{\varepsilon}{2}. \tag{B.120}$$

Since X is open and u is continuous, there exist $x^1, x^2 \in \mathrm{NDiag}(X^p)$ such that

$$\sum_{i \in \operatorname{supp} p} u(x_i^0) + \frac{\varepsilon}{2} > \sum_{i \in \operatorname{supp} p} u(x_i^2) > \sum_{i \in \operatorname{supp} p} u(x_i^0) > \sum_{i \in \operatorname{supp} p} u(x_i^1) > \sum_{i \in \operatorname{supp} p} u(x_i^0) - \frac{\varepsilon}{2}. \tag{B.121}$$

Now by the representation, we have $(p, x^2) \succ (p, x^0) \succ (p, x^1)$. By Axiom 2.3, there exists $\alpha^* \in (\delta, 1)$ such that for all $\alpha \in (\alpha^*, 1)$,

$$(p, x^2) \succ (\alpha p + (1-\alpha) q, x^0) \succ (p, x^1).$$

Let $\alpha \in (\alpha^*, 1)$. So we have by the representation,

$$\sum_{i \in \operatorname{supp} p} p_i^\lambda u(x^2) + c(p) > \sum_{i \in \operatorname{supp} p \cup \operatorname{supp} q} (\alpha p_i + (1-\alpha) q_i)^\lambda u(x^0) + c(\alpha p + (1-\alpha) q) > \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x^1) + c(p).$$

Thus by (B.121)

$$\Big| \sum_{i \in \operatorname{supp} p \cup \operatorname{supp} q} (\alpha p_i + (1-\alpha) q_i)^\lambda u(x^0) + c(\alpha p + (1-\alpha) q) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x^0) - c(p) \Big| < \frac{\varepsilon}{2}. \tag{B.122}$$

Finally, we have

$$\begin{aligned}
&\big| c(\alpha p + (1-\alpha) q) - c(p) \big| \\
&\leq \Big| \sum_{i \in \operatorname{supp} p \cup \operatorname{supp} q} (\alpha p_i + (1-\alpha) q_i)^\lambda u(x^0) + c(\alpha p + (1-\alpha) q) - \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x^0) - c(p) \Big| \\
&\qquad + \Big| \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x^0) - \sum_{i \in \operatorname{supp} p \cup \operatorname{supp} q} (\alpha p_i + (1-\alpha) q_i)^\lambda u(x^0) \Big| \\
&\overset{(B.120, B.122)}{<} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

Thus since $\varepsilon > 0$ was arbitrary,

$$\lim_{\alpha \to 1} c(\alpha p + (1-\alpha)q) = c(p)$$

and so c is linearly continuous.

By adding a constant to c, c will be an information function with for all $x \in X$, $c(\delta_x) = 0$.

□

The next result puts the above together and establishes the necessity of the axioms.

Proposition B.72 $\succsim$ satisfies Axioms 2.1 to 2.3, 2.10, and 2.12 iff there exist a continuous, strictly increasing, $u : X \to \mathbb{R}$, $\lambda > 0$ and a linearly continuous information function $c : \Delta(X) \to \mathbb{R}$ such that for all simple lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x) + c(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^\lambda u(x) + c(Q).$$

Proof. If direction follows from Lemmas B.61 and B.71.

We show the only if direction. By the representation with a continuous and strictly increasing utility $u : X \to \mathbb{R}$, Axioms 2.1 and 2.2 hold.

Axiom 2.3: Let $P, Q, R \in \Delta(X)$ and $(\alpha_i)_{i=1}^\infty \subseteq [0,1]$ be such that $\lim_{i \to \infty} \alpha_i \to \infty = \alpha^*$ be such that for all $i \in \mathbb{N}$, $\alpha_i P + (1-\alpha_i)Q \succsim R$. First, either the set $\{i \in \mathbb{N} | \alpha_i > \alpha^*\}$ or the set $\{i \in \mathbb{N} | \alpha_i \leq \alpha^*\}$ is countable. Assume w.l.o.g. that the set $\{i \in \mathbb{N} | \alpha_i \leq \alpha^*\}$ is countable. Then, there exists a subsequence $(\alpha_{i_j})_{j=1}^\infty$ such that for all $j \in \mathbb{N}$, $\alpha_{i_j} \leq \alpha^*$. Now we have for all $j \in \mathbb{N}$,

$$\alpha_{i_j} P + (1-\alpha_{i_j})Q = \Big(\frac{\alpha_{i_j}}{\alpha^*}\Big)(\alpha^* P + (1-\alpha^*)Q) + \Big(1 - \frac{\alpha_{i_j}}{\alpha^*}\Big)P$$

by the definition of mixtures, since for all $j \in \mathbb{N}$, $\alpha_{i_j} \leq \alpha^*$. We have by the representation, for all $j \in \mathbb{N}$,

$$\sum_{x \in \operatorname{supp} P \cup Q} \big(\alpha_{i_j} P(x) + (1-\alpha_{i_j})Q(x)\big)^\lambda u(x) + c\bigg(\Big(\frac{\alpha_{i_j}}{\alpha^*}\Big)(\alpha^* P + (1-\alpha^*)Q) + \Big(1 - \frac{\alpha_{i_j}}{\alpha^*}\Big)P\bigg)$$
$$\geq \sum_{x \in \operatorname{supp} R} R(x)^\lambda u(x) + c(R).$$

So by taking $j \to \infty$, we have by the linear continuity of c since $\alpha^* P + (1-\alpha^*)Q \in \Delta(X)$,

$$\sum_{x \in \operatorname{supp} P \cup Q} \big(\alpha^* P(x) + (1-\alpha^*)Q(x)\big)^\lambda u(x) + c\big(\alpha^* P + (1-\alpha^*)Q\big)$$
$$= \sum_{x \in \operatorname{supp} P \cup Q} \lim_{j \to \infty} \big(\alpha_{i_j} P(x) + (1-\alpha_{i_j})Q(x)\big)^\lambda u(x) +$$
$$\lim_{l \to \infty} c\bigg(\Big(\frac{\alpha_{i_l}}{\alpha^*}\Big)(\alpha^* P + (1-\alpha^*)Q) + \Big(1 - \frac{\alpha_{i_l}}{\alpha^*}\Big)P\bigg)$$

$$\geq \sum_{x \in \operatorname{supp} R} R(x)^\lambda u(x) + c(R)$$

and so $\alpha^* P + (1-\alpha^*)Q \succsim R$. Thus the set

$$\{\alpha \in [0,1] | \alpha P + (1-\alpha)Q \succsim R\}$$

is closed. Symmetrically, the set $\{\alpha \in [0,1] | R \succsim \alpha P + (1-\alpha)Q\}$ is closed.

Axiom 2.10: Let $p \in \Delta(\mathbb{N}), x, y \in \mathrm{NDiag}(X^p), \alpha \in (0,1), z \in X$ such that for all $i \in \operatorname{supp} p$, $x_i \neq z \neq y_i$. We have

$$\begin{aligned}
(p,x) \succsim (p,y) \qquad &\iff \qquad \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) + c(p) \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u(y_i) + c(p) \\
&\iff \qquad \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u(y_i) \\
&\iff \qquad \sum_{i \in \operatorname{supp} p} (\alpha p_i)^\lambda u(x_i) + (1-\alpha)^\lambda u(z) + c\big(\alpha p, (1-\alpha)\big) \\
&\qquad\qquad \geq \sum_{i \in \operatorname{supp} p} (\alpha p_i)^\lambda u(y_i) + (1-\alpha)^\lambda u(z) + c\big(\alpha p, (1-\alpha)\big) \\
&\iff \qquad \alpha(p,x) + (1-\alpha)z \succsim \alpha(p,y) + (1-\alpha)z.
\end{aligned}$$

Axiom 2.12: Let $P, Q \in \Delta(X)$, $z, z' \notin \operatorname{supp} P \cup \operatorname{supp} Q$ and $\alpha \in (0,1)$. Since c is an information function, we have for $R \in \{P, Q\}$,

$$\big(\alpha R + (1-\alpha)z\big) = c\big(\alpha R + (1-\alpha)z'\big).$$

So we have

$$\begin{aligned}
&\alpha P + (1-\alpha)z \succsim \alpha Q + (1-\alpha)z \\
\iff &\sum_{x \in \operatorname{supp} P} \big(\alpha P(x)\big)^\lambda u(x) + (1-\alpha)^\lambda u(z) + c(\alpha P + (1-\alpha)z) \\
&\qquad \geq \sum_{x \in \operatorname{supp} Q} \big(\alpha Q(x)\big)^\lambda u(x) + (1-\alpha)^\lambda u(z) + c(\alpha P + (1-\alpha)z) \\
\iff &\sum_{x \in \operatorname{supp} P} \big(\alpha P(x)\big)^\lambda u(x) + (1-\alpha)^\lambda u(z') + c(\alpha P + (1-\alpha)z') \\
&\qquad \geq \sum_{x \in \operatorname{supp} Q} \big(\alpha Q(x)\big)^\lambda u(x) + (1-\alpha)^\lambda u(z') + c(\alpha P + (1-\alpha)z') \\
\iff &\qquad \alpha P + (1-\alpha)z' \succsim \alpha Q + (1-\alpha)z'.
\end{aligned}$$

□

Finally, with Restricted FOSD, we rule out probability weighting.

Proposition B.73 $\succsim$ satisfies Axioms 2.1 to 2.3, 2.5, 2.10, and 2.12 iff there exist a continuous, strictly increasing, $u : X \to \mathbb{R}$ and a linearly continuous information function $c : \Delta(X) \to \mathbb{R}$ such that for all simple lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)u(x) + c(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x) + c(Q).$$

Proof. If direction follows from Lemmas B.61, B.62, and B.71. We show the only if direction. By Proposition B.72, the only Axiom left to show is Axiom 2.5. Let $p \in \Delta(\mathbb{N})$. We have for all $x, y \in \operatorname{NDiag}(X^p)$,

$$\begin{aligned}(p, x) \succsim (p, y) &\iff \sum_{i \in \operatorname{supp} p} p_i u(x_i) + c(p) \geq \sum_{i \in \operatorname{supp} p} p_i u(y_i) + c(p) \\ &\iff \sum_{i \in \operatorname{supp} p} p_i u(x_i) \geq \sum_{i \in \operatorname{supp} p} p_i u(y_i)\end{aligned}$$

that is an expected utility representation with a strictly increasing utility. Especially we have if $x \in \operatorname{NDiag}(X^p)$ first order stochastically dominates $y \in \operatorname{NDiag}(X^p)$,

$$\sum_{i \in \operatorname{supp} p} p_i u(x_i) \geq \sum_{i \in \operatorname{supp} p} p_i u(y_i) \Rightarrow (p, x) \succsim (p, y).$$

□

B.4 Entropy Modified Expected Utility

In this section, we establish the existence of entropy modified expected utility representation and partly show Theorem 2.1. We build on Appendices B.2 and B.3 and add Common Consequence Independence, Axiom 2.4, that implies Axioms 2.10 and 2.12. We show that under Common Consequence Independence, the previous variational representation will be an entropy modified expected utility. We will do this by showing the next recursivity property for c, which is a known characterization for a generalized entropy and for entropy when $\lambda = 1$.

Lemma B.74 Assume that there exist a utility function $u : X \to \mathbb{R}$, $\lambda > 0$, and a linearly continuous information function $c : \Delta(X) \to \mathbb{R}$ such that $\succsim$ represented by for all lotteries $P \in \Delta(X)$,

$$\sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x) + c(P).$$

Let $\succsim$ satisfy Axioms 2.1 to 2.4. Then for all co-probabilities $p = (p_1, \ldots, p_n)$,

$$c(p) = \sum_{i=1}^{n} c\Big(\frac{p_i}{\sum_{j=i}^{n} p_j}\Big)\Big(\sum_{j=i}^{n} p_j\Big)^\lambda$$

Proof. We omit the last index from the information function c

We show that for all co-probabilities $p=(p_1,\dots,p_n)\in\Delta(\mathbb{N})$ and $i^*\in\operatorname{supp} p$ such that $i^*<|\operatorname{supp} p|$,

$$c((p_j)_{j\in\operatorname{supp} p})-c(p_{i^*})=(1-p_{i^*})^\lambda c\Big(\Big(\frac{p_j}{1-p_{i^*}}\Big)_{j\neq i^*}\Big).$$

This shows the claim by recursion.

Let $i^*\in\operatorname{supp} p$. Denote $p_{-i^*}=\big(\frac{p_j}{1-p_{i^*}}\big)_{j\neq i^*}$ and $n=|\operatorname{supp} p|$. We do this by showing that for all $\alpha\in[0,1],\beta\in[0,1]$ such that there exist $x,y\in\operatorname{NDiag}(X^{p_{-i^*}})$,

$$\alpha p_{-i^*}x+(1-\alpha)x_n\sim\beta p_{-i^*}y+(1-\beta)y_n, \tag{B.123}$$

we have

$$c\big((\beta p_j)_{\substack{j\neq i^*\\ j\neq n}},p_{i^*}\big)-c\big((\alpha p_j)_{\substack{j\neq i^*\\ j\neq n}},p_{i^*}\big)=(1-p_{i^*})^\lambda\Big[c\Big(\Big(\frac{\beta p_j}{1-p_{i^*}}\Big)_{\substack{j\neq i^*\\ j\neq n}}\Big)-c\Big(\Big(\frac{\alpha p_j}{1-p_{i^*}}\Big)_{\substack{j\neq i^*\\ j\neq n}}\Big)\Big]. \tag{B.124}$$

Let $\alpha\in(0,1),\beta\in[0,1],x,y\in\operatorname{NDiag}(X^{p_{-i^*}})$ be as above. From (B.123), we by the representation since we are omitting the last index from the information function,

$$\begin{aligned}&\sum_{\substack{j=1\\ j\neq i^*}}^{n-1}\Big(\frac{\alpha p_j}{1-p_{i^*}}\Big)^\lambda u(x_j)+\Big(\frac{\alpha p_n}{1-p_{i^*}}+1-\alpha\Big)^\lambda u(x_n)+c\Big(\Big(\frac{\alpha p_j}{1-p_{i^*}}\Big)_{\substack{j\neq i^*\\ j\neq n}}\Big) \qquad\text{(B.125)}\\ =&\sum_{\substack{j=1\\ j\neq i^*}}^{n-1}\Big(\frac{\beta p_j}{1-p_{i^*}}\Big)^\lambda u(y_j)+\Big(\frac{\beta p_n}{1-p_{i^*}}+1-\beta\Big)^\lambda u(y_n)+c\Big(\Big(\frac{\beta p_j}{1-p_{i^*}}\Big)_{\substack{j\neq i^*\\ j\neq n}}\Big).\end{aligned}$$

Let $z_{i^*}\in X$ be such that for all $i\in\operatorname{supp} p_{-i^*}$, $x_i\neq z\neq y_i$. By Axiom 2.4 and (B.123), we have

$$p_{i^*}z_{i^*}+(1-p_{i^*})\Big(\alpha p_{-i^*}x+(1-\alpha)x_n\Big)\sim p_{i^*}z+(1-p_{i^*})\Big(\alpha p_{-i^*}y+(1-\alpha)y_n\Big).$$

By the representation, we have

$$\begin{aligned}&p_{i^*}^\lambda u(z_{i^*})+(1-p_{i^*})^\lambda\Big(\sum_{\substack{j=1\\ j\neq i^*}}^{n-1}\Big(\frac{\alpha p_j}{1-p_{i^*}}\Big)^\lambda u(x_j)+\Big(\frac{\alpha p_n}{1-p_{i^*}}+1-\alpha\Big)^\lambda u(x_n)\Big)+c\big((\alpha p_j)_{\substack{j\neq i^*\\ j\neq n}},p_{i^*}\big) \qquad\text{(B.126)}\\ =&p_{i^*}^\lambda u(z_{i^*})+(1-p_{i^*})^\lambda\Big(\sum_{\substack{j=1\\ j\neq i^*}}^{n-1}\Big(\frac{\beta p_j}{1-p_{i^*}}\Big)^\lambda u(y_j)+\Big(\frac{\beta p_n}{1-p_{i^*}}+1-\beta\Big)^\lambda u(y_n)\Big)+c\big((\beta p_j)_{\substack{j\neq i^*\\ j\neq n}},p_{i^*}\big)\end{aligned}$$

Thus by solving for one term in (B.125) and plugging this into (B.126), we have

$$\begin{aligned}&(1-p_{i^*})^\lambda\Big[\sum_{\substack{j=1\\ j\neq i^*}}^{n-1}\Big(\frac{\beta p_j}{1-p_{i^*}}\Big)^\lambda u(y_j)+\Big(\frac{\beta p_n}{1-p_{i^*}}+1-\beta\Big)^\lambda u(y_n)\\ &\qquad\qquad+c\Big(\Big(\frac{\beta p_j}{1-p_{i^*}}\Big)_{\substack{j\neq i^*\\ j\neq n}}\Big)-c\Big(\Big(\frac{\alpha p_j}{1-p_{i^*}}\Big)_{\substack{j\neq i^*\\ j\neq n}}\Big)\Big]+c\big((\alpha p_j)_{\substack{j\neq i^*\\ j\neq n}},p_{i^*}\big)\\ =&(1-p_{i^*})^\lambda\Big(\sum_{\substack{j=1\\ j\neq i^*}}^{n-1}\Big(\frac{\beta p_j}{1-p_{i^*}}\Big)^\lambda u(y_j)+\Big(\frac{\beta p_n}{1-p_{i^*}}+1-\beta\Big)^\lambda u(y_n)\Big)+c\big((\beta p_j)_{\substack{j\neq i^*\\ j\neq n}},p_{i^*}\big).\end{aligned}$$

So finally, we have

$$c\big((\beta p_j)_{\substack{j\neq i^*\\ j\neq n}},p_{i^*}\big)-c\big((\alpha p_j)_{\substack{j\neq i^*\\ j\neq n}},p_{i^*}\big)=(1-p_{i^*})^\lambda\Big[c\Big(\Big(\frac{\beta p_j}{1-p_{i^*}}\Big)_{\substack{j\neq i^*\\ j\neq n}}\Big)-c\Big(\Big(\frac{\alpha p_j}{1-p_{i^*}}\Big)_{\substack{j\neq i^*\\ j\neq n}}\Big)\Big].$$

Next, we show that for each $\alpha \in [0,1]$, there exists a neighborhood U_α of α such that for all $\beta \in U_\alpha \cap [0,1]$, there exist $x, y \in \mathrm{NDiag}(X^{p_{-i^*}})$ such that

$$\alpha p_{-i^*} x + (1-\alpha) x_n \sim \beta p_{-i^*} y + (1-\beta) y_n.$$

Let $\alpha \in [0,1]$. First, for all $\tilde{\alpha} \in [0,1]$,

$$\operatorname{supp} \tilde{\alpha} p_{-i^*} + (1-\tilde{\alpha})\delta_n \subseteq \operatorname{supp} p_{-i^*}. \tag{B.127}$$

Since X is open, there exist $x^1, x^2, x^3 \in \mathrm{NDiag}(X^{p_{-i^*}})$ such that for all $i \in \operatorname{supp} p_{-i^*}$, $x_i^3 > x_i^2 > x_i^1$. By Lemma B.31,

$$\alpha p_{-i^*}^3 x + (1-\alpha) x_n^3 \succ \alpha p_{-i^*} x^2 + (1-\alpha) x_n^2 \succ \alpha p_{-i^*} x^1 + (1-\alpha) x_n^1.$$

Thus if $\alpha > 0$, there exists $\alpha^1 < \alpha$ such that for all $\tilde{\alpha} \in (\alpha^1, \alpha]$,

$$\alpha p_{-i^*}^3 x + (1-\alpha) x_n^3 \succ \tilde{\alpha} p_{-i^*} x^2 + (1-\tilde{\alpha}) x_n^2 \succ \alpha p_{-i^*} x^1 + (1-\alpha) x_n^1.$$

Additionally, if $\alpha < 1$, then there exists $\alpha^2 > \alpha$ such that for all $\tilde{\alpha} \in [\alpha, \alpha^2]$,

$$\alpha p_{-i^*}^3 x + (1-\alpha) x_n^3 \succ \tilde{\alpha} p_{-i^*} x^2 + (1-\tilde{\alpha}) x_n^2 \succ \alpha p_{-i^*} x^1 + (1-\alpha) x_n^1.$$

In the case of $\alpha = 1$ denote $\alpha^2 = 2$ and in the case of $\alpha = 0$ denote $\alpha^1 = -1$. Now by the above and Lemma B.36, for all $\tilde{\alpha} \in (\alpha^1, \alpha^2) \cap [0,1]$, there exists $x^{\tilde{\alpha}} \in \mathrm{NDiag}(X^{p_{-i^*}})$ such that

$$\tilde{\alpha} p_{-i^*} x^2 + (1-\tilde{\alpha}) x_n^2 \sim \alpha p_{-i^*} x^{\tilde{\alpha}} + (1-\alpha) x_n^{\tilde{\alpha}}.$$

So we define $U_\alpha = (\alpha^1, \alpha^2)$, that is a neighborhood of α.

Next since $[0,1]$ is a compact interval and $\{U_\alpha | \alpha \in [0,1]\}$ is a cover of $[0,1]$, there exists $\{\alpha_i\}_{i=1}^m \subseteq [0,1]$ such that $\{U_{\alpha_i} | 1 \le i \le m^\alpha\}$ is a finite subcover of $[0,1]$. Thus there exists a finite sequence $(\beta_k)_{k=1}^m$ such that $\beta_1 = 1$, $\beta_m = 0$ and for each $1 \le k \le m-1$, there exists $i_k \in \{1, \dots, m^\alpha\}$ such that

$$\beta_k, \beta_{k+1} \in U_{\alpha_{i_k}}. \tag{B.128}$$

Thus we have

$$\begin{aligned}
&c\big((p_j)_{j \in \operatorname{supp} p}\big) - c(p_{i^*}) = c\big((\beta_1 p_j)_{\substack{j \neq i^* \\ j \neq n}}, p_{i^*}\big) - c\big((\beta_m p_j)_{\substack{j \neq i^* \\ j \neq n}}, p_{i^*}\big) \\
&= \sum_{k=1}^{m-1} c\big((\beta_k p_j)_{\substack{j \neq i^* \\ j \neq n}}, p_{i^*}\big) - c\big((\beta_{k+1} p_j)_{\substack{j \neq i^* \\ j \neq n}}, p_{i^*}\big) \\
&\overset{(B.124, B.128)}{=} \sum_{k=1}^{m-1} (1-p_{i^*})^\lambda \left[c\left(\left(\frac{\beta_k p_j}{1-p_{i^*}} \right)_{\substack{j \neq i^* \\ j \neq n}} \right) - c\left(\left(\frac{\beta_{k+1} p_j}{1-p_{i^*}} \right)_{\substack{j \neq i^* \\ j \neq n}} \right) \right] \\
&= (1-p_{i^*})^\lambda \left[c\left(\left(\frac{\beta_1 p_j}{1-p_{i^*}} \right)_{\substack{j \neq i^* \\ j \neq n}} \right) - c\left(\left(\frac{\beta_m p_j}{1-p_{i^*}} \right)_{\substack{j \neq i^* \\ j \neq n}} \right) \right]
\end{aligned}$$

$$= (1-p_{i^*})^\lambda \Big[c\Big(\Big(\frac{p_j}{1-p_{i^*}} \Big)_{\substack{j \neq i^* \\ j \neq n}} \Big) - c(0) \Big] = (1-p_{i^*})^\lambda c\Big(\Big(\frac{p_j}{1-p_{i^*}} \Big)_{\substack{j \neq i^* \\ j \neq n}} \Big).$$

This shows the claim by recursion. □

The previous recursivity gives us generalized entropy $H^\lambda(p)$ up to a scaling γ.

Lemma B.75 Assume that there exist a utility function $u : X \to \mathbb{R}$, $\lambda > 0$, and a linearly continuous information function $c : \Delta(X) \to \mathbb{R}$ such that $\succsim$ represented by for all lotteries $P \in \Delta(X)$,

$$\sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x) + c(P).$$

Let $\succsim$ satisfy Axioms 2.1 to 2.4. Then there exists $\gamma \in \mathbb{R}$ such that for all co-probability $p=(p_1, \dots, p_n)$,

$$c(p) = \gamma H^\lambda(p)$$

Proof. Let $p_1, p_2 \in [0,1]$ be such that $p_1 + p_2 \leq 1$. By Lemma B.74, we have

$$\begin{aligned} c(p_1, p_2) &= c(p_1) + (1-p_1)^\lambda c\Big(\frac{p_2}{1-p_1} \Big) \\ &= c(p_2) + (1-p_2)^\lambda c\Big(\frac{p_1}{1-p_2} \Big). \end{aligned}$$

Additionally the mapping

$$[0,1] \ni x \mapsto c(x)$$

is continuous as a single dimensional linearly continuous function. Thus by Daróczy (1970), there exists $\gamma \in \mathbb{R}$ such that for all $x \in [0,1]$,

$$\gamma c(x) = \begin{cases} -x \log x - (1-x) \log(1-x), & \text{if } \lambda = 1, \\ \dfrac{1}{\lambda - 1} \big(x^\lambda + (1-x)^\lambda - 1 \big), & \text{if } \lambda \neq 1, \lambda > 0. \end{cases}$$

Thus by Lemma B.74, we have for all $p \in \Delta(\mathbb{N})$,

$$c(p) = \begin{cases} \displaystyle\sum_{i \in \operatorname{supp} p} -p_i \log p_i, & \text{if } \lambda = 1, \\ \dfrac{1}{\lambda - 1} \Big(\displaystyle\sum_{i \in \operatorname{supp} p} p_i^\lambda - 1 \Big), & \text{if } \lambda \neq 1, \lambda > 0. \end{cases}$$

□

The next result shows that these axioms characterize power weighted expected utility with generalized entropy cost function. We show that this representation has an alternative power weighted expected utility representation by the uniqueness of the representation in Proposition B.80.

Proposition B.76 $\succsim$ satisfies Axioms 2.1 to 2.4 iff there exist a continuous, strictly increasing, $u : X \to \mathbb{R}$, $\lambda > 0$ and $\gamma \in \mathbb{R}$ such that for two lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x) + \gamma H^\lambda(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^\lambda u(x) + \gamma H^\lambda(Q).$$

Proof. If direction follows from Lemmas B.61, B.71, and B.75. We show the only if direction. Since $H^\lambda : \Delta(\mathbb{N}) \to \mathbb{R}$ is continuous function, it is especially linearly continuous. By Proposition B.72, we only need to show Axiom 2.4. Let $P, Q \in \Delta(X), \alpha \in (0,1), z \notin \operatorname{supp} P \cup \operatorname{supp} Q$. First, assume that $\lambda \neq 1$. We have

$$\begin{aligned}
P \succsim Q &\iff \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x) + \gamma H^\lambda(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^\lambda u(x) + \gamma H^\lambda(Q) \\
&\overset{|\cdot \alpha^\lambda}{\iff} \sum_{x \in \operatorname{supp} P} \big(\alpha P(x)\big)^\lambda u(x) + \frac{\gamma}{\lambda - 1}\Big(\sum_{x \in \operatorname{supp} P} \big(\alpha P(x)\big)^\lambda \Big) \\
&\qquad\qquad \geq \sum_{x \in \operatorname{supp} Q} \big(\alpha Q(x)\big)^\lambda u(x) + \frac{\gamma}{\lambda - 1}\Big(\sum_{x \in \operatorname{supp} Q} \big(\alpha Q(x)\big)^\lambda \Big) \\
&\iff \sum_{x \in \operatorname{supp} P} \big(\alpha P(x)\big)^\lambda u(x) + (1-\alpha)^\lambda u(z) + \frac{\gamma}{\lambda - 1}\Big(\sum_{x \in \operatorname{supp} P} \big(\alpha P(x)\big)^\lambda + (1-\alpha)^\lambda \Big) \\
&\qquad\qquad \geq \sum_{x \in \operatorname{supp} Q} \big(\alpha Q(x)\big)^\lambda u(x) + (1-\alpha)^\lambda u(z) + \frac{\gamma}{\lambda - 1}\Big(\sum_{x \in \operatorname{supp} Q} \big(\alpha Q(x)\big)^\lambda + (1-\alpha)^\lambda \Big) \\
&\iff \sum_{x \in \operatorname{supp} P} \big(\alpha P(x)\big)^\lambda u(x) + (1-\alpha)^\lambda u(z) + \gamma H^\lambda(\alpha P + (1-\alpha)z) \\
&\qquad\qquad \geq \sum_{x \in \operatorname{supp} Q} \big(\alpha Q(x)\big)^\lambda u(x) + (1-\alpha)^\lambda u(z) + \gamma H^\lambda(\alpha Q + (1-\alpha)z) \\
&\iff \alpha P + (1-\alpha)z \succsim \alpha Q + (1-\alpha)z,
\end{aligned}$$

where the third equivalence follows from adding $(1-\alpha)^\lambda u(z) + \frac{\gamma(1-\alpha)^\lambda}{\lambda - 1}$ to both sides. Second, assume that $\lambda = 1$. We have

$$\begin{aligned}
P \succsim Q &\iff \sum_{x \in \operatorname{supp} P} P(x)u(x) + \gamma H^1(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x) + \gamma H^1(Q) \\
&\overset{|+\log \alpha}{\iff} \sum_{x \in \operatorname{supp} P} P(x)u(x) + \sum_{x \in \operatorname{supp} P} P(x)\big(\log \alpha + \log P(x)\big) \\
&\qquad\qquad \geq \sum_{x \in \operatorname{supp} Q} Q(x)u(x) + \sum_{x \in \operatorname{supp} Q} Q(x)\big(\log \alpha + \log Q(x)\big) \\
&\overset{|\cdot \alpha}{\iff} \sum_{x \in \operatorname{supp} P} \alpha P(x)u(x) + \sum_{x \in \operatorname{supp} P} \alpha P(x) \log\big(\alpha P(x)\big) \geq \sum_{x \in \operatorname{supp} Q} \alpha Q(x)u(x) + \sum_{x \in \operatorname{supp} Q} \alpha Q(x) \log\big(\alpha Q(x)\big) \\
&\iff \sum_{x \in \operatorname{supp} P} \alpha P(x)u(x) + (1-\alpha)u(z) + \sum_{x \in \operatorname{supp} P} \alpha P(x) \log\big(\alpha P(x)\big) + (1-\alpha)\log(1-\alpha) \\
&\qquad\qquad \geq \sum_{x \in \operatorname{supp} Q} \alpha Q(x)u(x) + (1-\alpha)u(z) + \sum_{x \in \operatorname{supp} Q} \alpha Q(x) \log\big(\alpha Q(x)\big) + (1-\alpha)\log(1-\alpha)
\end{aligned}$$

$$\iff \sum_{x\in\operatorname{supp} P} \alpha P(x)u(x) + (1-\alpha)u(z) + \gamma H^1\big(\alpha P + (1-\alpha z)\big)$$
$$\geq \sum_{x\in\operatorname{supp} Q} \alpha Q(x)u(x) + (1-\alpha)u(z) + \gamma H^1\big(\alpha Q + (1-\alpha z)\big)$$
$$\iff \alpha P + (1-\alpha z) \succsim \alpha Q + (1-\alpha z),$$

where the fourth equivalence follows from adding $(1-\alpha)u(z) + (1-\alpha)\log(1-\alpha)$ to both sides. □

Finally, we show under Restricted FOSD, we have entropy modified expected utility.

Proposition B.77 $\succsim$ satisfies Axioms 2.1-2.5 iff there exist a continuous, strictly increasing, $u : X \to \mathbb{R}$ and $\gamma \in \mathbb{R}$ such that for two lotteries P, Q,

$$P \succsim Q \iff \sum_{x\in\operatorname{supp} P} P(x)u(x) - \gamma H^1(P) \geq \sum_{x\in\operatorname{supp} Q} Q(x)u(x) - \gamma H^1(Q).$$

Proof. If direction follows from Lemmas B.61, B.62, B.71, and B.75. The only if direction follows from Propositions B.73 and B.76 □

B.5 Uniqueness of Variational Expected Utility

In this section, we consider the uniqueness of the power weighted variational expected utility. We show Theorems 2.6 and 2.7 and Corollary 2.8 As a corollary, this shows that any power weighted expected utility with a generalized entropy cost function H^λ has an alternative power weighted expected utility representation and finish up the proof for Theorem 2.1 from Proposition B.76.

Proposition B.78 Let $\succsim \subseteq \Delta(X)\times\Delta(X)$, strictly increasing and continuous $u, \tilde{u} : X \to \mathbb{R}, \lambda, \tilde{\lambda} > 0$, and linearly continuous information functions $c, \tilde{c} : \Delta(X) \to \mathbb{R}$ be such that for all lotteries $P, Q \in \Delta(X)$,

$$P \succsim Q \iff \sum_{x\in\operatorname{supp} P} P(x)^\lambda u(x) + c(P) \geq \sum_{x\in\operatorname{supp} Q} Q(x)^\lambda u(x) + c(Q)$$
$$\iff \sum_{x\in\operatorname{supp} P} P(x)^{\tilde{\lambda}} \tilde{u}(x) + \tilde{c}(P) \geq \sum_{x\in\operatorname{supp} Q} Q(x)^{\tilde{\lambda}} \tilde{u}(x) + \tilde{c}(Q).$$

Then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that for all $x \in X, P \in \Delta(X)$,

$$\tilde{\lambda} = \lambda, \tilde{u}(x) = \alpha u(x) + \beta, \text{ and } \tilde{c}(P) = \alpha c(P) + \Big(1 - \sum_{x\in\operatorname{supp} P} P(x)^\lambda\Big)\beta.$$

Proof. Consider lottery $p = (1/2, 1/4, 1/4)$ and the induced preferences $\succsim^p$ on $\operatorname{NDiag}(X^p)$. $(p_1^\lambda u, p_2^\lambda u, p_3^\lambda u)$ and $(p_1^{\tilde{\lambda}}\tilde{u}, p_2^{\tilde{\lambda}}\tilde{u}, p_3^{\tilde{\lambda}}\tilde{u})$ give an additive representation for $\succsim^p$. Since u is continuous, by the uniqueness of an additive representation for $\succsim^p$ (Krantz et al., 1971, Theorem 13), there exist $\alpha^* > 0$ and $\beta^*, \beta_2, \beta_3 \in \mathbb{R}$ such that

$$p_1^\lambda u = \alpha^* p_1^{\tilde{\lambda}}\tilde{u} + \beta^*, p_2^\lambda u = \alpha^* p_2^{\tilde{\lambda}}\tilde{u} + \beta_2, p_3^\lambda u = \alpha^* p_3^{\tilde{\lambda}}\tilde{u} + \beta_3.$$

Since X is non-trivial and u is strictly increasing, we have

$$p_1^{\lambda-\tilde{\lambda}}u - \beta^* = \alpha^*\tilde{u}$$

and so

$$p_2^{\lambda}u = p_2^{\tilde{\lambda}}p_1^{\lambda-\tilde{\lambda}}u - p_2^{\tilde{\lambda}}\beta^* + \beta_2$$

and so especially since this holds for all $x \in X$ and $p_1 = 1/2 = 2^{-1}, p_2 = 1/4 = 2^{-2}$,

$$p_2^{\lambda-\tilde{\lambda}} = p_1^{\lambda-\tilde{\lambda}} \iff 2^{2(\tilde{\lambda}-\lambda)} = 2^{\tilde{\lambda}-\lambda} \iff 2(\tilde{\lambda}-\lambda) = \tilde{\lambda}-\lambda \iff \tilde{\lambda} = \lambda.$$

So we have $u = \alpha^*\tilde{u} + \beta^*$.

Finally, we show that for all $P \in \Delta(X)$,

$$c(P) = \alpha^*\tilde{c}(P) + \Big(1 - \sum_{x\in\operatorname{supp} P} P(x)^{\lambda}\Big)\beta^*.$$

Assume, per contra, that there exists P such that

$$c(P) = \alpha^*\tilde{c}(P) + \Big(1 - \sum_{x\in\operatorname{supp} P} P(x)^{\lambda}\Big)\beta^*.$$

Let $x^0 \in X$ and denote

$$\alpha^{\ddagger} = \inf\Big\{\alpha\in[0,1]\Big|c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) \neq \alpha^*\tilde{c}\big(\alpha P + (1-\alpha)\delta_{x^0}\big) + \Big(1 - \sum_{x\in\operatorname{supp} P\cup\{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^{\lambda}\Big)\beta^*\Big\}.$$

First, we show that if $\alpha^{\dagger} \in [0,1)$ is such that

$$c\big(\alpha^{\dagger}P + (1-\alpha^{\dagger})\delta_{x^0}\big) = \alpha^*\tilde{c}\big(\alpha^{\dagger}P + (1-\alpha^{\dagger})\delta_{x^0}\big) + \Big(1 - \sum_{x\in\operatorname{supp} P\cup\{x_0\}} \big(\alpha^{\dagger}P(x) + (1-\alpha^{\dagger})\delta_{x^0}(x)\big)^{\lambda}\Big)\beta^*, \quad \text{(B.129)}$$

then there exists $\alpha^1 \in (\alpha^{\dagger}, 1)$ such that for all $\alpha \in [\alpha^{\dagger}, \alpha^1)$,

$$c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) = \alpha^*\tilde{c}\big(\alpha P + (1-\alpha)\delta_{x^0}\big) + \Big(1 - \sum_{x\in\operatorname{supp} P\cup\{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^{\lambda}\Big)\beta^*. \quad \text{(B.130)}$$

Let $p^{\dagger} \in \Delta(\mathbb{N})$ and $x^{\dagger} \in \operatorname{NDiag}(X^X)$ be such that $\alpha^{\dagger}P + (1-\alpha^{\dagger})\delta_{x^0} = (p^{\dagger}, x^{\dagger})$. Without loss of generality assume that $1 \in \operatorname{supp} p^{\dagger}$ by permuting indices. Since X is open, there exist $x_1^1, x_1^2 \in X$ such that $x_1^2 > x_1^{\dagger} > x_1^1$ and for all $x_1 \in [x_1^1, x_1^2]$, $(x_1, x_{-1}^{\dagger}) \in \operatorname{NDiag}(X^{p^{\dagger}})$.

Now the function

$$[0,1] \ni \alpha \mapsto \sum_{x\in\operatorname{supp} P\cup\{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^{\lambda}u(x)$$

is continuous. Thus there exists $\delta > 0$ such that $\delta < 1 - \alpha^{\dagger}$ and for all $\alpha \in [\alpha^{\dagger}, \alpha^{\dagger} + \delta)$,

$$\Big|\sum_{x\in\operatorname{supp} P\cup\{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^{\lambda}u(x) - \sum_{x\in\operatorname{supp} P\cup\{x_0\}} \big(\alpha^{\dagger}P(x) + (1-\alpha^{\dagger})\delta_{x^0}(x)\big)^{\lambda}u(x)\Big| < \frac{p_1^{\dagger}}{2}\min\big\{u(x_1^2) - u(x_1^{\dagger}), u(x_1^{\dagger}) - u(x_1^1)\big\}.$$

Additionally since c is linearly continuous, there exists $\alpha^\dagger < \alpha^1 < \alpha^\dagger + \delta$ such that for all $\alpha \in [\alpha^\dagger, \alpha^1)$,

$$\left|c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) - c\big(\alpha^\dagger P + (1-\alpha^\dagger)\delta_{x^0}\big)\right| < \frac{p_1^\dagger}{2} \min\big\{u(x_1^2) - u(x_1^\dagger), u(x_1^\dagger) - u(x_1^1)\big\}.$$

Now we have for all $\alpha \in (\alpha^\dagger, \alpha^1)$, by denoting $P\alpha\delta_{x^0} = \alpha P + (1-\alpha)\delta_{x^0}$ and $P\alpha^\dagger\delta_{x^0} = \alpha^\dagger P + (1-\alpha^\dagger)\delta_{x^0}$,

$$\begin{aligned}
&\sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(P\alpha\delta_{x^0}(x)\big)^\lambda u(x) + c\big(P\alpha\delta_{x^0}\big) - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(P\alpha^\dagger\delta_{x^0}(x)\big)^\lambda u(x) - c\big(P\alpha^\dagger\delta_{x^0}\big) \\
&\quad \le \Big| \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(P\alpha\delta_{x^0}(x)\big)^\lambda u(x) - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(P\alpha^\dagger\delta_{x^0}(x)\big)^\lambda u(x)\Big| + \big|c\big(P\alpha\delta_{x^0}\big) - c\big(P\alpha^\dagger\delta_{x^0}\big)\big| \\
&\quad < p_1^\dagger\big(u(x_1^2) - u(x_1^\dagger)\big)
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(P\alpha\delta_{x^0}(x)\big)^\lambda u(x) + c\big(P\alpha\delta_{x^0}\big) - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(P\alpha^\dagger\delta_{x^0}(x)\big)^\lambda u(x) - c\big(P\alpha^\dagger\delta_{x^0}\big) \\
&\quad \ge -\Big| \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(P\alpha\delta_{x^0}(x)\big)^\lambda u(x) - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(P\alpha^\dagger\delta_{x^0}(x)\big)^\lambda u(x)\Big| - \big|c\big(P\alpha\delta_{x^0}\big) - c\big(P\alpha^\dagger\delta_{x^0}\big)\big| \\
&\quad > -p_1^\dagger\big(u(x_1^\dagger) - u(x_1^1)\big) = p_1^\dagger\big(u(x_1^1) - u(x_1^\dagger)\big)
\end{aligned}$$

By the continuity of u, for all $\alpha \in (\alpha^\dagger, \alpha^1)$, there exists $x_1^\alpha \in X$, $x_1^2 \ge x_1^\alpha \ge x_1^1$ such that

$$\sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^\lambda u(x) + c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) = \sum_{i \in \operatorname{supp} p^\dagger} (p_i^\dagger)^\lambda u\big((x_1^\alpha, x_{-1}^\dagger)_i\big) + c\big(p^\dagger\big). \tag{B.131}$$

Let $\alpha \in (\alpha^\dagger, \alpha^1)$. First, we have

$$\begin{aligned}
&\sum_{i \in \operatorname{supp} p^\dagger} (p_i^\dagger)^\lambda \Big(\alpha^* u\big((x_1^\alpha, x_{-1}^\dagger)_i\big) + \beta^*\Big) + \tilde{c}\big(p^\dagger\big) \\
&\quad \overset{\text{(B.129)}}{=} \sum_{i \in \operatorname{supp} p^\dagger} (p_i^\dagger)^\lambda \Big(\alpha^* u\big((x_1^\alpha, x_{-1}^\dagger)_i\big) + \beta^*\Big) + \alpha^* c\big(p^\dagger\big) + \Big(1 - \sum_{i \in \operatorname{supp} p^\dagger} (p_i^\dagger)^\lambda\Big)\beta^* \\
&\quad = \alpha^*\Big(\sum_{i \in \operatorname{supp} p^\dagger} (p_i^\dagger)^\lambda u\big((x_1^\alpha, x_{-1}^\dagger)_i\big) + c\big(p^\dagger\big)\Big) + \beta^*.
\end{aligned}$$

Now we have by the alternative representation and (B.131),

$$\begin{aligned}
&\sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^\lambda (\alpha^* u(x) + \beta^*) + \tilde{c}\big(\alpha P + (1-\alpha)\delta_{x^0}\big) \\
&\quad = \sum_{i \in \operatorname{supp} p^\dagger} (p_i^\dagger)^\lambda \Big(\alpha^* u\big((x_1^\alpha, x_{-1}^\dagger)_i\big) + \beta^*\Big) + \tilde{c}\big(p^\dagger\big) = \alpha^*\Big(\sum_{i \in \operatorname{supp} p^\dagger} (p_i^\dagger)^\lambda u\big((x_1^\alpha, x_{-1}^\dagger)_i\big) + c\big(p^\dagger\big)\Big) + \beta^* \\
&\quad \overset{\text{(B.131)}}{=} \alpha^* \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^\lambda u(x) + c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) + \beta^*
\end{aligned}$$

So we have

$$\sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^\lambda \beta^* + \tilde{c}\big(\alpha P + (1-\alpha)\delta_{x^0}\big) = c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) + \beta^*$$

$$\Rightarrow \tilde{c}\big(\alpha P + (1-\alpha)\delta_{x^0}\big) = c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) + \Big(1 - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^\lambda\Big)\beta^*.$$

Since in the above $\alpha \in (\alpha^\dagger, \alpha^1)$ was arbitrary, this holds for all $\alpha \in [\alpha^\dagger, \alpha^1)$ and shows the claim.

By the counter assumption, $\alpha^\ddagger \in [0,1]$. We show that $\alpha^\ddagger = \infty$.

We first show that $\alpha^\ddagger > 0$. This follows (B.130) and since $c(\delta_{x^0}) = \tilde{c}(\delta_{x^0}) = 0$ and $1 - 1^\lambda = 0$.

Next assume that $0 < \alpha^\ddagger \leq 1$. Now we have for all $\alpha \in (0, \alpha^\dagger)$,

$$c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) = \alpha^* \tilde{c}\big(\alpha P + (1-\alpha)\delta_{x^0}\big) + \Big(1 - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^\lambda\Big)\beta^*.$$

Since $c, \tilde{c}$ are linearly continuous and the function

$$[0,1] \ni \alpha \mapsto \Big(1 - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^\lambda\Big)\beta^*$$

is continuous, we can take $\alpha \to \alpha^\dagger$ to get

$$c\big(\alpha^\dagger P + (1-\alpha^\dagger)\delta_{x^0}\big) = \alpha^* \tilde{c}\big(\alpha^\dagger P + (1-\alpha^\dagger)\delta_{x^0}\big) + \Big(1 - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha^\dagger P(x) + (1-\alpha^\dagger)\delta_{x^0}(x)\big)^\lambda\Big)\beta^*.$$

Now if $\alpha^\ddagger = 1$, this is a contradiction so assume that $\alpha^\ddagger < 1$. By (B.130), there exists $\alpha^1 \in (\alpha^\ddagger, 1)$ such that for all $\alpha \in [\alpha^\ddagger, \alpha^1)$,

$$c\big(\alpha P + (1-\alpha)\delta_{x^0}\big) = \alpha^* \tilde{c}\big(\alpha P + (1-\alpha)\delta_{x^0}\big) + \Big(1 - \sum_{x \in \operatorname{supp} P \cup \{x_0\}} \big(\alpha P(x) + (1-\alpha)\delta_{x^0}(x)\big)^\lambda\Big)\beta^*$$

which contradicts the definition of $\alpha^\ddagger$. This shows that $\alpha^\ddagger = \infty$ and so for all $P \in \Delta(X)$,

$$c(P) = \alpha^* \tilde{c}(P) + \Big(1 - \sum_{x \in \operatorname{supp} P} P(x)^\lambda\Big)\beta^*$$

□

As a corollary, we get the uniqueness for power weighted expected utility with generalized entropy cost function.

Proposition B.79 Let $\succsim \subseteq \Delta(X) \times \Delta(X)$, strictly increasing and continuous $u, \tilde{u} : X \to \mathbb{R}$, $\lambda, \tilde{\lambda} > 0$, and $\gamma, \tilde{\gamma} \in \mathbb{R}$ be such that for all lotteries $P, Q \in \Delta(X)$,

$$\begin{aligned} P \succsim Q &\iff \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x) + \gamma H^\lambda(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^\lambda u(x) + \gamma H^\lambda(Q) \\ &\iff \sum_{x \in \operatorname{supp} P} P(x)^{\tilde{\lambda}} \tilde{u}(x) + \tilde{\gamma} H^{\tilde{\lambda}}(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^{\tilde{\lambda}} \tilde{u}(x) + \tilde{\gamma} H^{\tilde{\lambda}}(Q). \end{aligned}$$

Then there exist $\alpha > 0, \beta \in \mathbb{R}$ such that for all $x \in X, P \in \Delta(X)$,

$$\tilde{\lambda} = \lambda, \tilde{u}(x) = \alpha u(x) + \beta, \text{ and } \tilde{\gamma} = \alpha\gamma + (1-\lambda)\beta.$$

Proof. By Proposition B.78, there exist $\alpha^* > 0, \beta^* \in \mathbb{R}$ such that for all $x \in X$, $P \in \Delta(X)$

$$\tilde{\lambda} = \lambda, \tilde{u}(x) = \alpha^* u(x) + \beta^*, \text{ and } \tilde{\gamma} H^\lambda(P) = \alpha^* \gamma H^\lambda(P) + \Big(1 - \sum_{x \in \operatorname{supp} P} P(x)^\lambda\Big)\beta^*.$$

Let $P \in \Delta(X)$. We have

$$\tilde{\gamma} H^\lambda(P) = \frac{\tilde{\gamma}}{\lambda - 1}\Big(\sum_{x \in \operatorname{supp} P} P(x)^\lambda - 1\Big) = \alpha^* \gamma H^\lambda(P) + \Big(1 - \sum_{x \in \operatorname{supp} P} P(x)^\lambda\Big)\beta^*$$
$$= \frac{\alpha^* \gamma + (1-\lambda)\beta^*}{\lambda - 1}\Big(\sum_{x \in \operatorname{supp} P} P(x)^\lambda - 1\Big)$$

So

$$\tilde{\gamma} = \alpha^* \gamma + (1 - \lambda)\beta^*.$$

□

This gives especially that any power weighted expected utility with generalized entropy cost function has an alternative representation without cost function.

Proposition B.80 $\succsim$ satisfies Axioms 2.1 to 2.4 and 2.10 and does not satisfy Axiom 2.5 iff there exist a continuous, strictly increasing, $u : X \to \mathbb{R}$, $\lambda > 0$, $\lambda \neq 1$ such that for two simple lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^\lambda u(x).$$

Proof. The only if direction follows from Propositions B.76 and B.77. We show the if direction. By Proposition B.76, there exists a continuous, strictly increasing, $u : X \to \mathbb{R}$, $\lambda > 0$ and $\gamma \in \mathbb{R}$ such that for two simple lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} P(x)^\lambda u(x) + \gamma H^\lambda(P) \geq \sum_{x \in \operatorname{supp} Q} Q(x)^\lambda u(x) + \gamma H^\lambda(Q).$$

By Proposition B.77, $\lambda \neq 1$. By Proposition B.79 and since $\lambda \neq 1$, there exists an alternative representation using utility $\tilde{u} = u + \frac{\gamma}{\lambda - 1}$ and $\tilde{\gamma} = 0$ which shows the claim. □

B.6 Normalized Power Weighted Expected Utility

In this section, we show Theorem 2.4 and consider preferences that satisfy Restricted Common Consequence Independence and CE Bounds. This section builds on Appendix B.2. First, we show that under CE Bounds and continuity in probabilities and lottery has a certainty equivalent.

Lemma B.81 Let $\succsim$ satisfy Axioms 2.1 to 2.3 and 2.11. Then for each $P \in \Delta(X)$, there exists $z \in X$ such that $P \sim \delta_z$.

Proof. Let $x^* = \max\{y \in \operatorname{supp} P\}$, $x_* = \min\{z \in \operatorname{supp} P\}$ and $\tilde{x} = 1/2 x^* + 1/2 x_*$ where the mixture is taken in X. Let $x^*, \tilde{x}, x_* \in X$ be such that $x^* > \tilde{x} > x_*$. Assume, per contra, that there exists $P \in \Delta(X)$ such that for all $z \in X$, $P \not\sim \delta_z$. Assume w.l.o.g. $P \succ \delta_{x_*}$.

Consider lotteries for all $\alpha \in [0,1]$, $\alpha P + (1-\alpha)\delta_{\tilde{x}}$. Denote the co-probability of this lottery by p^α. Denote

$$B = \{\alpha \in [0,1] | \text{for all } z \in [x_*, x^*], z \not\sim \alpha P + (1-\alpha)\delta_{\tilde{x}}\}.$$

We will show that $B = \varnothing$. Assume, per contra, $B \neq \varnothing$. Let $\alpha^* = \inf B$. First, we show that $\alpha^* \notin B$. If $\alpha^* = 0$, then we are done since $x^* > \tilde{x} > x_*$. So $\alpha^* > 0$. Now, for all $\alpha \in [0, \alpha^*)$, there exists $z^\alpha \in [x_*, x^*]$ such that $z^\alpha \sim \alpha P + (1-\alpha)\delta_{\tilde{x}}$. Since the set $D = \{\tilde{\alpha} \in [0,1] | \delta_{x^*} \succsim \alpha P + (1-\alpha)\delta_{\tilde{x}} \succsim \delta_{x_*}\}$ is closed by Axiom 2.3 and $[0, \alpha^*) \subseteq D$ so especially $[0, \alpha^*] = \mathrm{Cl}[0, \alpha^*) \subseteq \mathrm{Cl}\, D = D$, where Cl denotes the closure. Thus $\alpha^* \notin B$. Now if $\alpha^* = 1$, we are done. So assume that $\alpha^* < 1$.

Next, we show that there exists $\varepsilon > 0$ such that for all $\alpha \in [\alpha^*, \alpha^* + \varepsilon)$, $\alpha \notin B$. Since X is an open set, there exist $x^{**} > x^*$ and $x_* > x_{**}$. By Axiom 2.3, the set $E = \{\tilde{\alpha} \in [0,1] | \delta_{x^{**}} \succ \alpha P + (1-\alpha)\delta_{\tilde{x}} \succ \delta_{x_{**}}$ is open. Since $\alpha^* \in E$, there exists $\varepsilon > 0$ such that for all $\alpha \in (\alpha^*, \alpha^* + \varepsilon)$, $\alpha \in E$. Let $\alpha \in (\alpha^*, \alpha^* + \varepsilon)$. Now we have $\delta_{x^{**}} \succ \alpha P + (1-\alpha)\delta_{\tilde{x}} \succ \delta_{x_{**}}$ and so by Axiom 2.2, there exists $z^\alpha \in [x^{**}, x_{**}]$ such that $z^\alpha \sim \alpha P + (1-\alpha)\delta_{\tilde{x}}$. Thus by Axiom 2.11, $x^* \geq z^\alpha \geq x_*$. So for all $\alpha \in (\alpha^*, \alpha^* + \varepsilon)$, we have $\alpha \notin B$. This contradicts the choice that $\alpha^* = \inf B$. Thus $B = \varnothing$ which shows the claim. □

The next lemma shows that since the certainty equivalents always exist by above, they will give the desired representation for lotteries with at least 4 prizes.

Lemma B.82 Let $\succsim$ satisfy Axioms 2.1 to 2.3, 2.10, and 2.11. Then there exists $\lambda > 0$ such that for all simple lotteries $(p, x), (q, y)$ with $p, q \in \Delta^4(\mathbb{N})$,

$$(p,x) \succsim (q,y) \iff \sum_{i \in \operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j \in \operatorname{supp} p} p_j^\lambda} u(x_i) \geq \sum_{i \in \operatorname{supp} q} \frac{q_i^\lambda}{\sum_{j \in \operatorname{supp} q} q_j^\lambda} u(y_i)$$

Proof. Let $p \in \Delta^4(\mathbb{N})$. By Lemma B.81, for all $x \in X^{\operatorname{supp} p}$, there exists $c_x^p \in X$ such that

$$c_x^p \sim (p, x).$$

Define $\varphi(p,x) = u(c_x^p)$. Now φ represents $\widetilde{\succsim}^p$. By Lemma B.61, there exists a strictly increasing function ψ such that for all $x \in X^{\operatorname{supp} p}$,

$$u(c_x^p) = \varphi(p,x) = \psi\Big(\sum_{i \in \operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j \in \operatorname{supp} p} p_j^\lambda} u(x_i) \Big).$$

We show that ψ is the identity function.

Let $x \in X^{\operatorname{supp} p}$. Now for all $\varepsilon > 0$, since X is open, there exist $x^*, x_* \in \mathrm{NDiag}(X^p)$ such that for all $i \in \operatorname{supp} p$,

$$c_x^p + \varepsilon > x_i^{*,\varepsilon} > c_x^p > x_{*,i}^{\varepsilon} > c_x^p - \varepsilon.$$

Thus by Axioms 2.1 and 2.11,

$$c_x^p + \varepsilon \succsim \max_{i\in\operatorname{supp} p} x_i^{*,\varepsilon} \succsim (p, x^{*,\varepsilon}) \sim c_{x^{*,\varepsilon}} \succsim \min_{i\in\operatorname{supp} p} x_i^{*,\varepsilon} \succ c_x^p$$

$$\succ \max_{i\in\operatorname{supp} p} x_{*,i}^{\varepsilon} \succsim c_{x_*^\varepsilon} \sim (p, x_*^\varepsilon) \succsim \min_{i\in\operatorname{supp} p} x_{*,i}^{\varepsilon} \succsim c_x^p - \varepsilon.$$

First, since ψ and u are increasing, we have

$$\psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j\in\operatorname{supp} p} p_j^\lambda} u(x_i^{*,\varepsilon})\Big) \geq \psi(u(c_x^p)) \geq \psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j\in\operatorname{supp} p} p_j^\lambda} u(x_{*,i}^{\varepsilon})\Big).$$

and by the representation for $\mathrel{\widetilde{\succsim}}^p$ and since $x^{*,\varepsilon} \mathrel{\widetilde{\succsim}}^p x \mathrel{\widetilde{\succsim}}^p x_*^\varepsilon$,

$$\psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_j p_j^\lambda} u(x_i^{*,\varepsilon})\Big) \geq \psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_j p_j^\lambda} u(x_i)\Big) \geq \psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_j p_j^\lambda} u(x_{*,i}^{\varepsilon})\Big).$$

Thus for all $\varepsilon > 0$,

$$u(c_x^p + \varepsilon) \geq u(c_{x^{*,\varepsilon}}) = \psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j\in\operatorname{supp} p} p_j^\lambda} u(x_i^{*,\varepsilon})\Big)$$

$$\geq \psi(u(c_x^p)) \geq \psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j\in\operatorname{supp} p} p_j^\lambda} u(x_{*,i}^{\varepsilon})\Big) = u(c_{x_*^\varepsilon}) \geq u(c_x^p - \varepsilon)$$

and symmetrically

$$u(c_x^p + \varepsilon) \geq \psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j\in\operatorname{supp} p} p_j^\lambda} u(x_i)\Big) \geq u(c_x^p - \varepsilon).$$

Now by taking $\varepsilon \to 0$, we have by the continuity of u,

$$u(c_x^p) \geq \psi(u(c_x^p)) \geq u(c_x^p) \Rightarrow \psi(u(c_x^p)) = u(c_x^p)$$

and

$$u(c_x^p) \geq \psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j\in\operatorname{supp} p} p_j^\lambda} u(x_i)\Big) \geq u(c_x^p) \Rightarrow \psi\Big(\sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j\in\operatorname{supp} p} p_j^\lambda} u(x_i)\Big) = u(c_x^p)$$

Since in the above $x \in \operatorname{NDiag}(X^p)$ was arbitrary, this shows that ψ is an identity function.

Finally, since in the above $p \in \Delta^4(\mathbb{N})$ was arbitrary, we have for all $(p, x), (q, y)$ with $p, q \in \Delta^4(\mathbb{N})$, since $c_x^p \sim (p, x), c_y^q \sim (q, y)$,

$$(p, x) \succsim (q, y) \iff c_x^p \succsim c_y^p \iff u(c_x^p) \geq u(c_y^p)$$

$$\iff \sum_{i\in\operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j\in\operatorname{supp} p} p_j^\lambda} u(x_i) \geq \sum_{i\in\operatorname{supp} q} \frac{q_i^\lambda}{\sum_{j\in\operatorname{supp} q} q_j^\lambda} u(y_i).$$

□

Finally, we extend the previous representation for all lotteries by continuity. Second, we show the necessity of the axioms.

Proposition B.83 $\succsim$ satisfies Axioms 2.1 to 2.3, 2.10, and 2.11 iff there exist a continuous, strictly increasing, $u : X \to \mathbb{R}$ and $\lambda > 0$ such that for all lotteries P, Q,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) \geq \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x).$$

Proof. We will first show the if direction. By Lemma B.82, there exist a continuous and strictly increasing $u : X \to \mathbb{R}$ and $\lambda > 0$ such that for all $P, Q \in \Delta(X)$ with $|\operatorname{supp} P|, |\operatorname{supp} Q| \geq 4$,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) \geq \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x).$$

We will show that for all lotteries $P, Q \in \Delta(X)$,

$$P \succsim Q \iff \sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) \geq \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x)$$

by showing the equivalent claim that

$$P \succsim Q \Rightarrow \sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) \geq \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x) \tag{B.132}$$

and

$$P \succ Q \Rightarrow \sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) > \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x).$$

Assume that $P \succsim Q$. Let $p \in \Delta(\mathbb{N})$, $x \in \operatorname{NDiag}(X^p)$ be such that $P = (p, x)$. Since X is open, there exists $(y^j)_{j=1}^\infty \in \operatorname{NDiag}(X^p)$ such that for all $j \in \mathbb{N}$ and $i \in \operatorname{supp} p$, $y_i^j > x_i$ and $\lim_{j \to \infty} y^j = x$. Let $j^* \in \mathbb{N}$ and denote $P^{j^*} = (p, y^{j^*})$. By Lemma B.31, $P^{j^*} \succ (p, x) \succsim Q$. By Axiom 2.3, there exists $R \in \Delta(X)$ such that $P^{j^*} \succ R \succ Q$. Let $S \in \Delta(X)$ be such that $|\operatorname{supp} S| \geq 4$. By Axiom 2.3, there exists $\alpha^1 \in (0, 1)$ such that for all $\alpha \in (0, \alpha^1)$,

$$\alpha S + (1 - \alpha) P^{j^*} \succ R \succ \alpha S + (1 - \alpha) Q.$$

Now $\left|\operatorname{supp} \alpha S + (1 - \alpha) P^{j^*}\right|, |\operatorname{supp} \alpha S + (1 - \alpha) Q| \geq 4$ and so by Lemma B.82,

$$\frac{\sum_{x \in \operatorname{supp} P^{j^*} \cup S} \left(\alpha S(x) + (1 - \alpha) P^{j^*}(x)\right)^\lambda u(x)}{\sum_{y \in \operatorname{supp} P^{j^*} \cup S} \left(\alpha S(y) + (1 - \alpha) P^{j^*}(y)\right)^\lambda} > \frac{\sum_{x \in \operatorname{supp} Q \cup S} \left(\alpha S(x) + (1 - \alpha) Q(x)\right)^\lambda u(x)}{\sum_{y \in \operatorname{supp} Q \cup S} \left(\alpha S(y) + (1 - \alpha) Q(y)\right)^\lambda}. \tag{B.133}$$

Since for all $\alpha \in [0, 1]$,

$$\sum_{x \in \operatorname{supp} P^{j^*} \cup S} \left(\alpha S(x) + (1 - \alpha) P^{j^*}(x)\right)^\lambda, \sum_{x \in \operatorname{supp} Q \cup S} \left(\alpha S(x) + (1 - \alpha) Q(x)\right)^\lambda > 0,$$

in (B.133) the left-hand side and right-hand side are continuous as a function of $\alpha \in [0, 1]$ as a division of two continuous functions. So we can take the limit of $\alpha \to 0$, and by the definition of P^{j^*}, we have

$$\sum_{i \in \operatorname{supp} p} \frac{p_i^\lambda}{\sum_{k \in \operatorname{supp} p} p_k^\lambda} u(y_i^{j^*}) \geq \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x).$$

Since in the above $j^* \in \mathbb{N}$ was arbitrary, we have for all $j \in \mathbb{N}$,

$$\sum_{i \in \operatorname{supp} p} \frac{p_i^\lambda}{\sum_{k \in \operatorname{supp} p} p_k^\lambda} u(y_i^j) \geq \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x).$$

So by taking $j \to \infty$ and by the continuity of u, we have

$$\sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) = \sum_{i \in \operatorname{supp} p} \frac{p_i^\lambda}{\sum_{k \in \operatorname{supp} p} p_k^\lambda} u(x_i) \geq \sum_{x \in \operatorname{supp} Q} \frac{Q(x)^\lambda}{\sum_{y \in \operatorname{supp} Q} Q(y)^\lambda} u(x),$$

which shows the claim.

Next assume that $P \succ Q$. Let $q \in \Delta(\mathbb{N})$, $y \in \operatorname{NDiag}(X^q)$ be such that $Q = (q, y)$. Since X is open, by Axiom 2.2, there exists $y^* \in \operatorname{NDiag}(X^q)$ such that for all $i \in \operatorname{supp} q$, $y_i^* \geq y_i$, for some $i^* \in \operatorname{supp} p$, $y_{i^*}^* > y_{i^*}$, and $P \succ (q, y^*)$. By (B.132) and since u is strictly increasing, we have

$$\sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) \geq \sum_{i \in \operatorname{supp} q} \frac{q_i^\lambda}{\sum_{k \in \operatorname{supp} q} q_k^\lambda} u(y_i^*) > \sum_{i \in \operatorname{supp} q} \frac{q_i^\lambda}{\sum_{k \in \operatorname{supp} q} q_k^\lambda} u(y_i),$$

which shows the claim since $(q, y) = Q$.

Next show the only if direction. Axioms 2.1 and 2.2 follow from the representation with strictly increasing and continuous utility.

Axiom 2.3: For each $P, Q \in \Delta(X)$, the representing function is as a function of α,

$$[0, 1] \ni \alpha \mapsto \frac{\sum_{x \in \operatorname{supp} P \cup \operatorname{supp} Q} (\alpha P(x) + (1 - \alpha) Q(x))^\lambda u(x)}{\sum_{y \in \operatorname{supp} P \cup \operatorname{supp} Q} (\alpha P(y) + (1 - \alpha) Q(y))^\lambda}$$

is continuous as a division of two continuous functions and since for all $\alpha \in [0, 1]$,

$$\sum_{y \in \operatorname{supp} P \cup \operatorname{supp} Q} (\alpha P(y) + (1 - \alpha) Q(y))^\lambda > 0.$$

Axiom 2.3 follows from this.

Axiom 2.10: Let $p \in \Delta(\mathbb{N}), x, y \in \operatorname{NDiag}(X^p), \alpha \in (0, 1), z \in X$ such that for all $i \in \operatorname{supp} p$, $x_i \neq z \neq y_i$. We have

$$\begin{aligned}
(p, x) \succsim (p, y) &\iff \sum_{i \in \operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j \in \operatorname{supp} p} p_j^\lambda} u(x_i) \geq \sum_{i \in \operatorname{supp} p} \frac{p_i^\lambda}{\sum_{j \in \operatorname{supp} p} p_j^\lambda} u(y_i) \\
&\iff \sum_{i \in \operatorname{supp} p} p_i^\lambda u(x_i) \geq \sum_{i \in \operatorname{supp} p} p_i^\lambda u(y_i) \\
&\iff \sum_{i \in \operatorname{supp} p} (\alpha p_i)^\lambda u(x_i) + (1 - \alpha)^\lambda u(z) \geq \sum_{i \in \operatorname{supp} p} (\alpha p_i)^\lambda u(y_i) + (1 - \alpha)^\lambda u(z) \\
&\iff \frac{\sum_{i \in \operatorname{supp} p} (\alpha p_i)^\lambda u(x_i) + (1 - \alpha)^\lambda u(z)}{\sum_{j \in \operatorname{supp} p} (\alpha p_j)^\lambda + (1 - \alpha)^\lambda} \geq \frac{\sum_{i \in \operatorname{supp} p} (\alpha p_i)^\lambda u(y_i) + (1 - \alpha)^\lambda u(z)}{\sum_{j \in \operatorname{supp} p} (\alpha p_j)^\lambda + (1 - \alpha)^\lambda} \\
&\iff \alpha(p, x) + (1 - \alpha) z \succsim \alpha(p, y) + (1 - \alpha) z.
\end{aligned}$$

Axiom 2.11: Follows from the observation that for all $P \in \Delta(X)$,

$$\max_{x \in \operatorname{supp} P} u(x) \geq \sum_{x \in \operatorname{supp} P} \frac{P(x)^\lambda}{\sum_{y \in \operatorname{supp} P} P(y)^\lambda} u(x) \geq \min_{x \in \operatorname{supp} P} u(x).$$

□

B.7 Proofs for Remarks on Continuity

This section considers the topology created by weighted weak convergence. First, we show Remark 2.1 that is under Common Consequence Independence, this is equivalent to our continuity in prizes conditional on a co-probability and mixture continuity.

Lemma B.84 Let $\succsim$ satisfy Axioms 2.1 and 2.4. Then $\succsim$ satisfies Axioms 2.2 and 2.3 iff $\succsim$ is continuous in weighted weak convergence as defined in Remark 2.1.

Proof. Weighted weak convergence implies Axioms 2.2 and 2.3.

We will show the other direction. By Theorem 2.1, $\succsim$ has either a power weighted expected utility representation or entropy modified expected utility representation.

First, if $\succsim$ has a power weighted expected utility representation, then it implies weighted weak convergence since power weighting is a continuous and increasing weighting function for $\lambda > 0$.

Second, assume that $\succsim$ has an entropy modified expected utility representation.

Let $(P_n)_{n=1}^\infty \subseteq \Delta(X)$ converge to $P \in \Delta(X)$ in weighted weak convergence. We show that $H(P_n) \to H(P)$ as $n \to \infty$, where H denotes the entropy. Take $f(x) = 1$ for all $x \in X$, $w^1(p) = p \log p + \sqrt{p}$, and $w^2(p) = \sqrt{p}$ for all $p \in [0,1]$. w^1 and w^2 are increasing and continuous functions with $w^1(0) = w^2(0) = 0$ and $w^1(1) = w^2(1) = 1$. By weighted weak convergence, we have

$$\sum_{x \in \operatorname{supp} P_n} P_n(x) \log(P_n(x)) + \sqrt{P_n(x)} \to \sum_{x \in \operatorname{supp} P} P(x) \log(P(x)) + \sqrt{P(x)} \text{ as } n \to \infty$$

and

$$\sum_{x \in \operatorname{supp} P_n} \sqrt{P_n(x)} \to \sum_{x \in \operatorname{supp} P} \sqrt{P(x)} \text{ as } n \to \infty$$

Especially we also have as a subtraction of two convergence sequences

$$\sum_{x \in \operatorname{supp} P_n} P_n(x) \log(P_n(x)) \to \sum_{x \in \operatorname{supp} P} P(x) \log(P(x)) \text{ as } n \to \infty.$$

We have

$$\sum_{x \in \operatorname{supp} P_n} u(x) P_n(x) P_n(x) \log(P_n(x)) \to \sum_{x \in \operatorname{supp} P} u(x) P_n(x) + P(x) \log(P(x)) \text{ as } n \to \infty$$

and this shows the continuity in weak convergence. □

Second, we show Remark 2.3 that is under FOSD preferences are continuous in weighted weak convergence if and only if they are continuous in weak convergence.

Lemma B.85 Let X be an interval. Let $\succsim$ satisfy Axiom 1 and Axiom 2.8. Then $\succsim$ satisfies Weak Convergence Continuity, Axiom 2.9, iff $\succsim$ is continuous in weighted weak convergence as defined in Remark 2.1.

Proof. "$\Rightarrow$" direction is clear. We will show "$\Leftarrow$". Denote $x^* = \sup X$.

Let $(P_n) \subseteq \Delta(X), P, Q \in \Delta(X)$ be such that $P_n \to P$ in weak convergence. Assume that for all $n \in \mathbb{N}$, $P_n \succsim Q$ since the other case follows symmetrically. Let $i^* \in \mathbb{N}$ be such that

$$\frac{1}{i^*} < \frac{1}{2}\min\{|x-y||x,y \in \operatorname{supp} P \cup \{x^*\}, x \neq y\}.$$

For each $i > i^*$, define the lottery $Q_i \in \Delta(X)$ as $\frac{1}{i}\varepsilon$ shift of P: by $Q_i = \sum_{x \in \operatorname{supp} P} P(x)\Big(x + \min\big\{\frac{1}{i}, x^* - x\big\}\Big)$ and $Q_i \in \Delta(X)$ by the definition of i^*.

Denote $x^\dagger = \max \operatorname{supp} P$.

For $i > i^*$, we will construct sequence R_i and a subsequence P_{n_i} inductively such that

$$R_i = \frac{1}{i}\Big(\max \operatorname{supp} P_{n_i} \cup \Big\{x^\dagger + \min\big\{\frac{1}{i}, x^* - x^\dagger\big\}\Big\}\Big) + (1 - \frac{1}{i})Q_i$$

and R_i first order stochastically dominates P_{n_i}. Assume that R_i has been defined for all $i^* < i < i^0$. Since $P_n \to P$ in weak convergence and $\operatorname{supp} P$ is finite, there exists $n^* \in \mathbb{N}$ such that for all $n > n^*$, $x^0 \in \operatorname{supp} P$

$$\Big|\sum_{x \in \operatorname{supp} P, |x - x^0| < \frac{1}{i^0}} P(x) - P(x^0)\Big| < \frac{1}{2i^0|\operatorname{supp} P|}.$$

Let n_{i^0} be such that $n_{i^0} > n_{i^0-1}, n^*$. Let

$$x^{i^0} = \max \operatorname{supp} P_{n_{i^0}} \cup \Big\{x^\dagger + \min\big\{\frac{1}{i^0}, x^* - x^\dagger\big\}\Big\} \in X.$$

Let $R_{i^0} = \frac{1}{i^0}x^{i^0} + (1 - \frac{1}{i^0})Q_{i^0}$. Now R_{i^0} first order stochastically dominates $P_{n_{i^0}}$ since

$$\sum_{x \in \operatorname{supp} P, x > x^\dagger + \frac{1}{i^0}} P_{n_{i^0}}(x) < \frac{1}{2i^0}.$$

Thus we have for all $i > i^*$, by Axiom 2.8,

$$R_i \succsim P_{n_i} \succsim Q.$$

Now R_i converges to P in weighted weak convergence: Let $w : [0,1] \to [0,1]$ be increasing function and continuous such that $w(0) = 0$ and $w(1) = 1$ and $f : X \to \mathbb{R}$ be bounded and continuous function bounded by $M \in \mathbb{R}_{++}$ from above and $-M$ from below. Let $\varepsilon > 0$. Now since $\operatorname{supp} P$ is finite, there exists $\delta > 0$ such that for all $x^0 \in \operatorname{supp} P$, if $x \in X$ and $p \in [0,1]$ are such that $|x^0 - x|, |p - P(x^0)| < \delta$,

then $|f(x)w(p) - f(x^0)w(P(x^0))| < \frac{\varepsilon}{|\operatorname{supp} P|+1}$ and if $|p-0| < \delta$, then $|w(p)| < \frac{\varepsilon}{M(|\operatorname{supp} P|+1)}$. Now there exists $i^0 > i^*$ such that $\frac{1}{i^0} < \delta$. So for all $i > i^0$, we have if $x^i = x^\dagger + \min\{\frac{1}{i}, x^* - x^\dagger\}$

$$\begin{aligned}
&\Big| \sum_{x\in\operatorname{supp} R_i} f(x)w(R_i(x)) - \sum_{x\in\operatorname{supp} P} f(x)w(P(x))\Big| \\
&\leq \sum_{x\in\operatorname{supp} P\setminus\{x^\dagger\}} \Big|f(x+\frac{1}{i})w((1-\frac{1}{i})P(x)) - f(x)w(P(x))\Big| \\
&\qquad\qquad + \Big|f(x+\min\{\frac{1}{i}, x^*-x\})w(P(x^\dagger)+\frac{1}{i}) - f(x^\dagger)w(P(x^\dagger))\Big| \\
&< \sum_{x\in\operatorname{supp} P\setminus\{x^\dagger\}} \frac{\varepsilon}{|\operatorname{supp} P|+1} + \frac{\varepsilon}{|\operatorname{supp} P|+1} \leq \varepsilon
\end{aligned}$$

and if $x^i \neq x^\dagger + \min\{\frac{1}{i}, x^* - x^\dagger\}$

$$\begin{aligned}
&\Big| \sum_{x\in\operatorname{supp} R_i} f(x)w(R_i(x)) - \sum_{x\in\operatorname{supp} P} f(x)w(P(x))\Big| \\
&\leq \sum_{x\in\operatorname{supp} P} \Big|f(x+\min\{\frac{1}{i}, x^*-x\})w((1-\frac{1}{i})P(x)) - f(x)w(P(x))\Big| + \Big|f(x^{i^0})w(\frac{1}{i})\Big| \\
&< \sum_{x\in\operatorname{supp} P} \frac{\varepsilon}{|\operatorname{supp} P|+1} + M\frac{\varepsilon}{M(|\operatorname{supp} P|+1)} = \varepsilon.
\end{aligned}$$

This shows that R_i converges to P in weighted weak convergence.

Finally, by continuity in weighted weak convergence, $P \succsim Q$ that shows the claim. □

Third, we show Remark 2.2 that is under FOSD preferences are continuous conditional on a co-probability if and only if they are continuous in shifts of prizes when prize space is an open interval.

Lemma B.86 Let X be an open interval. Let $\succsim$ satisfy Axiom 2.1 and Axiom 2.8. Then $\succsim$ satisfies Axiom 2.2 iff $\succsim$ satisfies Axiom 2.7.

Proof. $\Rightarrow$: Let $P, Q \in \Delta(X)$ be such that $P \succ Q$ and denote

$$\varepsilon = \min\{|\sup\operatorname{supp} Q - \sup X|, |\inf\operatorname{supp} P - \inf X|, 1\} > 0.$$

For all $n \in \mathbb{N}$, denote

$$P_n = P - \varepsilon\frac{1}{n} \text{ and } Q_n = Q + \varepsilon\frac{1}{n}.$$

Now for all $n \in \mathbb{N}$, $P_n, Q_n \in \Delta(X)$, P and P_n have the same co-probability, Q and Q_n have the same co-probability, P_n converges weakly to P and Q_n converges weakly to Q.

First, there exists $n^P \in \mathbb{N}$ such that $P_{n^P} \succ Q$: If for all $n \in \mathbb{N}$, $Q \succsim P_n$, then by Axiom 2.2, $Q \succsim P$ that is a contradiction.

Second, symmetrically, there exists $n^Q \in \mathbb{N}$ such that $P_{n^P} \succ Q_{n^Q}$. Thus

$$P - \varepsilon \frac{1}{n^P} \succ Q + \varepsilon \frac{1}{n^Q}$$

which shows Axiom 2.7.

$\Leftarrow$: Let $Q, P \in \Delta(X)$ and $(P_n)_{n=1}^{\infty} \subseteq \Delta(X)$ be such that for all $n \in \mathbb{N}$, P_n and P have the same co-probability and P_n converges weakly to P as $n \to \infty$. We only show the case for closed upper contour sets since lower contour sets follow symmetrically. Assume for all $n \in \mathbb{N}$,

$$P_n \succsim Q. \tag{B.134}$$

Assume, per contra, that $Q \succ P$. By Axiom 2.7 and Axiom 2.8, there exists $\varepsilon, \gamma > 0$ such that by

$$Q \succsim Q - \varepsilon \succ P + \gamma. \tag{B.135}$$

By the proof of Proposition B.1, there exist a co-probability $p \in \Delta(\mathbb{N})$ and for all $n \in \mathbb{N}$, $x^n \in \mathrm{NDiag}(X^p)$ such that $P_n = (p, x^n)$, $P = (p, x)$, and $x^n \to x$ as $n \to \infty$. Thus there exists $n \in \mathbb{N}$ such that $\|x^n - x\| < \gamma$. So especially for all $i \in \operatorname{supp} p$, $|x_i^n - x_i| < \gamma$ and so

$$x_i + \gamma > x_i^n. \tag{B.136}$$

Now $P + \gamma = (p, x + \gamma)$ and so by , $P + \gamma$ first order stochastically dominates P^n. By Axiom 2.8,

$$P + \gamma \succsim P^n.$$

This contradicts $(B.134, B.135)$ by Axiom 2.1. Thus $P \succsim Q$. This shows Axiom 2.2 since $P, Q \in \Delta(X)$ were arbitrary. □

Fourth, we show a general version of Remark 2.2 that is under FOSD preferences are continuous conditional on a co-probability if and only if they are continuous in shifts of prizes for a general interval and Borel lotteries.

Lemma B.87 Let X be an interval. Let $\succsim$ satisfy Axiom $1'^{\mathrm{EU}}$ and Axiom $4'^{\mathrm{EU}}$. Then $\succsim$ satisfies Axiom $3'^{\mathrm{EU}}$ iff for all lotteries $Q \in \Delta^{\mathrm{Bor}}(X)$ and simple lotteries $(P_n)_{n=1}^{\infty} \subseteq \Delta(X)$ and $P \in \Delta(X)$ such that for all $n \in \mathbb{N}$, P_n and P have the same co-probability and P_n converges weakly to P as $n \to \infty$,

$$\text{if } P_n \succsim Q \text{ for all } n \in \mathbb{N}, \text{ then } P \succsim Q$$

and

$$\text{if } Q \succsim P_n \text{ for all } n \in \mathbb{N}, \text{ then } Q \succsim P.$$

Proof. $\Leftarrow$: Let $P \in \Delta(X)$ and $Q \in \Delta^{\mathrm{Bor}}(X)$. Assume that $P \succ Q$ since the other case follows symmetrically. If $\inf X \in \operatorname{supp} P$, then X is closed from below and denote

$$\varepsilon = \min\{|\inf(\operatorname{supp} P \setminus \{\inf X\}) - \inf X|, 1\} > 0.$$

If $\inf X \notin \operatorname{supp} P$, then denote

$$\varepsilon = \min\{|\inf \operatorname{supp} P - \inf X|, 1\} > 0.$$

For all $n \in \mathbb{N}$, denote

$$P_n = P -_{|X} \varepsilon \frac{1}{n}.$$

Now for all $n \in \mathbb{N}$ in both cases, $P_n \in \Delta(X)$, P and P_n have the same co-probability by the choice of ε, and P_n converges weakly to P.

Now there exists $n \in \mathbb{N}$ such that $P -_{|X} \varepsilon\frac{1}{n} = P_n \succ Q$: If for all $n \in \mathbb{N}$, $Q \succsim P_n$, then by the assumption, $Q \succsim P$ that is a contradiction. This shows Axiom $3'^{\mathrm{EU}}$ since the case $Q \succ P$ follows symmetrically.

$\Rightarrow$: Let $Q \in \Delta^{\mathrm{Bor}}(X)$, $P \in \Delta(X)$ and $(P_n)_{n=1}^{\infty} \subseteq \Delta(X)$ be such that for all $n \in \mathbb{N}$, P_n and P have the same co-probability and P_n converges weakly to P as $n \to \infty$. We only show the case for closed upper contour sets since lower contour sets follow symmetrically. Assume for all $n \in \mathbb{N}$,

$$P_n \succsim Q. \tag{B.137}$$

Assume, per contra, that $Q \succ P$. By Axiom $3'^{\mathrm{EU}}$, there exists $\tilde{\varepsilon} > 0$ such that by

$$Q \succ P +_{|X} \tilde{\varepsilon}.$$

If $\sup X \in \operatorname{supp} P$, then denote

$$\varepsilon = \frac{1}{2}\min\{|\sup(\operatorname{supp} P \setminus \{\sup X\}) - \sup X|, \tilde{\varepsilon}\} > 0.$$

If $\sup X \notin \operatorname{supp} P$, then denote

$$\varepsilon = \frac{1}{2}\min\{|\sup \operatorname{supp} P - \sup X|, \tilde{\varepsilon}\} > 0.$$

Then $P +_{|X} \varepsilon$ and P have the same co-probability and by Axiom $4'^{\mathrm{EU}}$,

$$Q \succ P +_{|X} \tilde{\varepsilon} \succsim P +_{|X} \varepsilon. \tag{B.138}$$

By the proof of Proposition B.1, there exist a co-probability $p \in \Delta(\mathbb{N})$ and for all $n \in \mathbb{N}$, $x^n \in \mathrm{NDiag}(X^p)$ such that $P_n = (p, x^n)$, $P = (p, x)$, and $x^n \to x$ as $n \to \infty$. Thus there exists $n \in \mathbb{N}$ such that $\|x^n - x\| < \varepsilon$. So especially for all $i \in \operatorname{supp} p$, $|x_i^n - x_i| < \varepsilon$ and so

$$x_i + \varepsilon > x_i^n \text{ and } \sup X \leq x_i^n. \tag{B.139}$$

Now

$$P +_{|X} \varepsilon = (p, (\min\{x_i + \varepsilon, \sup X\})_{i \in \operatorname{supp} p}).$$

By , $P +_{|X} \varepsilon$ first order stochastically dominates P^n. By Axiom $4'^{\mathrm{EU}}$,

$$P +_{|X} \varepsilon \succsim P^n.$$

This contradicts $(B.137, B.138)$ by Axiom $1'^{\mathrm{EU}}$. Thus $P \succsim Q$. This shows the claim since $P \in \Delta(X)$ and $Q \in \Delta^{\mathrm{Bor}}(X)$ were arbitrary. □

Appendix C

Appendix for Chapter 3

C.1 Additional Graphs

C.1.1 Absolute Hit Rates

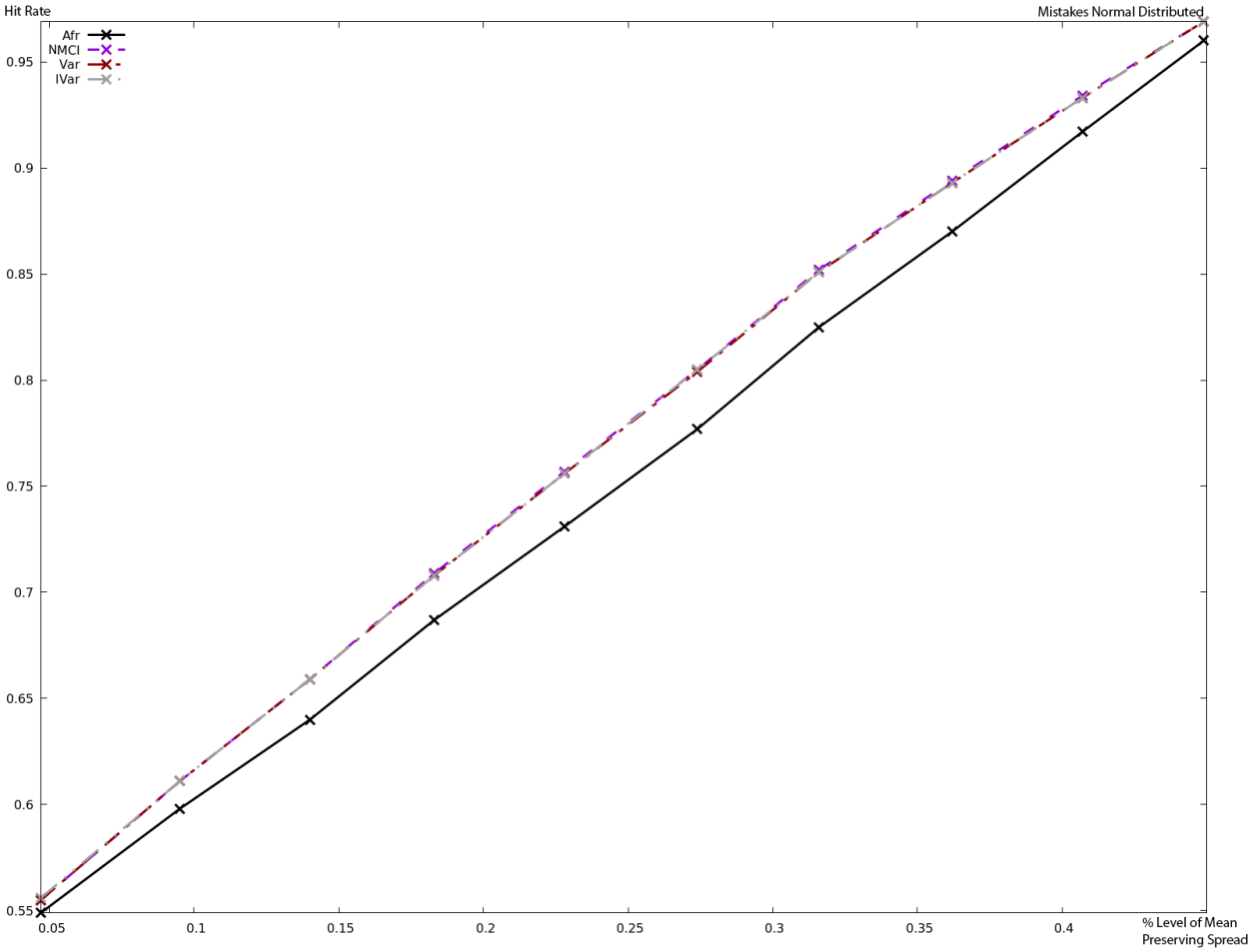

Figure C.1. Measures of rationality's hit rates when mistakes are drawn from Normal distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread. For example at 5% the low mistake population's share of rational people is 55% and the high mistake population's share of rational people is 45%.

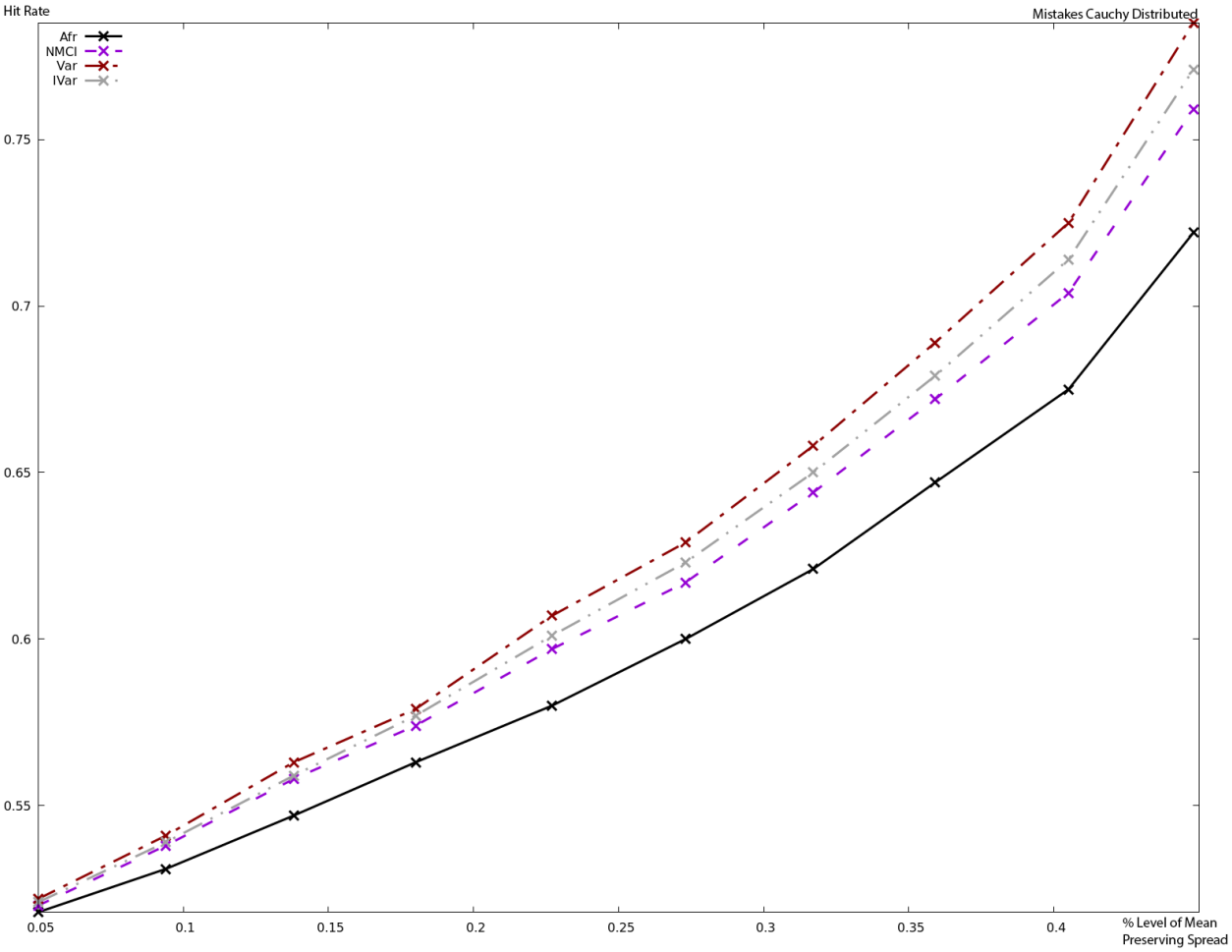

Figure C.2. Measures of rationality's hit rates when mistakes are drawn from Cauchy distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread.

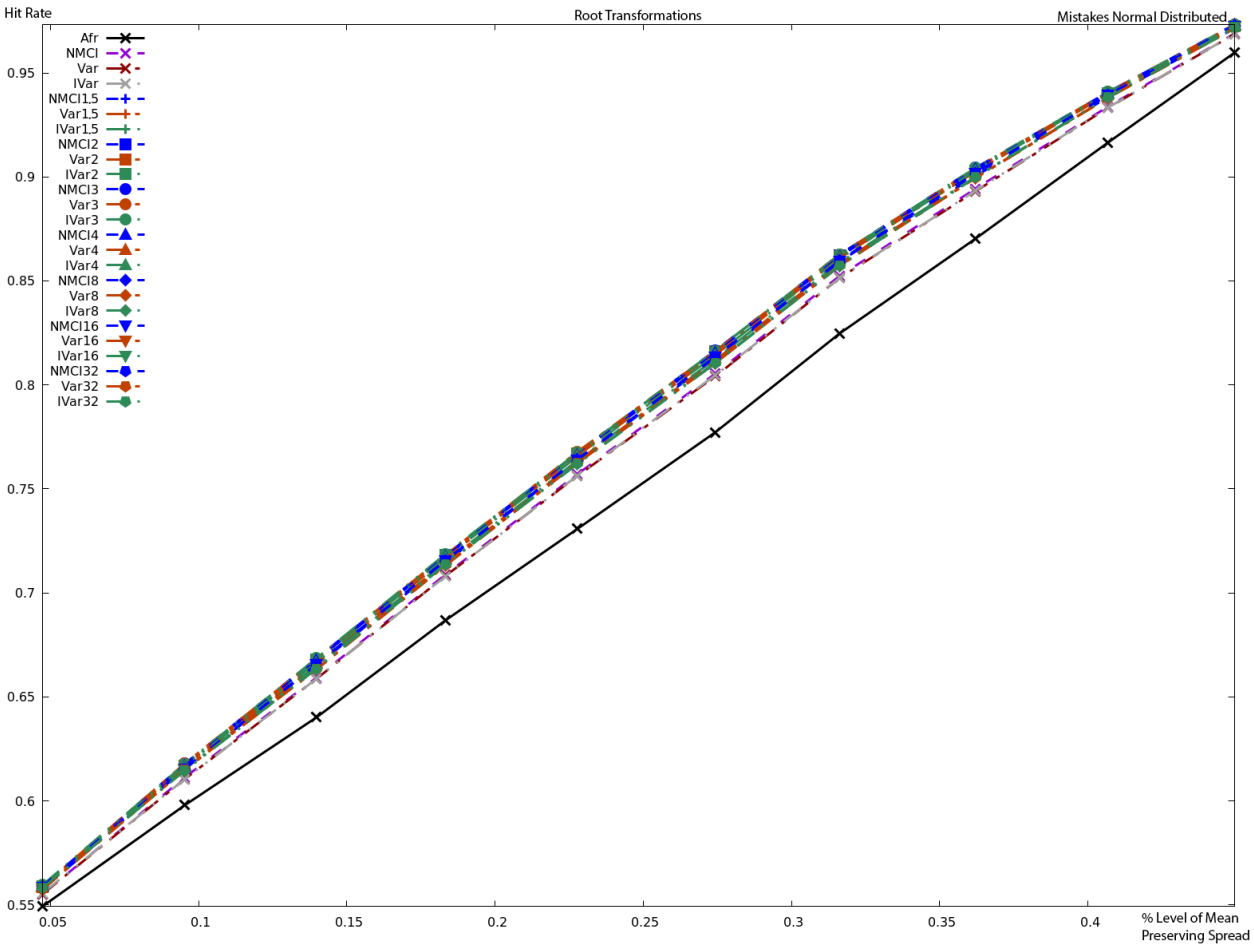

Figure C.3. Concavifications of measures of rationality using root transformations for mistakes. Root transformations are varied between 1 and 32 and denoted in the labels. Measures' hit rates when mistakes are drawn from Normal distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread.

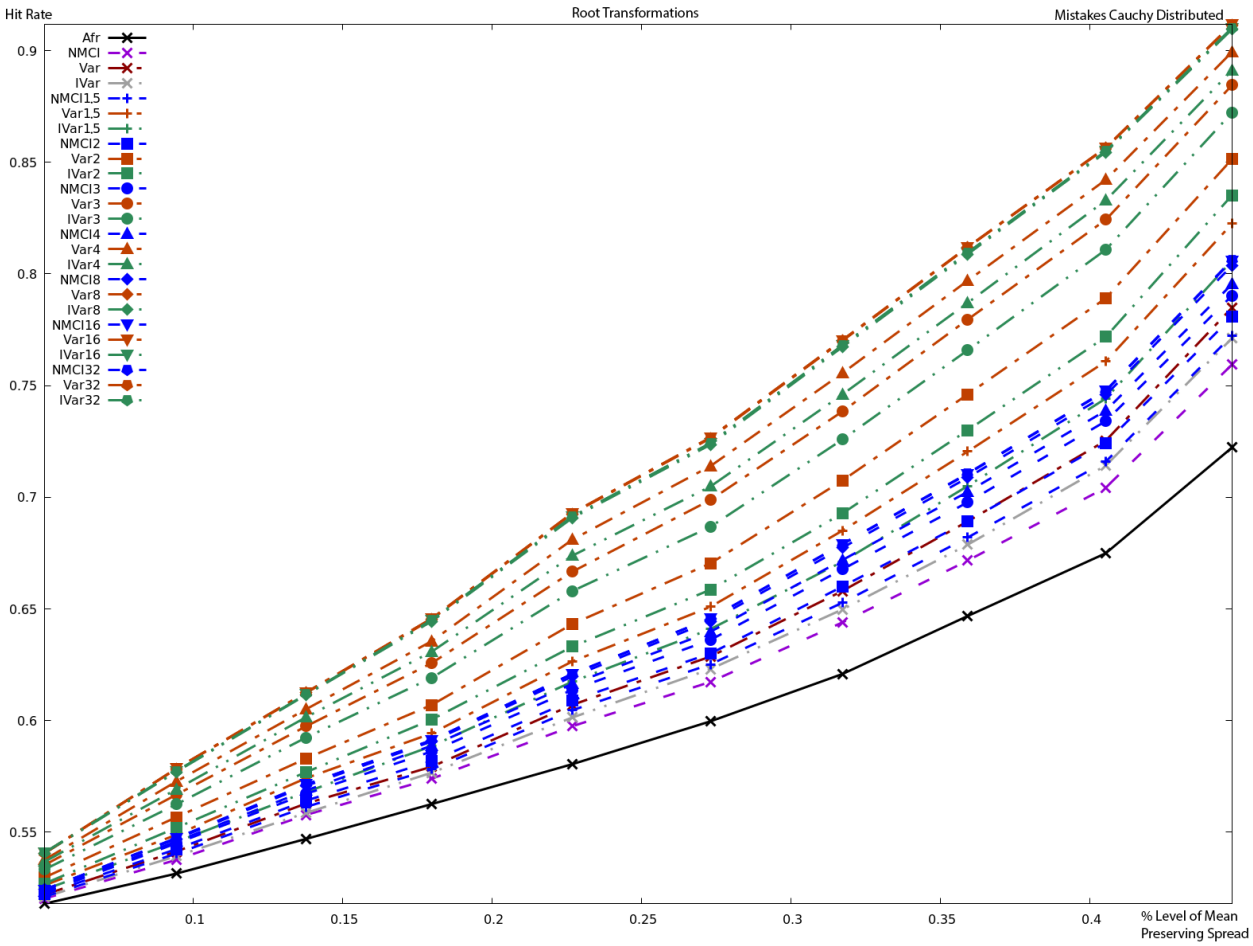

Figure C.4. Concavifications of measures of rationality using root transformations for mistakes. Root transformations are varied between 1 and 32 and denoted in the labels. Measures' hit rates when mistakes are drawn from Cauchy distribution. The two populations share of rational people are calibrated to be a mean preserving spread from 50%. The x-axis changes the percentage level of mean preserving spread.

C.1.2 Alternative Calibration Method

We consider an alternative way to calibrate the mistake level for the populations. In this method, we keep the difficulty of the comparison constant and change the average level of rationality. We measure the difficulty of comparison using Afriat's Efficiency Index as a baseline. We calibrate the mistake level for the population with the higher mistake level to keep the Afriat's Efficiency Index at a fixed value. Then, we vary the mistake level for the population with the lower mistake level from a level such that 95% of the population is rational to a level such that 5% of the population is rational. This gives as comparisons of measures as we keep the difficulty of comparison constant but change the level of rationality.

We calibrate the Afriat's Index to 0.65. As in the text, we have 6 goods and 45 time periods. Each point consists of 100,000 pairwise comparisons when both of the data sets compared have rationality violations.

The results are similar to the main text.

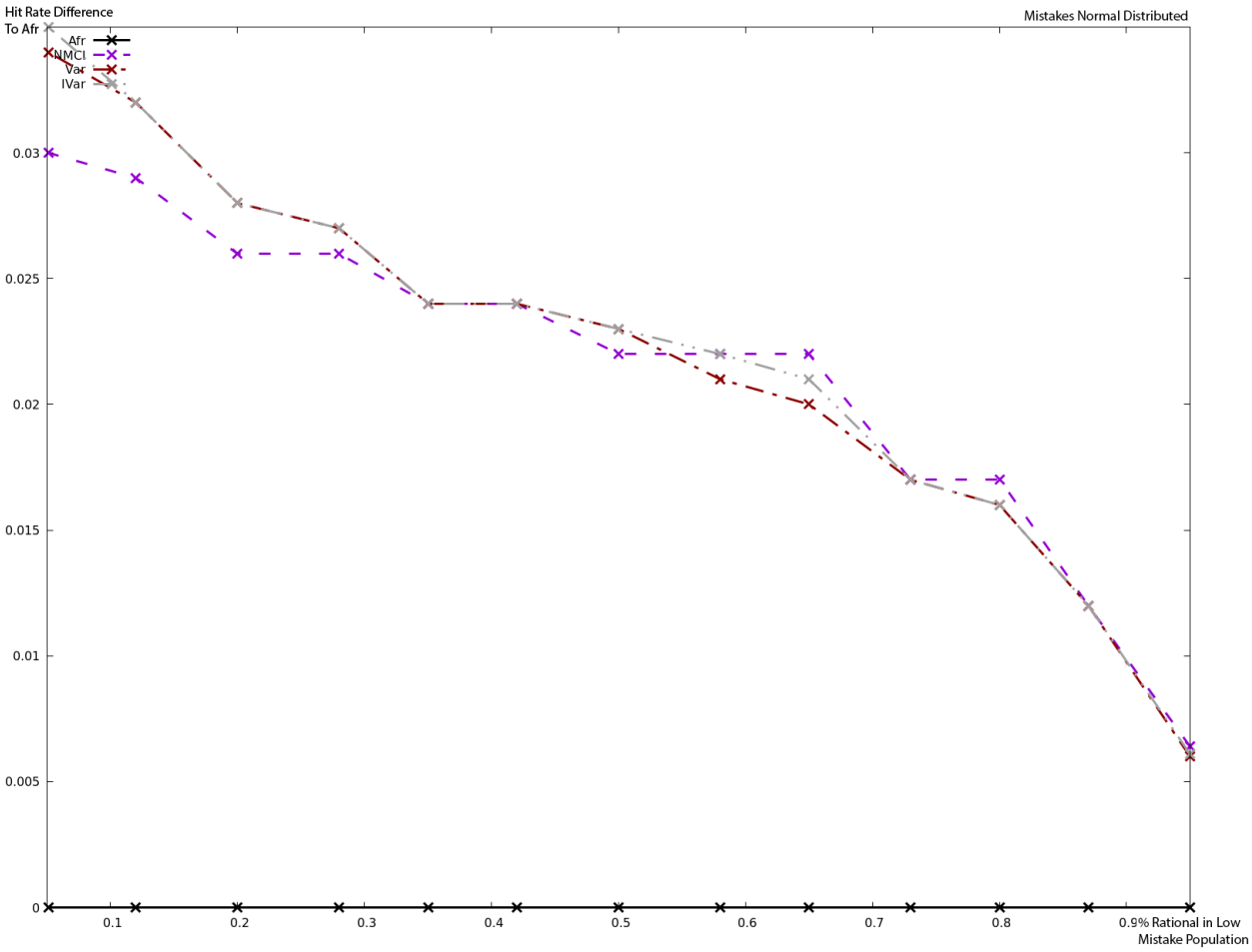

Figure C.5. Measures of rationality's hit rate differences to Afriat's Index's hit rate at predicting which of two populations has a higher level of mistakes when mistakes are drawn from Normal distribution. Afriat's Index's hit rate is calibrated to be at 0.65 using the mistake level for the higher mistake rate population. The x-axis changes the percentage of rational people in the population with lower mistakes.

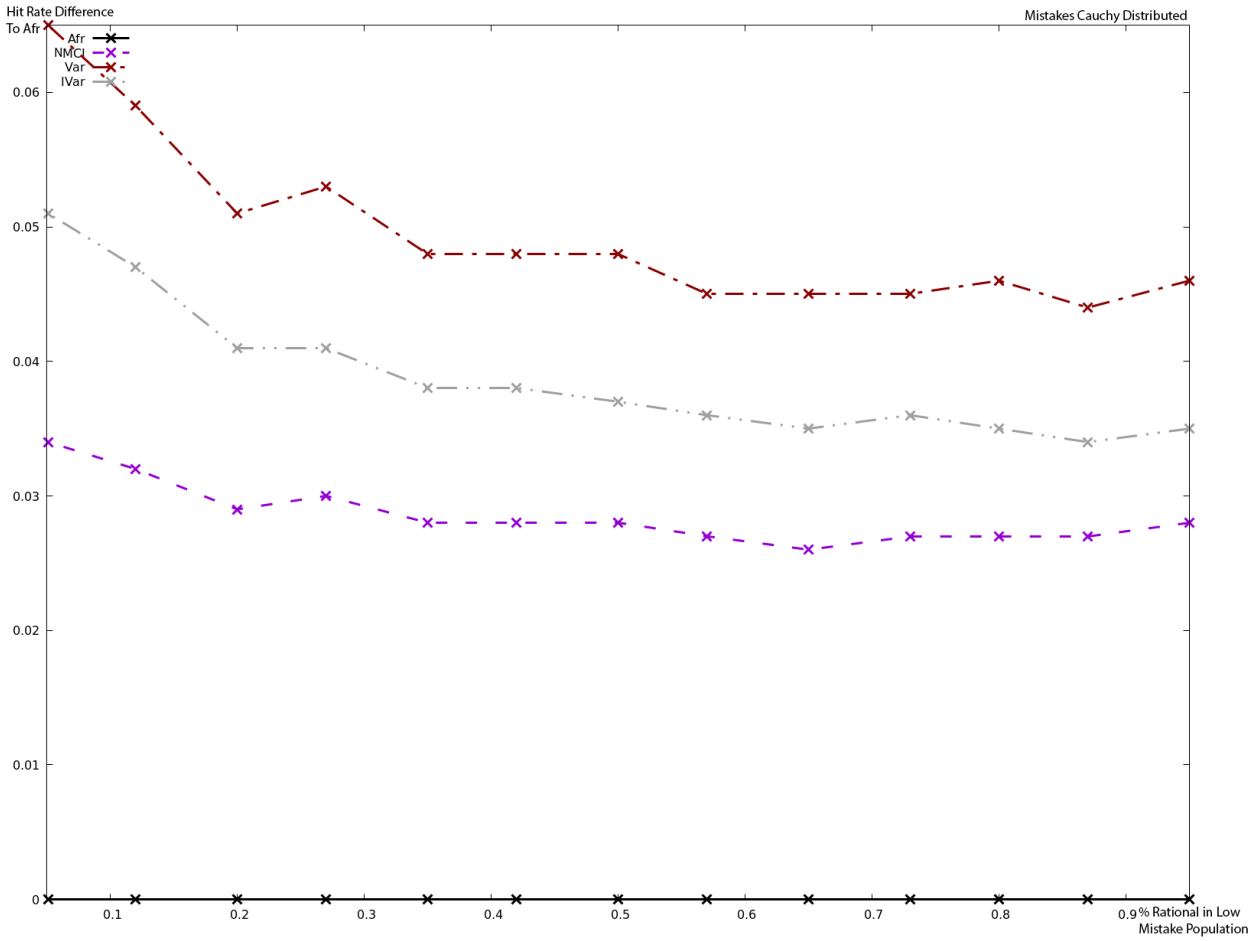

Figure C.6. Measures of rationality's hit rate differences to Afriat's Index's hit rate at predicting which of two populations has a higher level of mistakes when mistakes are drawn from Cauchy distribution. Afriat's Index's hit rate is calibrated to be at 0.65 using the mistake level for the higher mistake rate population. The x-axis changes the percentage of rational people in the population with lower mistakes.

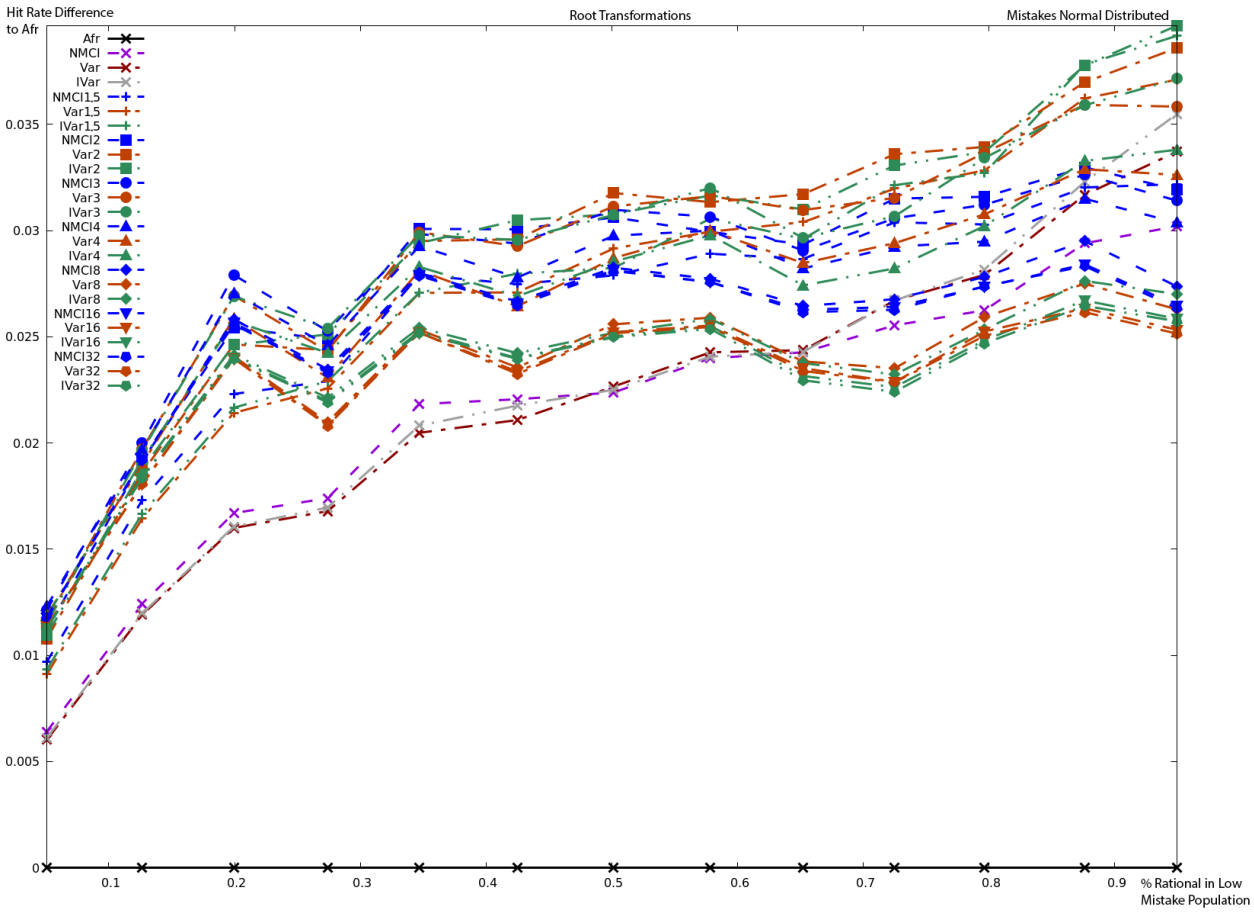

Figure C.7. Concavifications of measures of rationality using root transformations for mistakes. Root transformations are varied between 1 and 32 and denoted in the labels. Measures' hit rate differences to Afriat's Index's hit rate at predicting which of two populations has a higher level of mistakes when mistakes are drawn from Normal distribution. Afriat's Index's hit rate is calibrated to be at 0.65 using the mistake level for the higher mistake rate population. The x-axis changes the percentage of rational people in the population with lower mistakes.

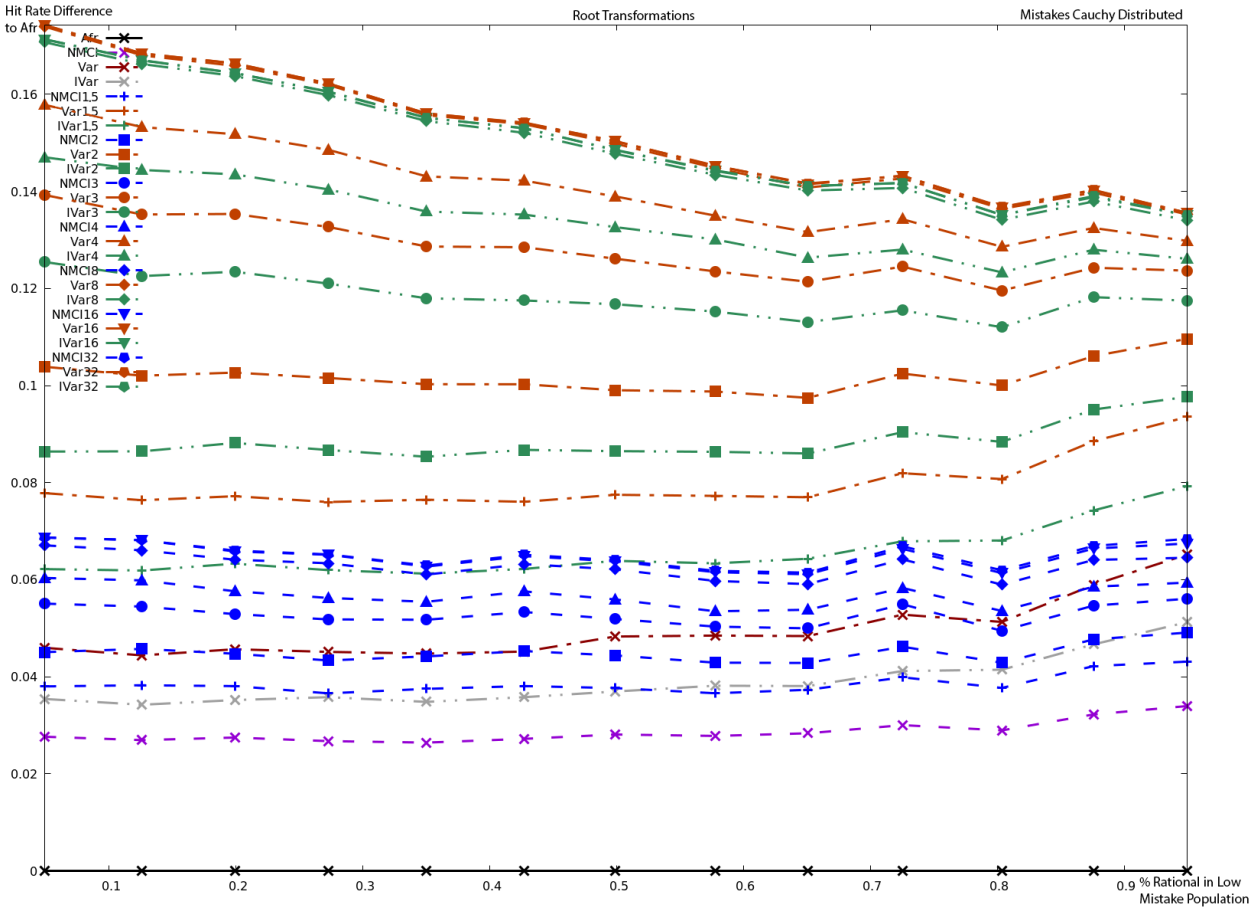

Figure C.8. Concavifications of measures of rationality using root transformations for mistakes. Root transformations are varied between 1 and 32 and denoted in the labels. Measures' hit rate differences to Afriat's Index's hit rate at predicting which of two populations has a higher level of mistakes when mistakes are drawn from Cauchy distribution. Afriat's Index's hit rate is calibrated to be at 0.65 using the mistake level for the higher mistake rate population. The x-axis changes the percentage of rational people in the population with lower mistakes.

C.2 Afriat's Efficiency Index: Algorithm

This section provides a novel algorithm to calculate Afriat's Efficiency Index based on Depth First Search. The algorithm is presented in Algorithm C.1. Here, q and p are the purchased bundles and vectors of prices respectively.

The idea of the algorithm is to update the estimate for Afriat's index upwards. This is done by iteratively in the revealed preference graph removing all relations that are cheaper than the current estimate for Afriat's Index. Then by finding cycles in the leftover graph and updating the estimate to be the cheapest trade in the cycle. The search of cycles is done in single depth-first search. So

always when a cycle is found, we roll back the search to the beginning of the cycle and continue the search from there with the updated leftover graph.

C.3 Proofs

As in the text, we will take all sums of indices modulo the length of the sequence.

C.3.1 Preliminary Results

Lemma C.1 Let $a > b > 0$ and $g = T$. Then there exists $P \in \mathbb{R}_{++}^{g \times T}$ such that for each $A \in (b, a)^{T \times T}$, there exists $X(A) \in \mathbb{R}_{++}^{g \times T}$ such that

$$\text{trans}(P) * X(A) = A.$$

Additionally, $X(A)$ is continuous function of A in the Euclidean topology.

Proof. Let $\varepsilon \in \mathbb{R}_{++}$ be such that

$$b\Big(1 - \frac{\varepsilon}{1 + \varepsilon T}\Big) - a(T-1)\frac{\varepsilon}{1 + \varepsilon T} > 0.$$

We define $P \in \mathbb{R}_{++}^{g \times T}$ by for all $1 \le i, j, k \le T$,

$$P_{i,j} = \begin{cases} 1 + \varepsilon, & \text{if } i = j \\ \varepsilon, & \text{if } i \neq j. \end{cases}$$

We define $X(A) \in \mathbb{R}_{+}^{g \times T}$ by for all $1 \le i, j \le T$,

$$X(A)_{i,j} = A_{i,j}\Big(1 - \frac{\varepsilon}{1 + \varepsilon T}\Big) - \sum_{\substack{k=1 \\ k \neq i}}^{T} A_{k,j}\frac{\varepsilon}{1 + \varepsilon T}.$$

By the choice of ε, $X(A)$ is positive. Additionally, $X(A)$ is a continuous function of A in the Euclidean product topology. Denote by $\bar{1}$ a T-dimensional vector of ones and by I the T-dimensional identity matrix. Now $P = I + \varepsilon\bar{1} * \text{trans}(\bar{1})$. By Sherman–Morrison formula, P is invertible with inverse $(I - \frac{\varepsilon}{1+\varepsilon T}\bar{1} * \text{trans}(\bar{1}))$. Now $X(A) = P^{-1} * A$.

Finally, we have since P is symmetric

$$\text{trans}(P) * X(A) = P * P^{-1} * A = A.$$

□

Lemma C.2 Let I satisfy Axiom 3.1. Let $D = \big((p_1, x_1), \ldots, (p_T, x_T)\big) \in \mathcal{D}$ be such that for all $1 \le i \le T$, $p_i \cdot x_i > 0$. Then

$$\tilde{D} = \big((\frac{p_1}{p_1 \cdot x_1}, x_1), \ldots, (\frac{p_T}{p_T \cdot x_T}, x_T)\big)$$

Algorithm C.1: Afriat's Depth First Search

Input: T, $(\mathsf{p}(t))_{t=1}^{\mathsf{T}}$, $(\mathsf{q}(t))_{t=1}^{\mathsf{T}}$,
Output: e
Global variables: T,e, $(\mathsf{p}(t))_{t=1}^{\mathsf{T}}$, $(\mathsf{q}(t))_{t=1}^{\mathsf{T}}$, $(\mathsf{visited}(t))_{t=1}^{\mathsf{T}}$, $(\mathsf{stacked}(t))_{t=1}^{\mathsf{T}}$
begin
    $\mathsf{e} := 0$ /* Set starting point at 0 */
    for $v = 1$ **to** T **do**
        if $\mathsf{visited}(v) = \text{false}$ **then** /* Search unvisited elements */
            $\mathsf{visited}(v) := \mathsf{stacked}(v) := \text{true}$
            `Afriat(`v`)`
            $\mathsf{stacked}(v) := \text{false}$

Function `Afriat(`v`)`**:**
    for $w = 1$ **to** T **do**
        if $\frac{\mathsf{p}(v)\cdot(\mathsf{q}(v)-\mathsf{q}(w))}{\mathsf{p}(v)\cdot\mathsf{q}(v)} > \mathsf{e}$ **then** /* Non-removed relation */
            if $\mathsf{visited}(w) = \text{false}$ **then** /* Visit */
                $\mathsf{visited}(w) := \mathsf{stacked}(w) := \text{true}$
                while $\mathsf{stacked}(w) = \text{true}$ **do**
                    $\mathsf{cost} :=$ `Afriat(`w`)`
                    if $\mathsf{cost} > 0$ **then** /* $\mathsf{cost} > 0$ iff found cycle */
                        $\mathsf{cost} := \min\left(\mathsf{cost}, \frac{\mathsf{p}(v)\cdot(\mathsf{q}(v)-\mathsf{q}(w))}{\mathsf{p}(v)\cdot\mathsf{q}(v)}\right)$ /* Update min cost of cycle */
                        if $\mathsf{visited}(v) = \text{true}$ **then** /* Not at the beginning, rollback */
                            $\mathsf{visited}(w) := \text{false}$
                            return cost
                        else /* At the beginning of cycle */
                            $\mathsf{visited}(v) := \text{true}$
                            $\mathsf{e} := \max(\mathsf{e}, \mathsf{cost})$ /* Update estimate & search again */
                    else
                        $\mathsf{stacked}(w) := \text{false}$ /* No cycles found, exit */
            else if $\mathsf{stacked}(w) = \text{true}$ **then**/* visited & stacked, found cycle */
                $\mathsf{visited}(w) := \text{false}$ /* Mark the first element of cycle visited */
                $\mathsf{cost} := \frac{\mathsf{p}(v)\cdot(\mathsf{q}(v)-\mathsf{q}(w))}{\mathsf{p}(v)\cdot\mathsf{q}(v)}$ /* Collect min cost of cycle */
                return cost /* Return current min cost of cycle */
    return 0 /* If no cycles, return 0 */

is unit income normalized data such that $I(D) = I(\tilde{D})$.

Proof. Follows from applying Axiom 3.1 T times since for all $1 \leq i \leq T$, $p_i \cdot x_i > 0$. □

Lemma C.3 If $D = \big((p_1, x_1), \ldots, (p_T, x_T)\big) \in \mathcal{D}$ is such that for all $1 \leq i \leq T$, $p_i \cdot x_i > 0$, then for all $1 \leq i, j \leq T$, $x_j \neq 0$ and $p_i \cdot x_j > 0$.

Proof. Let $1 \leq i, j \leq T$. Since $p_i \cdot x_i > 0$ and p_i, x_i are non-negative, $x_i \neq 0$. Thus since p_j is strictly positive, $p_j \cdot x_i > 0$. □

Lemma C.4 Let I satisfy Axioms 3.1 to 3.3. If $D = \big((p_1, x_1), \ldots, (p_T, x_T)\big) \in \mathcal{D}$ is such that for all $1 \leq i \leq T$, $p_i \cdot x_i > 0$ and for all $(x_{i_1}, \ldots, x_{i_n}) \in \mathcal{C}^D$, there exists $1 \leq j \leq n$ such that

$$p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}}) = 0, \tag{C.1}$$

then

$$I(D) = 0.$$

Proof. Define

$$\tilde{D} = \big((\frac{p_1}{p_1 \cdot x_1}, x_1), \ldots, (\frac{p_T}{p_T \cdot x_T}, x_T)\big) = \big((\tilde{p}_1, \tilde{x}_1), \ldots, (\tilde{p}_T, \tilde{x}_T)\big).$$

By Lemma C.2, $I(\tilde{D}) = I(D)$. Additionally $\tilde{D}$ inherits (C.1) and for all $(\tilde{x}_{i_1}, \ldots, \tilde{x}_{i_n}) \in \mathcal{C}^{\tilde{D}}$, there exists $1 \leq j \leq n$ such that

$$\tilde{p}_{i_j} \cdot (\tilde{x}_{i_j} - \tilde{x}_{i_{j+1}}) = 0. \tag{C.2}$$

Denote

$$B = \big\{(i, j) | \tilde{p}_i \cdot (\tilde{x}_i - \tilde{x}_j) = 0\big\}.$$

Let $\alpha \in [0, 1)$. Define $A^\alpha \in \mathbb{R}^{T \times T}$ as

$$A_{i,j} = \begin{cases} 1 + \alpha, & \text{if } (i, j) \in B \\ \tilde{p}_i \cdot \tilde{x}_j, & \text{if } (i, j) \notin B. \end{cases}$$

By Lemmas C.1 and C.3, there exist P, X^α such that $\text{trans}(P) * X^\alpha = A^\alpha$. Additionally, P and X^α are continuous in α in the Euclidean topology. Denote

$$D^\alpha = \big((\text{pr}_{\bar{T} \times \{1\}} P, \text{pr}_{\bar{T} \times \{1\}} X^\alpha), \ldots, \text{pr}_{\bar{T} \times \{T\}} P, \text{pr}_{\bar{T} \times \{T\}} X^\alpha\big) = \big((P_1, x_1^\alpha), \ldots, (P_T, x_T^\alpha)\big).$$

Now, $D^\alpha \in \mathcal{D}$ and is unit income normalized data.

By Axiom 3.3, $I(D^0) = I(\tilde{D}) = I(D)$. By (C.2), for all $\alpha \in (0, 1)$, $\mathcal{C}^{D^\alpha} = \varnothing$ and so $I(D^\alpha)$. Thus by Axiom 3.2,

$$I(D^0) = \lim_{\alpha \to 0} I(D^\alpha) = 0.$$

□

C.3.2 Afriat's Efficiency Index

Lemma C.5 Let I satisfy Axioms 3.2 to 3.4. Let $D = \big((p_1, x_1), \ldots, (p_T, x_T)\big) \in \mathcal{D}$ be unit income normalized data. Then

$$I(D) \geq I_{\text{Afr}}(D).$$

Proof. Denote $\bar{T} = \{1, \ldots, T\}$. Let $c = (x_{i_1}, \ldots, x_{i_n}) \in \mathcal{C}^D$. Denote

$$B = \{(i_j, i_{j+1}) | j \in \{1, \ldots, n\}\}$$

and

$$d = \min_{1 \leq j \leq n} p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}}).$$

We will show that $I(D) \geq d$. If $d = 0$, then it is clear. So assume that $d > 0$. First, assume that $d < 1$.

First, let $\alpha \in (-1, d)$. Define $A^\alpha \in \mathbb{R}^{T \times T}_{++}$ as such that

$$A_{i,j} = \begin{cases} 1, & \text{if } i = j \\ 1 - d, & \text{if } (i,j) \in B \\ \max\{p_i \cdot x_j, 1 - \alpha\}, & \text{if } p_i \cdot x_j \leq 1, (i,j) \notin B \\ 2, & \text{if } p_i \cdot x_j > 1, (i,j) \notin B. \end{cases}$$

By Lemma C.1, there exist P, X^α such that $\text{trans}(P) * X^\alpha = A^\alpha$. Additionally, P and X^α are continuous in α in the Euclidean topology. Denote

$$D^\alpha = \big((\text{pr}_{\bar{T} \times \{1\}} P, \text{pr}_{\bar{T} \times \{1\}} X^\alpha), \ldots, \text{pr}_{\bar{T} \times \{T\}} P, \text{pr}_{\bar{T} \times \{T\}} X^\alpha) = \big((p_1^0, x_1^\alpha), \ldots, (p_T^0, x_T^\alpha)\big).$$

Now, $D^\alpha \in \mathcal{D}$ and is unit income normalized data. Next, let $\alpha \in (0, d)$

Now we have for all $i, j \in \bar{T}$,

$$p_i \cdot (x_i - x_j) \geq 0 \iff p_i^0 \cdot (x_i^\alpha - x_j^\alpha) \geq 0,$$

$$p_i \cdot (x_i - x_j) > 0 \iff p_i^0 \cdot (x_i^\alpha - x_j^\alpha) > 0,$$

and

$$p_i \cdot (x_i - x_j) \geq 0 \Rightarrow p_i \cdot (x_i - x_j) \leq p_i^0 \cdot (x_i^\alpha - x_j^\alpha).$$

Additionally $D^\alpha \in \mathcal{D}$ and is unit income normalized data. Thus by Axiom 3.3, we have

$$I(D) \geq I(D^\alpha).$$

Thus we have for all $\alpha \in (0, d)$,

$$I(D) \geq I(D^\alpha).$$

Since D^α is continuous in α in the Euclidean topology, we can take $\alpha \to 0$ for

$$I(D) \geq I(D^0).$$

On the other hand, let $\alpha \in (-1, 0)$. Now D^α has a single constant cycle of size d. Thus by Axiom 3.5, $I(D^\alpha) = d$. Since D^α is continuous in α, we can take $\alpha \to 0$ to have

$$I(D^0) = d.$$

Thus we have $I(D) \geq d$.

Next assume that $d = 1$. Following above, we have for all $\tilde{d} \in (0, 1)$, $I(D) \geq \tilde{d}$. Thus especially, $I(D) \geq \sup_{\tilde{d} \in (0,1)} \tilde{d} = 1$.

This holds for all $c \in \mathcal{C}^D$. Thus we have

$$I(D) \geq \max_{(x_{i_1}, \ldots, x_{i_n}) \in \mathcal{C}^D} \min_{1 \leq j \leq n} p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}}),$$

which shows the claim since D is unit income normalized. □

Lemma C.6 Let I satisfy Axioms 3.2 to 3.5. Let $D = \big((p_1, x_1), \ldots, (p_T, x_T)\big) \in \mathcal{D}$ be unit income normalized data. Then

$$I(D) \leq I_{\mathrm{Afr}}(D).$$

Proof. Denote $\beta = I_{\mathrm{Afr}}(D)$. Define $\tilde{A} \in \mathbb{R}_{++}^{T \times T}$ as such that

$$\tilde{A}_{i,j} = \begin{cases} 1, & \text{if } i = j \\ p_i \cdot x_j + \beta, & \text{if } i \neq j. \end{cases}$$

By Lemma C.1, there exist $\tilde{P}, \tilde{X}$ such that $\mathrm{trans}(\tilde{P}) * \tilde{X} = \tilde{A}$. Denote

$$\tilde{D} = \big((\mathrm{pr}_{\bar{T} \times \{1\}} \tilde{P}, \mathrm{pr}_{\bar{T} \times \{1\}} \tilde{X}), \ldots, \mathrm{pr}_{\bar{T} \times \{T\}} \tilde{P}, \mathrm{pr}_{\bar{T} \times \{T\}} \tilde{X}\big) = \big((\tilde{p}_1, \tilde{x}_1), \ldots, (\tilde{p}_T, \tilde{x}_T)\big).$$

Now, $\tilde{D} \in \mathcal{D}$ and is unit income normalized data. By the definition of β, for all $(x_{i_1}, \ldots, x_{i_n}) \in \mathcal{C}^D$, there exists $1 \leq j \leq n$ such that $p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}}) = 0$. By Lemma C.4, $I(\tilde{D}) = 0$. By Axiom 3.5,

$$I_{\mathrm{Afr}}(D) = \beta = I(\tilde{D}) + \beta \geq I(D).$$

□

Theorem C.7 Let I be a measure of rationality on $\mathcal{D}$. I satisfies Axioms 3.1-3.5 iff

$$I = I_{\mathrm{Afr}}.$$

Proof. We first show that Axioms 3.1-3.5 are necessary. Let $D = ((p_1, x_1), \ldots, (p_T, x_T)) \in \mathcal{D}$.

Axiom 3.1: Let $i \in \{1, ..., T\}, \alpha \in \mathbb{R}_{++}$. Denote $\tilde{D} = ((p_1, x_1), ..., (\alpha p_i, x_i), ...(p_T, x_T))$. First, we have $\mathcal{C}^D = \mathcal{C}^{\tilde{D}}$. If $\mathcal{C}^D = \varnothing$, then the claim is clear. So assume that $\mathcal{C}^D \neq \varnothing$. Denote for all $j \in \{1, \dots, T\} \setminus \{i\}$, $\alpha_j = 1$ and $\alpha_i = 1$. Now we have

$$I_{\text{Afr}}(D) = \max_{(x_{i_1}, \dots, x_{i_n}) \in \mathcal{C}^D} \min_{1 \leq j \leq n} \frac{p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}})}{p_{i_j} \cdot x_{i_j}} = \max_{(x_{i_1}, \dots, x_{i_n}) \in \mathcal{C}^{\tilde{D}}} \min_{1 \leq j \leq n} \frac{\alpha_{i_j} p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}})}{\alpha_{i_j} p_{i_j} \cdot x_{i_j}} = I_{\text{Afr}}(\tilde{D}).$$

Axiom 3.2: Let $(p_1, \dots, p_T)$ be prices, $g \in \mathbb{N}$ be the number of goods, and for each $n \in \mathbb{N}$, let $\big((p_1, x_1^n), \dots, (p_T, x_T^n)\big) \in \mathcal{D}$ be such that for each $t \in \{1, \dots, T\}$, there exists x_t^0 such that $\lim_{n \to \infty} x_t^n = x_t^0$ and $\big((p_1, x_1^0), \dots, (p_T, x_T^0)\big) \in \mathcal{D}$. We consider two cases.

First, assume that for all $t \in \{1, \dots, T\}$, $p_t \cdot x_t^0 > 0$. Denote the collection of all the possible indices for cycles for data of length T by

$$\mathcal{E} = \{(i_1, \dots, i_n) \subseteq \{1, \dots, T\} | n \in \mathbb{N}, \forall\, j, k \{1, \dots, n\}, j \neq k, i_j \neq i_k\}.$$

Now we have by identity, for all $n \in \mathbb{N} \cup \{0\}$,

$$I_{\text{Afr}}\big((p_1, x_1^n), \dots, (p_T, x_T^n)\big) = \max_{(i_1, \dots, i_n) \in \mathcal{E}} \min_{1 \leq j \leq n} \max \Big\{ \frac{p_{i_j} \cdot (x_{i_j}^n - x_{i_{j+1}}^n)}{p_{i_j} \cdot x_{i_j}^n}, 0 \Big\}.$$

Now, since $\mathcal{E}$ does not depend on n, and $\max, \min$ are continuous operations and quotient is continuous since for all $t \in \{1, \dots, T\}$, $p_t \cdot x_t^0 > 0$, we have

$$\lim_{n \to \infty} I_{\text{Afr}}\big((p_1, x_1^n), \dots, (p_T, x_T^n)\big) = I_{\text{Afr}}\big((p_1, x_1^0), \dots, (p_T, x_T^0)\big).$$

Second, assume that there exists $t \in \{1, \dots, T\}$ such that $p_t \cdot x_t = 0$. Since p_t is strictly positive, and x_t is non-negative, this is possible only if $x_t = 0$. Now especially for all $t' \neq t$, $x_{t'} \neq 0$. Additionally since for all $x \in \mathbb{R}_+^g \setminus \{0\}$, and $p \in \mathbb{R}_{++}^g$, $p \cdot x > 0$, there exists $n^0 \in \mathbb{N}$ such that for all $n > n^0$ and $t' \in \{1, \dots, T\} \setminus \{t\}$, $p_{t'} \cdot x_{t'} > p_{t'} \cdot x_t = 0$. Thus x_t is not a part of any cycle. So for all $n > n^0$ or $n = 0$,

$$I_{\text{Afr}}\big((p_1, x_1^n), \dots, (p_T, x_T^n)\big) = I_{\text{Afr}}\big((p_1, x_1^n), \dots, (p_{t-1}, x_{t-1}^n), (p_{t+1}, x_{t+1}^n), \dots, (p_T, x_T^n)\big).$$

So the claim follows from the first part.

Axioms 3.3-3.5: Follow directly from the above used definition.

Next, we will show the sufficiency of these axioms. First, by above, I_{Afr} satisfies Axiom 3.1. Let $D = \big((p_1, x_1), \dots, (p_T, x_T)\big) \in \mathcal{D}$ be such that for all $1 \leq i \leq T$, $p_i \cdot x_i > 0$. Then the claim follows from Lemmas C.2, C.5, and C.6. By above, I_{Afr} satisfies Axiom 3.2, and so this shows the claim for all $D \in \mathcal{D}$ by taking the limit. □

C.3.3 Normalized Money Pump Index

Lemma C.8 Let I satisfy Axiom 3.3^{NMPI} and let $D \in \mathcal{D}$ be unit income normalized data. Then there exists unit income normalized $D' \in \mathcal{D}$ such that for all $c, c' \in \mathcal{C}^{D'}, c \neq c',\ c \cap c' = \varnothing$ and $I(D) = I(D')$.

Proof. Let $T' = \sum_{c \in \mathcal{C}^D} |c|$. We enumerate $\mathcal{C}^D = \{c_1 \ldots, c_n\}$. For each $i \in \{1, \ldots, n\}, j \in \{1, \ldots, |c^i|\}$, define $t(i,j)$ as such that $c^i = (x_{t(i,1)}, \ldots, x_{t(i,|c^i|)})$.

Define $A \in \mathbb{R}_{++}^{T' \times T'}$ as such that for all $i, j \in \{1, \ldots, T'\}$,

$$A_{i,j} = \begin{cases} 1, & \text{if } i = j \\ p_{t(k,i-\sum_{l=1}^{k-1}|c_l|)} \cdot x_{t(k,i-\sum_{l=1}^{k-1}|c_l|+1)}, & \text{if } \exists\, k \in \{1,...,n\}, \sum_{l=1}^{k-1}|c_l| < i < \sum_{l=1}^{k}|c_l|, j = i+1 \\ p_{t(k,|c_k|)} \cdot x_{t(k,1)}, & \text{if } \exists\, k \in \{1,...,n\}, i = \sum_{l=1}^{k}|c_l|, j = \sum_{l=1}^{k-1}|c_l| \\ 2, & \text{if } i \neq j, \forall\, k \in \{1,...,n\}, i \neq \sum_{l=1}^{k}|c_l|, j \neq i+1 \\ 2, & \text{if } i \neq j, \exists\, k \in \{1,...,n\}, i = \sum_{l=1}^{k}|c_l|, j \neq \sum_{l=1}^{k-1}|c_l|. \end{cases}$$

By Lemma C.1, there exist $\tilde{P}, \tilde{X}$ such that $\text{trans}(\tilde{P}) * \tilde{X} = A$. Denote

$$\tilde{D} = \left((\text{pr}_{g' \times \{1\}} \tilde{P}, \text{pr}_{g' \times \{1\}} \tilde{X}), \ldots, \text{pr}_{g' \times \{T'\}} \tilde{P}, \text{pr}_{g' \times \{T'\}} \tilde{X}\right) = \left((\tilde{p}_1, \tilde{x}_1), \ldots, (\tilde{p}_{T'}, \tilde{x}_{T'})\right).$$

By the definition of A, we have for all $i \in \{1, \ldots, n\}, k \in \{1, \ldots, |\,|c| - 1\}$,

$$p_{t(i,k)} \cdot (x_{t(i,k)} - t(i, k+1)) = \tilde{p}_{\sum_{l=1}^{i-1}|c_l|+k} \cdot (\tilde{x}_{\sum_{l=1}^{i-1}|c_l|+k} - \tilde{x}_{\sum_{l=1}^{i-1}|c_l|+k+1})$$

and

$$p_{t(i,|c_i|)} \cdot (x_{t(i,|c_i|)} - t(i,1)) = \tilde{p}_{\sum_{l=1}^{i}|c_l|} \cdot (\tilde{x}_{\sum_{l=1}^{i}|c_l|} - \tilde{x}_{\sum_{l=1}^{i-1}|c_l|}).$$

Additionally, by the definition of A, $|\mathcal{C}^D| = |\mathcal{C}^{D'}|$. So there exists an enumeration of $\mathcal{C}^{\tilde{D}} = \{\tilde{c}_1 \ldots, \tilde{c}_n\}$ such that for all $i \in \{1, \ldots, n\}$,

$$|c_i| = |\tilde{c}_i| \text{ and } c_i \geq \tilde{c}_i \geq c_i.$$

By Axiom 3.3^{NMPI}, $I(D) = I(D')$. Additionally for all $i, j \in \{1, \ldots, n\}, i \neq j,\ c_i \cap c_j = \varnothing$. □

Lemma C.9 Let I satisfy Axioms 3.3^{NMPI}, 3.4^{NMPI} and let $D \in \mathcal{D}$ be unit income normalized data such that for all $c, c' \in \mathcal{C}^D, c \neq c',\ c \cap c' = \varnothing$. Then there exists unit income normalized $\tilde{D} \in \mathcal{D}$ such that $I(D) = I(\tilde{D})$, for all $c, c' \in \mathcal{C}^{\tilde{D}}, c \neq c',\ c \cap c' = \varnothing$, and for all $(\tilde{x}_{i_1}, \ldots, \tilde{x}_{i_n}) \in \mathcal{C}^{\tilde{D}}, j, k \in \{1, \ldots, n\}$, $\tilde{p}_{i_j} \cdot (\tilde{x}_{i_j} - \tilde{x}_{i_{j+1}}) = \tilde{p}_{i_k} \cdot (\tilde{x}_{i_k} - \tilde{x}_{i_{k+1}})$.

Proof. Let $d \in \mathbb{N}$ be the least common denominator for the set $\{|c||c \in \mathcal{C}^D\}$. We enumerate $\mathcal{C}^D = \{c_1 \ldots, c_n\}$. For each $i \in \{1, \ldots, n\}, j \in \{1, \ldots, |c^i|\}$, define $t(i,j)$ as such that $c^i = (x_{t(i,1)}, \ldots, x_{t(i,|c^i|)})$. Define a rotation permutation $\pi : T \to T$ by for all $i \in \{1, \ldots, n\}, j \in \{1, \ldots, |c^i|-1\}$, $\pi(t(i,j)) = t(i, j+1)$, $\pi(t(i, |c^i|)) = t(i,1)$, and for all $t \notin \cup_{i=1}^{n} \cup_{j=1}^{|c^i|} t(i,j)$, $\pi(t) = t$. Define for each $n \in \mathbb{N}$, $\pi^n : T \to T$ by for all $t \in T$,

$$\pi^n(t) = \overbrace{\pi \circ \cdots \circ \pi}^{n \text{ times}}(t).$$

Denote

$$B = \bigcup_{k=1}^{n} \bigcup_{l=1}^{|c^k|-1} \{(t(k,l), t(k,l+1))\} \cup \bigcup_{m=1}^{n} \{(t(m,|c^m|), t(m,1))\}$$

For each $n \in \{1, \ldots, d\}$, define $A^n \in \mathbb{R}_{++}^{T \times T}$ as such that for all $i, j \in \{1, \ldots, T\}$,

$$A^n_{i,j} = \begin{cases} 1, & \text{if } i = j \\ p_{(\pi^n)^{-1}(i)} \cdot x_{(\pi^n)^{-1}(j)}, & \text{if } (i,j) \in B \\ 2, & \text{if } i \neq j, (i,j) \notin B, \end{cases}$$

where $(\pi^n)^{-1}$ is the inverse function π^n that exists since π^n is a permutation.

By Lemma C.1, there exists $\tilde{P}$ such that for each $n \in \{1, \ldots, d\}$, there exists $\tilde{X}^n$ such that $\text{trans}(\tilde{P}) * \tilde{X}^n = A^n$. Denote

$$\tilde{D}^n = ((\text{pr}_{g' \times \{1\}} \tilde{P}, \text{pr}_{g' \times \{1\}} \tilde{X}^n), \ldots, \text{pr}_{g' \times \{T\}} \tilde{P}, \text{pr}_{g' \times \{T\}} \tilde{X}^n) = ((\tilde{p}_1, \tilde{x}_1^n), \ldots, (\tilde{p}_T, \tilde{x}_T^n)).$$

Now we have for all $n \in \mathbb{N}$ and $(i,j) \in B$, $p_i \cdot x_j^n \leq 1$, and for all $i \neq j$, $(i,j) \notin B$, $p_i \cdot x_j^n = 2$. Additionally for all $n \in \mathbb{N}$, since π^n is a rotation of cycles,

$$(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D \iff (x_{t_1}^n, \ldots, x_{t_n}^n) \in \mathcal{C}^{\tilde{D}^n}.$$

Finally, by above and Axiom 3.3^{NMPI}, for all $n \in \mathbb{N}$, $I(D) = I(\tilde{D}^n)$. Thus by Axiom 3.4^{NMPI},

$$I(\sum_{n=1}^{d} \frac{1}{d} D^n) = \sum_{n=1}^{d} \frac{1}{d} I(D^n) = I(D).$$

Let $(i,j) \in B$ and $k \in \{1, \ldots, n\}$ be the corresponding cycle that (i,j) is a part of. Now $\frac{d}{|c^k|}$ is an integer.

$$\begin{aligned} & p_i \cdot \Big(\sum_{n=1}^{d} \frac{1}{d} x_i^n - \sum_{m=1}^{d} \frac{1}{d} x_j^m \Big) = \sum_{n=1}^{d} \frac{1}{d} p_i \cdot (x_i^n - x_j^n) \\ & = \sum_{n=1}^{d} \frac{1}{d} (A^n_{i,i} - A^n_{i,j}) = \sum_{n=1}^{d} \frac{1}{d} p_{(\pi^n)^{-1}(i)} \cdot (x_{(\pi^n)^{-1}(i)} - x_{(\pi^n)^{-1}(j)}) \\ & = \frac{1}{d} \sum_{l=1}^{d/|c^k|} \Big(\sum_{l=1}^{|c^k|-1} p_{t(k,|c^k|-l)} \cdot (x_{t(k,|c^k|-l)} - x_{t(k,|c^k|-l+1)}) + p_{t(k,|c^k|)} \cdot (x_{t(k,|c^k|)} - x_{t(k,1)}) \Big) \end{aligned}$$

$$= \frac{1}{|c^k|}\Big(\sum_{l=1}^{|c^k|-1} p_{t(k,l)} \cdot (x_{t(k,l)} - x_{t(k,l+1)}) + p_{t(k,|c^k|)} \cdot (x_{t(k,|c^k|)} - x_{t(k,1)})\Big).$$

Where the second to last equality follows from the definition of π^n. This depends only on (i,j) only through the cycle k and so it shows the claim. □

Lemma C.10 Let I satisfy Axioms 3.2^{NMPI}, 3.3^{NMPI}, 3.4^{NMPI} and let $D \in \mathcal{D}$ be unit income normalized data such that for all $c, c' \in \mathcal{C}^{D'}, c \neq c', c \cap c' = \varnothing$ and for all $(x_{i_1}, \ldots, x_{i_n}) \in \mathcal{C}^D, j, k \in \{1, \ldots, n\}$,

$$p_{i_j} \cdot (x_{i_j} - x_{i_{j+1}}) = p_{i_k} \cdot (x_{i_k} - x_{i_{k+1}}). \tag{C.3}$$

Then

$$I(D) = I_{\text{NMPI}}(D).$$

Proof. We enumerate $\mathcal{C}^D = \{c_1, \ldots, c_n\}$. For each $k \in \{1, \ldots, n\}$, when $c_k = (x_{i_1}, \ldots, x_{i_m})$ define $\theta(i) = p_{i_1} \cdot x_{i_2}$. Denote $T' = 2n$. Define $A \in \mathbb{R}_{++}^{T' \times T'}$ as such that for all $i, j \in \{1, \ldots, T'\}$,

$$A_{i,j} = \begin{cases} 1, & \text{if } i = j \\ \theta(k), & \text{if } \exists\, k \in \{1, \ldots, n\}, i = 2k-1, j = 2k \text{ or } i = 2k, j = 2k-1 \\ 2, & \text{if } i \neq j, \forall\, k \in \{1, \ldots, n\}, (i,j) \neq (2k-1, 2k), (i,j) \neq (2k, 2k-1). \end{cases}$$

By Lemma C.1, there exist $\tilde{P}, \tilde{X}$ such that $\text{trans}(\tilde{P}) * \tilde{X} = A$. Denote

$$\tilde{D} = \big((\text{pr}_{\bar{T}' \times \{1\}} \tilde{P}, \text{pr}_{\bar{T}' \times \{1\}} \tilde{X}), \ldots, \text{pr}_{\bar{T}' \times \{T'\}} \tilde{P}, \text{pr}_{\bar{T}' \times \{T'\}} \tilde{X}\big) = \big((\tilde{p}_1, \tilde{x}_1), \ldots, (\tilde{p}_{T'}, \tilde{x}_{T'})\big).$$

Now $|\mathcal{C}^D| = n = |\mathcal{C}^{\tilde{D}}|$ and by (C.3) and the definition of θ, we have for all $i \in \{1, \ldots, n\}$, $(\tilde{x}_{2i-1}, \tilde{x}_{2i}) \in \mathcal{C}^{\tilde{D}}$ and

$$c_i \gg (\tilde{x}_{2i-1}, \tilde{x}_{2i}) \gg c_i.$$

Thus by Axiom 3.3^{NMPI}, $I(D) = I(\tilde{D})$.

Define a rotation permutation $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ by for all $i \in \{1, \ldots, n-1\}$, $\pi(i) = i+1$ and $\pi(n) = 1$. Define for each $m \in \mathbb{N}$, $\pi^m : T \to T$ by for all $i \in \{1, \ldots, n\}$,

$$\pi^m(i) = \overbrace{\pi \circ \cdots \circ \pi}^{m \text{ times}}(i).$$

For each $m \in \{1, \ldots, n\}$, define $A^m \in \mathbb{R}_{++}^{T' \times T'}$ as such that for all $i, j \in \{1, \ldots, T'\}$,

$$A^m_{i,j} = \begin{cases} 1, & \text{if } i = j \\ \theta\big((\pi^m)^{-1}(k)\big), & \text{if } \exists\, k \in \{1, \ldots, n\}, (i,j) \in \{(2k-1, 2k), (2k, 2k-1)\} \\ 2, & \text{if } i \neq j, (i,j) \notin \bigcup_{k=1}^{n} \{(2k, 2k-1), (2k-1, 2k)\}. \end{cases}$$

where $(\pi^m)^{-1}$ is the inverse function π^m that exists since π^m is a permutation.

By Lemma C.1, there exists $\hat{P}$ such that for each $n \in \{1, \dots, d\}$, there exists $\hat{X}^n$ such that $\text{trans}(\hat{P}) * \hat{X}^n = A^n$. Denote

$$\hat{D}^n = \big((\text{pr}_{g' \times \{1\}} \hat{P}, \text{pr}_{g' \times \{1\}} \hat{X}^n), \dots, \text{pr}_{g' \times \{T'\}} \hat{P}, \text{pr}_{g' \times \{T'\}} \hat{X}^n\big) = \big((\hat{p}_1, \hat{x}_1^n), \dots, (\hat{p}_{T'}, \hat{x}_{T'}^n)\big).$$

Now for all $m \in \{1, \dots, n\}$, $|\mathcal{C}^D| = n = |\mathcal{C}^{\hat{D}^m}|$ and we have for all $i \in \{1, \dots, n\}$, $(\hat{x}_{2i-1}^m, \hat{x}_{2i}^m) \in \mathcal{C}^{\hat{D}}$ and

$$(\tilde{x}_{2i-1}, \tilde{x}_{2i}) \gtrdot (\hat{x}_{2\pi^m(i)-1}^m, \hat{x}_{2\pi^m(i)}^m) \gtrdot (\tilde{x}_{2i-1}, \tilde{x}_{2i})$$

and so by Axiom 3.3^{NMPI}, $I(D) = I(\tilde{D}) = I(\hat{D}^m)$.

Denote

$$B = \bigcup_{k=1}^{n} \{(2k-1, 2k), (2k, 2k-1)\}$$

Now we have for all $m \in \{1, \dots, n\}, (i,j) \in B$, $\hat{p}_i \cdot \hat{x}_j^m \leq 1$, and for all $i \neq j$, $(i,j) \notin B$, $\hat{p}_i \cdot \hat{x}_j^m = 2$. Additionally, for all $m \in \{1, \dots, n\}$,

$$\mathcal{C}^{\hat{D}^m} = \bigcup_{k=1}^{n} \{(\hat{x}_{2k-1}^m, \hat{x}_{2k}^m)\}.$$

Thus by Axiom 3.4^{NMPI},

$$I\Big(\sum_{m=1}^{n} \frac{1}{n} \hat{D}^m\Big) = \sum_{m=1}^{n} \frac{1}{n} I(\hat{D}^m) = I(D).$$

Let $k \in \{1, \dots, n\}$ and either $i = 2k-1$ and $j = 2k$ or $i = 2k$ and $j = 2k-1$. Then we have

$$\hat{p}_i \cdot \Big(\sum_{m=1}^{n} \frac{1}{n} \hat{x}_i^m - \sum_{m=1}^{n} \frac{1}{n} \hat{x}_j^m\Big) = \sum_{m=1}^{n} \frac{1}{n} p_i \cdot (\hat{x}_i^m - \hat{x}_j^m) = \sum_{m=1}^{n} \frac{1}{n} (A_{i,i}^m - A_{i,j}^m) \tag{C.4}$$

$$= \sum_{m=1}^{n} \frac{1}{n} \big[1 - \theta\big((\pi^m)^{-1}(k)\big)\big] = \sum_{m=1}^{n} \frac{1}{n} \big[1 - \theta\big((k - m \bmod n)\big)\big] = \frac{1}{n} \sum_{m=1}^{n} \big[1 - \theta(m)\big]$$

Where the second to last equality follows from the definition of π^m. This does not depend on (i,j). Now

$$\mathcal{C}^{\sum_{m=1}^{n} \frac{1}{n} D^m} = \bigcup_{k=1}^{n} \Big\{\Big(\sum_{m=1}^{n} \frac{1}{n} \hat{x}_{2k-1}^m, \sum_{m=1}^{n} \frac{1}{n} \hat{x}_{2k}\Big)\Big\}.$$

By (C.4) and Axiom 3.2^{NMPI}

$$I(D) = I\Big(\sum_{m=1}^{n} \frac{1}{n} D^m\Big) = \frac{1}{n} \sum_{m=1}^{n} \big[1 - \theta(m)\big] \overset{\text{(C.3)}}{=} \frac{1}{\mathcal{C}^D} \sum_{(x_{t_1}, \dots, x_{t_n}) \in \mathcal{C}^D} \frac{1}{n} \sum_{i=1}^{n} \frac{p_{t_i} \cdot (x_{t_i} - x_{t_{i+1}})}{p_{t_i} \cdot x_{t_i}} = I_{\text{NMPI}}(D).$$

□

Theorem C.11 Let I be a measure of rationality on $\mathcal{D}$. I satisfies Axiom 3.1 and Axioms 3.1^{NMPI}-3.4^{NMPI} iff

$$I = I_{\text{NMPI}}.$$

Proof. We show first the only if direction. Axioms 3.1 and 3.1^{NMPI} follow as in the proof of Theorem 3.1. Axioms 3.2^{NMPI} and 3.3^{NMPI} follow from averaging a constant number and from monotonicity of average tradewise and cyclewise. Axiom 3.4^{NMPI} follows from the observation that if $(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D$ and $(x'_{t_1}, \ldots, x'_{t_n}) \in \mathcal{C}^{D'}$, then

$$(\alpha x_{t_1} + (1-\alpha)x'_{t_1}, \ldots, \alpha x_{t_n} + (1-\alpha)x'_{t_n}) \in \mathcal{C}^{\alpha D + (1-\alpha)D'}.$$

Thus the condition $\left|\mathcal{C}^D\right| = \left|\mathcal{C}^{\alpha D+(1-\alpha)D'}\right|$ guarantees that the set of cycles in the convex combination is the same as in D as a function of indices. Additionally, we have for all t, t', since the data are unit income normalized,

$$\frac{p_t \cdot (\alpha x_t + (1-\alpha)x'_t - \alpha x_{t'} + (1-\alpha)x'_{t'})}{p_t \cdot (\alpha x_t + (1-\alpha)x'_t)} = \alpha p_t \cdot (x_t - x_{t'}) + (1-\alpha) p_t \cdot (x'_t - x'_{t'}).$$

The axiom then follows from the linearity of arithmetic average since the averaged set remains the same in the convex combination.

Next, we show the if direction. First, let $D \in \mathcal{D}$ be such that for all $1 \leq i \leq T$, $p_i \cdot x_i > 0$. Then by the only if direction and Lemmas C.2 and C.8 to C.10, $I(D) = I_{\text{NMPI}}(D)$.

Finally, let $D \in \mathcal{D}$ be such that for some $1 \leq i^* \leq T$, $p_{i^*} \cdot x_{i^*} = 0$ and $g \in \mathbb{N}$ be the number of goods. Since p_{i^*} is strictly positive and x_{i^*} is non-negative, we must have $x_{i^*} = 0$. Since $D \in \mathcal{D}$ without duplicate quantities, we must have for all $j \in \{1, \ldots, T\} \setminus \{i^*\}$, $x_j \neq 0$. Let $(x_i^n) \subseteq \mathbb{R}^g_{++} \setminus \{0\}$ be such that for all $j \in \{1, \ldots, T\}$, $x_i^n \neq x_j$ and $x_i^n \to 0$ as $n \to \infty$. For each $n \in \mathbb{N}$, denote $D^n = ((p_1, x_1), \ldots, (p_i, x_i^n), \ldots, (p_T, x_T))$. By the above, only if direction, and Axiom 3.1^{NMPI}, we have

$$I(D) = \lim_{n\to\infty} I(D^n) = \lim_{n\to\infty} I_{\text{NMPI}}(D^n) = I_{\text{NMPI}}(D).$$

□

C.3.4 Normalized Minimum Cost Index

Lemma C.12 Let I satisfy Axioms 3.2-3.3 and 3.2^{NMCI}. Then for all total income normalized $D \in \mathcal{D}$,

$$I_{\text{NMCI}}(D) \geq I(D).$$

Proof. Let $D = ((p_1, x_1), \ldots, (p_T, x_T))$. Let $B \subseteq \{T \times T\}$ be the solution to the minimization problem from the definition of NMCI since the data is total income normalized

$$I_{\text{NMCI}}(D) = \min_{B \subseteq \text{P}} \sum_{(x_t, x_{t'}) \in B} p_t \cdot (x_t - x_{t'}) \text{ such that P} \setminus B \text{ is acyclical.}$$

Define $A \in \mathbb{R}^{T \times T}_{++}$ as such that for all $i, j \in \{1, \ldots, T\}$,

$$A_{i,j} = \begin{cases} \frac{1}{T}, & \text{if } (i,j) \in B \\ p_i \cdot x_j, & \text{if } (i,j) \notin B \end{cases}.$$

By Lemma C.1, there exist $\tilde{P}, \tilde{X}$ such that $\text{trans}(\tilde{P}) * \tilde{X} = A$. Denote

$$\tilde{D} = ((\text{pr}_{\bar{T} \times \{1\}} \tilde{P}, \text{pr}_{\bar{T} \times \{1\}} \tilde{X}), \ldots, \text{pr}_{\bar{T} \times \{T\}} \tilde{P}, \text{pr}_{\bar{T} \times \{T\}} \tilde{X}) = ((\tilde{p}_1, \tilde{x}_1), \ldots, (\tilde{p}_T, \tilde{x}_T)).$$

By Lemma C.4 and the definition of B, we have $I(\tilde{D}) = 0$. Additionally by Axiom 3.2^{NMCI}, we have

$$I(D) \leq I(\tilde{D}) + \sum_{(i,j) \in B} p_i \cdot (x_i - x_j) = I_{\text{NMCI}}(D).$$

□

Lemma C.13 Let I satisfy Axioms 3.1-3.3 and 3.3^{NMCI}. Then for all total income normalized $D \in \mathcal{D}$,

$$I_{\text{RNMCI}}(D) \leq I(D).$$

Proof. Define the collection of cycles on time indices as

$$\mathcal{I}^D = \{((t_1, t_2), (t_2, t_3), \ldots, (t_{n-1}, t_n), (t_n, t_1)) \subseteq T \times T \big| (x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D\}.$$

$I_{\text{RNMCI}}(D)$ is the value to the linearized cycle covering problem

$$I_{\text{RNMCI}}(D) = \frac{1}{T} \min_{[0,1]^{\text{P}}} \sum_{(x_t, x_{t'}) \in \text{P}} \beta(x_t, x_{t'}) \frac{p_t \cdot (x_t - x_{t'})}{p_t \cdot x_t} \text{ s.t. P} - \beta \text{ is acyclical}$$

$$= \min_{\beta \in \mathbb{R}^{T \times T}_+} \sum_{(t,t') \in T \times T} \beta(t, t') \max\{p_t \cdot (x_t - x_{t'}), 0\} \text{ s.t.} \forall ((t_1, t_2), ..., (t_n, t_1)) \in \mathcal{I}^D, \sum_{i=1}^{n} \beta(t_i, t_{i+1}) \geq 1.$$

This follows from the following observations: D is total income normalized and so $p_t \cdot x_t = \frac{1}{T}$. For all cycles $(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D$, $\sum_{i=1}^{n} (x_{t_i} \not\mathrel{P} x_{t_{i+1}}) = 0$. So P$-\beta$ is acyclical iff for all cycles $(x_{t_1}, \ldots, x_{t_n}) \in \mathcal{C}^D$, $\sum_{i=1}^{n} \beta(x_{t_i}, x_{t_{i+1}}) \geq 1$. It is not never optimal to have $\beta(x_t, x_{t'}) > 1$ if $p_t \cdot (x_t - x_{t'}) > 0$. So we can relax the problem to have all non-negative values for β. Finally, cycles only use pairs in P, so we can relax the problem to have β on $T \times T$ when objective is changed to $\max\{p_t \cdot (x_t - x_{t'}), 0\}$.

With an abuse of notation, for $c = ((t_1, t_2), \ldots, (t_n, t_1)))$, write $(t, t') \in c$ if there exists $1 \leq i \leq n$ such that $(t_i, t_{i+1}) = (t, t')$. The dual of this problem is

$$\max_{\gamma \in \mathbb{R}^{\mathcal{I}^D}_+} \sum_{c \in \mathcal{I}^D} \gamma(c) \text{ such that } \forall (t, t') \in T \times T, \sum_{c \in \mathcal{I}^D | (t,t') \in c} \gamma(c) \leq \max\{p_t \cdot (x_t - x_{t'}), 0\}, \tag{C.5}$$

where the sum in the constraint is over cycles c such that $(t,t') \in c$. By strong duality theorem, the value of this problem is $I_{\mathrm{RNMCI}}(D)$. Let γ^* be a solution to the problem.

We enumerate the cycles in $\mathcal{I}^D = \{c_1, \dots, c^{n^I}\}$. For $(i,j) \in T \times T$ and $c_l \in \mathcal{I}^D$, denote $\mathbb{1}((i,j) \in c_l)$ for an indicator function. For each $k \in \{0, \dots, n^I\}, \alpha \in [0,1]$, define $A^{k,\alpha} \in \mathbb{R}_{++}^{T \times T}$ as such that for all

$$A_{i,j}^{k,\alpha} = p_i \cdot x_j + \sum_{l=1}^{k-1} \mathbb{1}((i,j) \in c_l)\gamma^*(c_l) + \alpha \mathbb{1}((i,j) \in c_k)\gamma^*(c_k). \tag{C.6}$$

By Lemma C.1, there exists $\tilde{P}$ such that for all $k \in \{0, \dots, n^I\}, \alpha \in [0,1]$, there exists $\tilde{X}^{k,\alpha}$ such that $\mathrm{trans}(\tilde{P}) * \tilde{X}^{k,\alpha} = A^{k,\alpha}$. Denote

$$\tilde{D}^{k,\alpha} = ((\mathrm{pr}_{\bar{T} \times \{1\}} \tilde{P}, \mathrm{pr}_{\bar{T} \times \{1\}} \tilde{X}^{k,\alpha}), \dots, \mathrm{pr}_{\bar{T} \times \{T\}} \tilde{P}, \mathrm{pr}_{\bar{T} \times \{T\}} \tilde{X}^{k,\alpha}) = ((\tilde{p}_1, \tilde{x}_1^{k,\alpha}), \dots, (\tilde{p}_T, \tilde{x}_T^{k,\alpha})).$$

Now especially for all $i \in T$, $\gamma^*(i,i) = 0$ and so $\tilde{D}^{k,\alpha}$ is total income normalized.

We show that for all $k \in \{1, \dots, n^I\}$,

$$I(\tilde{D}^{k-1,1}) \geq I(\tilde{D}^{k,1}) + \gamma^*(c_k). \tag{C.7}$$

Let $k \in \{1, \dots, n^I\}$. First, assume that $\gamma^*(c_k) = 0$. Then $A^{k-1,1} = A^{k,1}$ and since $\tilde{D}^{k-1,\alpha}, \tilde{D}^{k,\alpha}$ are total income normalized, by Axiom 3.1, after a price scaling we can apply Axiom 3.3 to get that $I(\tilde{D}^{k-1,\alpha}) = I(\tilde{D}^{k,\alpha})$. This shows (C.7) in this case.

Next, assume that $\gamma^*(c_k) > 0$. Let $\alpha \in [0,1)$. We show that $I(\tilde{D}^{k-1,1}) \geq I(\tilde{D}^{k,\alpha}) + \alpha\gamma^*(c_k)$. By (C.5) and since $\alpha < 1, \gamma^*(c_k) > 0$, we have for all $(t,t') \in c_k$,

$$\sum_{l=1}^{k-1} \mathbb{1}((t,t') \in c_l)\gamma^*(c_l) + \alpha\gamma^*(c_k) < p_t \cdot (x_t - x_{t'}).$$

Thus we have for all $(t,t') \in c_k$, by (C.6),

$$\tilde{p}_t \cdot (\tilde{x}_t^{k,\alpha} - \tilde{x}_{t'}^{k,\alpha}) > 0$$

and

$$\tilde{p}_t \cdot (\tilde{x}_t^{k-1,1} - \tilde{x}_{t'}^{k-1,1}) = \frac{1}{T} - p_i \cdot x_j - \sum_{l=1}^{k-1} \mathbb{1}((i,j) \in c_l)\gamma^*(c_l)) = \tilde{p}_t \cdot (\tilde{x}_t^{k,\alpha} - \tilde{x}_{t'}^{k,\alpha}) + \alpha\gamma^*(c_k).$$

Additionally, we have for all $(t,t') \notin c_k$, by (C.6),

$$\tilde{p}_t \cdot (\tilde{x}_t^{k-1,1} - \tilde{x}_{t'}^{k-1,1}) = \tilde{p}_t \cdot (\tilde{x}_t^{k,\alpha} - \tilde{x}_{t'}^{k,\alpha}).$$

Thus by Axiom 3.3^{NMCI}, for all $\alpha \in [0,1)$,

$$I(\tilde{D}^{k-1,1}) \geq I(\tilde{D}^{k,\alpha}) + \alpha\gamma^*(c_k).$$

By Lemma C.1, $\tilde{X}^{k,\alpha}$ is continuous in α. Thus we can take $\alpha \to 1$ to have by Axiom 3.2,

$$I(\tilde{D}^{k-1,1}) \geq \lim_{\alpha \to 1} I(\tilde{D}^{k,\alpha}) + \alpha\gamma^*(c_k) = I(\tilde{D}^{k,1}) + \gamma^*(c_k).$$

This shows (C.7) since $k \in \{1, \ldots, n^I\}$ was arbitrary.

Next, we show that $I(D) = I(\tilde{D}^{0,1})$. By (C.6), for all $(t, t') \in T \times T$, we have

$$\tilde{p}_t \cdot (\tilde{x}_t^{k,\alpha} - \tilde{x}_{t'}^{k,\alpha}) = p_t \cdot x_{t'}.$$

After rescaling with Axiom 3.1, we have by Axiom 3.3, $I(D) = I(\tilde{D}^{0,1})$.

Finally, by Equation (C.7), we have

$$I(D) = I(\tilde{D}^{0,1}) \geq I(\tilde{D}^{1,1}) + \gamma^*(c_1) \geq \cdots \geq I(\tilde{D}^{n^I,1}) + \sum_{l=1}^{n^I} \gamma^*(c_l).$$

Since I is a measure of rationality and non-negative, we have

$$I(D) \geq I(\tilde{D}^{n^I,1}) + \sum_{l=1}^{n^I} \gamma^*(c_l) \geq \sum_{l=1}^{n^I} \gamma^*(c_l) = I_{\text{RNMCI}}(D).$$

□

Theorem C.14 If a measure of rationality I satisfies Axioms 3.1-3.3 and Axioms 3.1^{NMCI}-3.3^{NMCI}, then for all $D \in \mathcal{D}$,

$$I_{\text{NMCI}}(D) \geq I(D) \geq I_{\text{RNMCI}}(D).$$

Additionally, all the above axioms are necessary for I_{NMCI} and I_{RNMCI}.

Proof. We show first the necessity of the axioms. Denote the collection of all the possible indices for cycles for data of length T by

$$\mathcal{E} = \{(i_1, \ldots, i_n) \subseteq \{1, \ldots, T\} | n \in \mathbb{N}, \forall\, j, k \{1, \ldots, n\}, j \neq k, i_j \neq i_k\}.$$

First, we observe that if $D = \big((p_1, x_1), \ldots, (p_T, x_T)\big) \in \mathcal{D}$ is such that for all $1 \leq i \leq T$, $p_i \cdot x_i > 0$, then we can write NMCI and RNMCI as follows

$$I_{\text{NMCI}}(D) =$$
$$\frac{1}{T} \min_{\beta \in \{0,1\}^{T \times T}} \sum_{(t,t') \in T \times T} \beta(t, t') \max\Big\{\frac{p_t \cdot (x_t - x_{t'})}{p_t \cdot x_t}, 0\Big\} \text{ s.t. } \forall (i_1, \ldots, i_n) \in \mathcal{E}, \sum_{j=1}^{n} \beta(i_j, i_{j+1}) \geq 1$$

and

$$I_{\text{RNMCI}}(D) =$$
$$\frac{1}{T} \min_{\beta \in [0,1]^{T \times T}} \sum_{(t,t') \in T \times T} \beta(t, t') \max\Big\{\frac{p_t \cdot (x_t - x_{t'})}{p_t \cdot x_t}, 0\Big\} \text{ s.t. } \forall (i_1, \ldots, i_n) \in \mathcal{E}, \sum_{j=1}^{n} \beta(i_j, i_{j+1}) \geq 1$$

by the same observations as in the proof of Lemma C.13.

Axioms 3.1 and 3.2: Follows using the above definitions as in Theorem 3.1. By Axiom 3.2, dropping an observation that is not a part of any cycles only affects the normalization $\frac{1}{T}$.

Axioms 3.1^{NMCI} to 3.3^{NMCI} follow directly from the observation that if D is total income normalized, then as in the proof of Lemma C.13, we can write NMCI and RNMCI as

$$I_{\text{NMCI}}(D)$$
$$= \min_{\beta \in \{0,1\}^{T\times T}} \sum_{(t,t')\in T\times T} \beta(t,t') \max\{p_t \cdot (x_t - x_{t'}), 0\} \text{ s.t. } \forall (i_1,\ldots,i_n) \in \mathcal{E}, \sum_{j=1}^{n} \beta(i_j, i_{j+1}) \geq 1$$

and

$$I_{\text{RNMCI}}(D)$$
$$= \min_{\beta \in [0,1]^{T\times T}} \sum_{(t,t')\in T\times T} \beta(t,t') \max\{p_t \cdot (x_t - x_{t'}), 0\} \text{ s.t. } \forall (i_1,\ldots,i_n) \in \mathcal{E}, \sum_{j=1}^{n} \beta(i_j, i_{j+1}) \geq 1.$$

So then Axiom 3.2^{NMCI} follows from for all (t,t'), $\beta(t,t') \leq 1$ and Axiom 3.3^{NMCI} follows from $\forall (i_1,\ldots,i_n) \in \mathcal{E}, \sum_{j=1}^{n} \beta(i_j, i_{j+1}) \geq 1$.

Next, we show that if a measure of rationality I satisfies Axioms 3.1-3.3 and Axioms 3.1^{NMCI}-3.3^{NMCI}, then for all $D \in \mathcal{D}$,

$$I_{\text{NMCI}}(D) \geq I(D) \geq I_{\text{RNMCI}}(D).$$

First, let $D = ((p_1,x_1),\ldots,(p_T,x_T)) \in \mathcal{D}$ be such that for all $1 \leq i \leq T$, $p_i \cdot x_i > 0$. Now by the necessity part and as in Lemma C.2, we can normalize D for a total income normalized data $\tilde{D}$ such that $I(D) = I(\tilde{D})$, $I_{\text{NMCI}}(D) = I_{\text{NMCI}}(\tilde{D})$, and $I_{\text{RNMCI}}(D) = I_{\text{RNMCI}}(\tilde{D})$. By Lemmas C.12 and C.13, we have

$$I_{\text{NMCI}}(\tilde{D}) \geq I(\tilde{D}) \geq I_{\text{RNMCI}}(\tilde{D})$$

and so

$$I_{\text{NMCI}}(D) \geq I(D) \geq I_{\text{RNMCI}}(D). \tag{C.8}$$

Next, if $D = ((p_1,x_1),\ldots,(p_T,x_T)) \in \mathcal{D}$ is such that there exists $1 \leq i \leq T$ such that $p_i \cdot x_i = 0$, then the claim follows from continuity, Axiom 3.2, and (C.8). □

Bibliography

Abdellaoui, Mohammed; Li, Chen; Wakker, Peter P., and Wu, George (2020). A Defense of Prospect Theory in Bernheim & Sprenger's Experiment.

Aczél, János and Daróczy, Zoltán (1975). *On Measures of Information and their Characterizations*. Vol. 115. Mathematics in Science and Engineering. Academic Press.

Afriat, Sydney N. (1967). The construction of utility functions from expenditure data. *International economic review* 8(1), pp. 67–77.

— (1972). Efficiency estimation of production functions. *International economic review*, pp. 568–598.

Alefeld, G. E.; Potra, F. A., and Shi, Yixun (1995). Algorithm 748: Enclosing Zeros of Continuous Functions. *ACM Trans. Math. Softw.* 21(3), pp. 327–344.

Amarante, Massimiliano (2009). Foundations of neo-Bayesian statistics. *Journal of Economic Theory* 144(5), pp. 2146–2173.

Anscombe, Francis J. and Aumann, Robert J. (1963). A Definition of Subjective Probability. *The annals of mathematical statistics* 34(1), pp. 199–205.

Apesteguia, Jose and Ballester, Miguel (2015). A measure of rationality and welfare. *Journal of Political Economy* 123(6), pp. 1278–1310.

Arrow, Kenneth J. (1951). Alternative Approaches to the Theory of Choice in Risk-Taking Situations. *Econometrica* 19(4), pp. 404–437.

— (1971). *Essays in the Theory of Risk-Bearing*. Chicago: Markham Publishing Company.

Aumann, Robert J. (1971). *Letter from Robert Aumann to Leonard Savage, 8 January 1971*. Published in Drèze (1987).

Austin, A. K. (1985). 69.8 Two Curiosities. *The Mathematical Gazette* 69(447), pp. 42–44.

Baccelli, Jean (2019). The Problem of State-Dependent Utility: A Reappraisal. *The British Journal for the Philosophy of Science*.

Bernheim, B. Douglas and Sprenger, Charles (2020). On the Empirical Validity of Cumulative Prospect Theory: Experimental Evidence of Rank-Independent Probability Weighting. *Econometrica* 88(4), pp. 1363–1409.

Bewley, Truman F. (2002). Knightian decision theory. Part I. *Decisions in Economics and Finance* 25(2). First version 1986, pp. 79–110.

Birnbaum, Michael H. (1999). The Paradoxes of Allais, Stochastic Dominance, and Decision Weights. *Decision Science and Technology*. Ed. by James Shanteau; Barbara A. Mellers, and David A. Schum. Springer, pp. 27–52.

— (2008). New Paradoxes of Risky Decision Making. *Psychological Review* 115(2), pp. 463–501.

Birnbaum, Michael H. (2018). Behavioral models of decision making under risk. *Psychological Perspectives on Risk and Risk Analysis: Theory, Models and Applications*. Ed. by Martina Raue; Eva Lermer, and Bernhard Streicher. Springer Verlag, pp. 181–200.

Blavatskyy, Pavlo R. (2008). Stochastic utility theorem. *Journal of Mathematical Economics* 44(11), pp. 1049–1056.

Border, Kim C. and Segal, Uzi (1994). Dynamic Consistency Implies Approximately Expected Utility Preferences. *Journal of Economic Theory* 63(2), pp. 170–188.

Bossert, Walter and Kamaga, Kohei (2020). An axiomatization of the mixed utilitarian–maximin social welfare orderings. *Economic Theory* 69(2), pp. 451–473.

Boyd, Stephen and Vandenberghe, Lieven (2004). *Convex Optimization*. Cambridge University Press.

Brent, Richard P. (1973). *Algorithms for minimization without derivatives*. Prentice-Hall.

Carvalho, Leandro and Silverman, Dan (2019). Complexity and Sophistication.

Castagnoli, Erio and Calzi, Marco Li (1996). Expected Utility without Utility. *Theory and Decision* 41(3), pp. 281–301.

Cerreia-Vioglio, Simone; Dillenberger, David, and Ortoleva, Pietro (2015). Cautious Expected Utility and the Certainty Effect. *Econometrica* 83(2), pp. 693–728.

— (2020). An explicit representation for disappointment aversion and other betweenness preferences. *Theoretical Economics* 15(4), pp. 1509–1546.

Cerreia-Vioglio, Simone; Ghirardato, Paolo; Maccheroni, Fabio; Marinacci, Massimo, and Siniscalchi, Marciano (2011). Rational preferences under ambiguity. *Economic Theory* 48(2-3), pp. 341–375.

Chambers, Christopher P. and Echenique, Federico (2016). *Revealed Preference Theory*. Econometric Society Monographs. Cambridge University Press.

Chandrasekher, Madhav; Frick, Mira; Iijima, Ryota, and Le Yaouanq, Yves (2020). Dual-self representations of ambiguity preferences.

Chew, Soo Hong and Wang, Wenqian (2020). On the robustness of indeterminacy in subjective probability. *Economics Letters* 188, p. 108939.

Choi, Syngjoo; Kariv, Shachar; Müller, Wieland, and Silverman, Dan (2014). Who Is (More) Rational? *American Economic Review* 104(6), pp. 1518–50.

Clarke, Frank H. (1983). *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society series of monographs and advanced texts. Wiley.

Conlisk, John (1993). The Utility of Gambling. *Journal of Risk and Uncertainty* 6(3), pp. 255–275.

Daróczy, Zoltán (1970). Generalized Information Functions. *Information and Control* 16(1), pp. 36–51.

Dean, Mark and Martin, Daniel (2016). Measuring rationality with the minimum cost of revealed preference violations. *Review of Economics and Statistics* 98(3), pp. 524–534.

Debreu, Gérard (1954). Representation of a preference ordering by a numerical function. *Decision processes*. Ed. by M. Thrall; R.C. Davis, and C.H. Coombs. John Wiley and Sons, pp. 159–165.

— (1960). Topological methods in cardinal utility theory. *Mathematical models in the social science*. Ed. by Kenneth J. Arrow; Samuel Karlin, and Patrick Suppes. Stanford University Press, pp. 16–26.

Diecidue, Enrico; Schmidt, Ulrich, and Wakker, Peter P. (2004). The Utility of Gambling Reconsidered. *Journal of Risk and Uncertainty* 29(3), pp. 241–259.

Drèze, Jacques (1958). *Individual Decision Making under Partially Controllable Uncertainty*. PhD thesis. Columbia University. Republished in 1971.

— (1961). Les fondements logiques de l'utilité cardinale et de la probabilité subjective. *La décision*, pp. 73–97. Translated and published as Chapter 3 in Drèze (1987).

— (1987). *Essays on economic decisions under uncertainty*. Cambridge University Press.

Drèze, Jacques and Rustichini, Aldo (2004). State-Dependent Utility and Decision Theory. *Handbook of Utility Theory: Volume 2 Extensions*. Ed. by Salvador Barberà; Peter Hammond, and Christian Seidl. Springer Science & Business Media, pp. 839–892.

Echenique, Federico; Lee, Sangmok, and Shum, Matthew (2011). The money pump as a measure of revealed preference violations. *Journal of Political Economy* 119(6), pp. 1201–1223.

Einy, Ezra (1989). On preference relations which satisfy weak independence property. *Journal of Mathematical Economics* 18(3), pp. 291–300.

Fishburn, Peter C. (1970). *Utility Theory for Decision Making*. Wiley, New York.

— (1983). Transitive measurable utility. *Journal of Economic Theory* 31(2), pp. 293–317.

Ghirardato, Paolo; Maccheroni, Fabio, and Marinacci, Massimo (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory* 118(2), pp. 133–173.

Ghirardato, Paolo and Marinacci, Massimo (2002). Ambiguity Made Precise: A Comparative Foundation. *Journal of Economic Theory* 102(2), pp. 251–289.

Gilboa, Itzhak; Maccheroni, Fabio; Marinacci, Massimo, and Schmeidler, David (2010). Objective and Subjective Rationality in a Multiple Prior Model. *Econometrica* 78(2), pp. 755–770.

Gilboa, Itzhak and Schmeidler, David (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics* 18(2), pp. 141–153.

Grant, Simon; Kajii, Atsushi; Polak, Ben, and Safra, Zvi (2010). Generalized Utilitarianism and Harsanyi's Impartial Observer Theorem. *Econometrica* 78(6), pp. 1939–1971.

Grant, Simon and Polak, Ben (2013). Mean-dispersion preferences and constant absolute uncertainty aversion. *Journal of Economic Theory* 148(4), pp. 1361–1398.

Hansen, Lars Peter and Sargent, Thomas J. (2001). Robust Control and Model Uncertainty. *American Economic Review* 91(2), pp. 60–66.

Hara, Kazuhiro; Ok, Efe A., and Riella, Gil (2019). Coalitional Expected Multi-Utility Theory. *Econometrica* 87(3), pp. 933–980.

Harsanyi, John C. (1953). Cardinal Utility in Welfare Economics and in the Theory of Risk-taking. *Journal of Political Economy* 61(5), pp. 434–435.

— (1955). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy* 63(4), pp. 309–321.

— (1977). *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*. Cambridge University Press.

Herstein, Israel N. and Milnor, John (1953). An Axiomatic Approach to Measurable Utility. *Econometrica* 21(2), pp. 291–297.

Hill, Brian (2019). A Non-Bayesian Theory of State-Dependent Utility. *Econometrica* 87(4), pp. 1341–1366.

Hong, Chew Soo (1983). A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox. *Econometrica* 51(4), pp. 1065–1092.

Houtman, Martijn and Maks, J. A. H. (1985). Determining All Maximal Data Subsets Consistent with Revealed Preference. *Kwantitatieve methoden* 19(1), pp. 89–104.

Huck, Steffen and Weizsäcker, Georg (1999). Risk, complexity, and deviations from expected-value maximization: Results of a lottery choice experiment. *Journal of Economic Psychology* 20(6), pp. 699–715.

Humphrey, Steven J. (1995). Regret Aversion or Event-Splitting Effects? More Evidence under Risk and Uncertainty. *Journal of Risk and Uncertainty* 11(3), pp. 263–274.

Kahneman, Daniel and Tversky, Amos (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2), pp. 263–291.

Karmarkar, Uday S. (1978). Subjectively Weighted Utility: A Descriptive Extension of the Expected Utility Model. *Organizational Behavior and Human Performance* 21(1), pp. 61–72.

Karni, Edi (2007). Foundations of Bayesian theory. *Journal of Economic Theory* 132(1), pp. 167–188.

— (2011a). A theory of Bayesian decision making with action-dependent subjective probabilities. *Economic Theory* 48(1), pp. 125–146.

— (2011b). Subjective Probabilities on a State Space. *American Economic Journal: Microeconomics* 3(4), pp. 172–85.

— (2020). On the indeterminacy of the representation of beliefs by probabilities. *Economics Letters* 196, p. 109535.

Karni, Edi and Schmeidler, David (2016). An expected utility theory for state-dependent preferences. *Theory and Decision* 81(4), pp. 467–478.

Karni, Edi; Schmeidler, David, and Vind, Karl (1983). On State Dependent Preferences and Subjective Probabilities. *Econometrica* 51(4), pp. 1021–1031.

Kovářík, Jaromír; Levin, Dan, and Wang, Tao (2016). Ellsberg paradox: Ambiguity and complexity aversions compared. *Journal of Risk and Uncertainty* 52(1), pp. 47–64.

Krantz, David; Luce, Duncan; Suppes, Patrick, and Tversky, Amos (1971). Foundations of measurement, Vol. I: Additive and polynomial representations.

Le Menestrel, Marc (2001). A Process Approach to the Utility for Gambling. *Theory and Decision* 50(3), pp. 249–262.

Lu, Jay (2019). Bayesian Identification: A Theory for State-Dependent Utilities. *American Economic Review* 109(9), pp. 3192–3228.

Luce, R. Duncan (2001). Reduction Invariance and Prelec's Weighting Functions. *Journal of Mathematical Psychology* 45(1), pp. 167–179.

Luce, R. Duncan and Marley, A. A. J. (2005). Independence Properties Vis-À-Vis Several Utility Representations. *Theory and Decision* 58, pp. 77–143.

Luce, R. Duncan; Ng, C. T., and Marley, A. A. J. (2009). Entropy-Related Measures of the Utility of Gambling. *The Mathematics of Preference, Choice and Order*. Springer, pp. 5–25.

Luce, R. Duncan; Ng, C. T.; Marley, A. A. J., and Aczél, János (2008a). Utility of gambling I: entropy modified linear weighted utility. *Economic Theory* 36(1), pp. 1–33.

Luce, R. Duncan; Ng, C. T.; Marley, A. A. J., and Aczél, János (2008b). Utility of gambling II: risk, paradoxes, and data. *Economic Theory* 36(2), pp. 165–187.

Luce, R. Duncan and Raiffa, Howard (1957). *Games and Decisions*. Wiley.

Maccheroni, Fabio; Marinacci, Massimo, and Rustichini, Aldo (2006). Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica* 74(6), pp. 1447–1498.

Maccheroni, Fabio; Marinacci, Massimo; Rustichini, Aldo, and Taboga, Marco (2009). Portfolio Selection with Monotone Mean-Variance Preferences. *Mathematical Finance* 19(3), pp. 487–521.

Machina, Mark J. (2001). Payoff Kinks in Preferences over Lotteries. *Journal of Risk and Uncertainty* 23(3), pp. 207–260.

Mador, Galit; Sonsino, Doron, and Benzion, Uri (2000). On complexity and lotteries' evaluation – three experimental observations. *Journal of Economic Psychology* 21(6), pp. 625–637.

Meginniss, James R. (1976). *A new class of symmetric utility rules for gambles, subjective marginal probability functions, and a generalized Bayes' rule*. Columbia University, Graduate School of Business.

Moffatt, Peter G.; Sitzia, Stefania, and Zizzo, Daniel John (2015). Heterogeneity in preferences towards complexitys. *Journal of Risk and Uncertainty* 51(2), pp. 147–170.

Mononen, Lasse (2020). On the Price Invariance of Measures of Rationality: The Case of Afriat's Efficiency Index.

Ortoleva, Pietro (2013). The price of flexibility: Towards a theory of Thinking Aversion. *Journal of Economic Theory* 148(3), pp. 903–934.

Prelec, Drazen (1998). The Probability Weighting Function. *Econometrica* 66(3), pp. 497–527.

Puri, Indira (2020). Preference for Simplicity.

Roberts, Kevin W. S. (1980). Interpersonal Comparability and Social Choice Theory. *The Review of Economic Studies* 47(2), pp. 421–439.

Saponara, Nick (2020). Revealed reasoning. *Journal of Economic Theory*.

Savage, Leonard J. (1954). *The Foundations of Statistics*. Wiley publications in statistics. Wiley.

Schmeidler, David (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica* 57(3), pp. 571–587.

Segal, Uzi (1988). Does the Preference Reversal Phenomenon Necessarily Contradict the Independence Axiom? *The American Economic Review* 78(1), pp. 233–236.

— (1990). Two-Stage Lotteries without the Reduction Axiom. *Econometrica* 58(2), pp. 349–377.

Selten, Reinhard (1991). Properties of a measure of predictive success. *Mathematical Social Sciences* 21(2), pp. 153–167.

Sen, Amartya (1986). Social choice theory. Ed. by Michael D. Intriligator and Kenneth J. Arrow. Handbook of Mathematical Economics: Volume III. North Holland, pp. 1073–1181.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal* 27(3), pp. 379–423.

Sims, Christopher A. (2003). Implications of rational inattention. *Journal of Monetary Economics* 50(3). Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information, pp. 665–690.

Siniscalchi, Marciano (2009). Vector Expected Utility and Attitudes Toward Variation. *Econometrica* 77(3), pp. 801–855.

Smeulders, Bart; Spieksma, Frits C. R.; Cherchye, Laurens, and De Rock, Bram (2014). Goodness-of-Fit Measures for Revealed Preference Tests: Complexity Results and Algorithms. 2(1).

Sonsino, Doron; Benzion, Uri, and Mador, Galit (2002). The Complexity Effects on Choice with Uncertainty – Experimental Evidence. *The Economic Journal* 112(482), pp. 936–965.

Starmer, Chris and Sugden, Robert (1993). Testing for Juxtaposition and Event-Splitting Effects. *Journal of Risk and Uncertainty* 6(3), pp. 235–254.

Stodder, James (1997). Complexity Aversion: Simplification in the Herrnstein and Allais Behaviors. *Eastern Economic Journal* 23(1), pp. 1–15.

Strzalecki, Tomasz (2011). Axiomatic Foundations of Multiplier Preferences. *Econometrica* 79(1), pp. 47–73.

Tverberg, Helge (1958). A New Derivation of the Information Function. *Mathematica Scandinavica* 6, pp. 297–298.

Tversky, Amos and Kahneman, Daniel (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4), pp. 297–323.

Vaart, A. W. van der (2000). Asymptotic Statistics.

Valenzuela-Stookey, Quitzé (2020). Subjective Complexity Under Uncertainty.

Varian, Hal R. (1990). Goodness-of-fit in optimizing models. *Journal of Econometrics* 46(1), pp. 125–140.

Viro, O. Ya.; Ivanov, O. A.; Netsvetaev, N. Yu., and Kharlamov, V. M. (2008). *Elementary Topology: Problem Textbook*. American Mathematical Society.

Wakker, Peter P. (1988). The Algebraic versus the Topological Approach to Additive Representations. *Journal of Mathematical Psychology* 32(4), pp. 421–435.

— (2010). *Prospect Theory: For Risk and Ambiguity*. Cambridge University Press.

Weymark, John A. (1991). A Reconsideration of the Harsanyi–Sen Debate on Utilitarianism. *Interpersonal Comparisons of Well-Being*. Ed. by Jon Elster and John E. Roemer. Cambridge University Press, pp. 255–320.

Yang, Jiping and Qiu, Wanhua (2005). A measure of risk and a decision-making model based on expected utility and entropy. *European Journal of Operational Research* 164(3). Recent Advances in Scheduling in Computer and manufacturing Systems, pp. 792–799.