

**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 copyright law. We do not charge any permission or usage fees for the publication of images of material in our 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 on the 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 in Decision Theory: Induced and Revealed Preference

Tang, Rui

Princeton, NJ : Princeton University

---

<http://arks.princeton.edu/ark:/88435/dsp01r781wk11j>

*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 IN DECISION THEORY: INDUCED  
AND REVEALED PREFERENCE

RUI TANG

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  
ADVISER: FARUK GUL

MAY 2021

© Copyright by Rui Tang, 2021.

All rights reserved.

# Abstract

This collection of essays investigates dynamic choices and the revealed preference of indirect choices.

In Chapter 1, I study dynamic choices when new information is ambiguous. I introduce a new updating rule, the conditional maximum likelihood rule (CML) for updating ambiguous information. CML satisfies a new axiom, increased sensitivity after updating, while existing updating rules do not. With CML, a decision maker's posterior is unaffected by the order in which independent signals arrive. CML accommodates recent experimental findings on updating ambiguous information and has applications on learning and information design.

In Chapter 2 (co-authored with Mu Zhang), we propose a new model of menu preference that captures the optimism of decision makers. We model a decision maker who anticipates her preference to change in the future and optimistically evaluates each menu according to the best choice that could possibly be made by her future self. We characterize this menu preference, discuss its uniqueness and propose a comparative measure of optimism. The model connects optimism with the naive quasi-hyperbolic discounting model and has novel predictions on decision makers' choices over menus.

In Chapter 3 (co-authored with Gaoji Hu, Jiangtao Li and John K.-H. Quah), we develop a theory for testing the rationality of decision makers when their indirect preferences over menus are observed. A preference over menus is said to be an indirect preference if it is induced by a preference over the objects that make up those menus. We consider partial information of a decision maker's ranking over certain menus and provide necessary and sufficient conditions on those rankings that guarantee the existence of a preference over objects that induces the observed menu rankings. We then apply this basic result to a wide variety of applications.

# Acknowledgements

I am deeply indebted to my adviser, Faruk Gul, for his continuous guidance, support, and encouragement throughout my doctoral studies. His invaluable advice helps improve the quality of my research and inspires me to pursue excellence in academics. His generosity has always motivated me to be a better person.

I would like to thank Pietro Ortoleva and Wolfgang Pesendorfer for their guidance and insightful suggestions. For many times when I had difficulty making progress on my research, their critical advice helped me find the way out. Their enthusiasm and wisdom for research have shaped my comprehension of economic theory.

I am grateful to John K.-H. Quah—with whom Chapter 3 of this dissertation is co-authored—for his endless help. He taught me how to ask interesting and important economic questions. From him, I also learned how to be more rigorous and meticulous about every detail of research.

I would like to thank Roland Bénabou, Sylvain Chassang, Alessandro Lizzeri, Stephen Morris, Xiaosheng Mu, Can Urgan, Juan Pablo Xandri Antuna and Leeat Yariv, for their insightful feedback on my research and advice on the job market.

I am thankful to Soo Hong Chew, Satoru Takahashi, Junjian Yi and Songfa Zhong for their help and encouragement. I am especially thankful to Satoru Takahashi for his insightful suggestions on many papers of mine. I also would like to thank Gaoji Hu (co-author of Chapter 3 ), Jiangtao Li (co-author of Chapter 3 ) and Mu Zhang (co-author of Chapter 2 ) for their enthusiasm and dedication that have been keeping our projects going.

I would like to thank Ellen Graf and Laura Hedden for providing extraordinary assistance. I thank Princeton University and Stephen Goldfeld Fund for financial support.

Many thanks to my wonderful classmates and friends at Princeton, especially to Zhang Chen, Shushen Hou, Lintong Li, Shumiao Ouyang, Peng Wang, Fangyuan Yi and Haiyue Yu. Their support and company have enriched my years at Princeton.

As always, I want to thank my parents for their unconditional love, support, and encouragement. I am especially grateful to my wife for her support and encouragement. She has broadened my research interests and helped me to improve my writing skills. She is my primary motivation to be better.

To my wife Han and my parents Huifang and Youjun.

# Contents

|                                                                  |          |
|------------------------------------------------------------------|----------|
| Abstract . . . . .                                               | iii      |
| Acknowledgements . . . . .                                       | iv       |
| List of Tables . . . . .                                         | x        |
| List of Figures . . . . .                                        | xi       |
| <b>1 Conditional Maximum Likelihood Updating</b>                 | <b>1</b> |
| 1.1 Introduction . . . . .                                       | 2        |
| 1.2 Model . . . . .                                              | 8        |
| 1.2.1 Evaluation Functions and Extended Evaluation Functions . . | 9        |
| 1.2.2 CML Updating . . . . .                                     | 12       |
| 1.2.3 Axiom ISU . . . . .                                        | 13       |
| 1.3 Characterization . . . . .                                   | 17       |
| 1.3.1 Axiomatic Foundation . . . . .                             | 17       |
| 1.3.2 CML Rationalizability . . . . .                            | 19       |
| 1.4 Application . . . . .                                        | 20       |
| 1.4.1 Divisibility . . . . .                                     | 21       |
| 1.4.2 Information with Ambiguous Accuracy . . . . .              | 22       |
| 1.4.3 Information Design . . . . .                               | 25       |
| 1.5 Discussion . . . . .                                         | 28       |
| 1.5.1 A More General Framework: Violation of Axiom ISU . . . . . | 28       |
| 1.5.2 Comparison with the Proxy Rule . . . . .                   | 30       |

|          |                                                              |           |
|----------|--------------------------------------------------------------|-----------|
| <b>2</b> | <b>Optimism in Choices over Menus</b>                        | <b>34</b> |
| 2.1      | Introduction . . . . .                                       | 35        |
| 2.2      | Model . . . . .                                              | 39        |
| 2.3      | Axiomatization . . . . .                                     | 41        |
| 2.4      | Uniqueness and Comparative Statics . . . . .                 | 46        |
| 2.5      | Application . . . . .                                        | 48        |
| 2.5.1    | Naive Quasi-Hyperbolic Discounting . . . . .                 | 48        |
| 2.5.2    | Disjunction Effect . . . . .                                 | 50        |
| 2.6      | Discussion of Axioms . . . . .                               | 52        |
| <b>3</b> | <b>A Theory of Revealed Indirect Preference</b>              | <b>54</b> |
| 3.1      | Introduction . . . . .                                       | 55        |
| 3.2      | Preliminaries . . . . .                                      | 60        |
| 3.3      | Rationalizability of menu preferences . . . . .              | 62        |
| 3.3.1    | The Never-Covered Property . . . . .                         | 63        |
| 3.3.2    | The Basic Result . . . . .                                   | 66        |
| 3.3.3    | Algorithm . . . . .                                          | 68        |
| 3.3.4    | Nice Rationalization when $\succeq$ is Trivial . . . . .     | 69        |
| 3.3.5    | Strict Menu Preferences and Strict Rationalization . . . . . | 72        |
| 3.3.6    | Menu Preferences in Kreps (1979) and Tyson (2018) . . . . .  | 74        |
| 3.3.7    | Continuous Cationalizability . . . . .                       | 76        |
| 3.4      | Application: Revealed Price Preference . . . . .             | 78        |
| 3.5      | Application: Coarse Rationalizability . . . . .              | 81        |
| 3.5.1    | Four Concepts of Rationalization . . . . .                   | 82        |
| 3.5.2    | Coarse Data Sets and Menu Preferences . . . . .              | 83        |
| 3.5.3    | A Generalization of Afriat's Theorem . . . . .               | 85        |
| 3.6      | Application: Multiple Preferences . . . . .                  | 88        |
| 3.7      | Application: Minimax Regret . . . . .                        | 92        |

|          |                                                                   |            |
|----------|-------------------------------------------------------------------|------------|
| <b>A</b> | <b>Appendices for Chapter 1</b>                                   | <b>97</b>  |
| A.1      | Proof of Proposition 1.2.5 . . . . .                              | 97         |
| A.2      | Proof of Theorem 1.3.1 . . . . .                                  | 98         |
| A.3      | Proof of Theorem 1.3.4 . . . . .                                  | 100        |
| A.4      | Proof of Theorem 1.4.2 . . . . .                                  | 102        |
| <b>B</b> | <b>Appendices for Chapter 2</b>                                   | <b>104</b> |
| B.1      | Proof of Theorem 2.3.1 . . . . .                                  | 104        |
| B.2      | Proof of Theorem 2.4.3 . . . . .                                  | 116        |
| B.3      | Proof of Theorem 2.4.6 . . . . .                                  | 121        |
| B.4      | Proof of Theorem 2.6.1 . . . . .                                  | 122        |
| B.5      | Proof of Theorem 2.6.2 . . . . .                                  | 123        |
| B.6      | Equivalence to C18 with Finite Alternatives . . . . .             | 124        |
| <b>C</b> | <b>Appendices for Chapter 3</b>                                   | <b>126</b> |
| C.1      | Proof of Theorem 3.3.5 . . . . .                                  | 126        |
| C.2      | Proof of Proposition 3.3.6 . . . . .                              | 128        |
| C.3      | Proof of Proposition 3.3.7 . . . . .                              | 129        |
| C.4      | Proof of Theorem 3.3.10 . . . . .                                 | 130        |
| C.5      | Menu Preference Rationalization and Upper Contour Rationalization | 131        |
| C.6      | Proof of Theorem 3.3.15 . . . . .                                 | 132        |
| C.7      | Proof of Theorem 3.4.1 . . . . .                                  | 132        |
| C.8      | Proof of Corollary 3.4.2 . . . . .                                | 134        |
| C.9      | Proof of Theorem 3.5.2 . . . . .                                  | 136        |
| C.10     | Proof of Theorem 3.6.2 . . . . .                                  | 136        |
| C.11     | Proof of Corollary 3.6.3 . . . . .                                | 138        |
| C.12     | Proof of Theorem 3.7.3 . . . . .                                  | 139        |
|          | <b>References</b>                                                 | <b>142</b> |

# List of Tables

|     |                                                                        |    |
|-----|------------------------------------------------------------------------|----|
| 1.1 | Distributions of the priors in Example 1.2.1 . . . . .                 | 13 |
| 1.2 | Distributions of the priors and extended acts in Example 1.2.4 . . . . | 15 |
| 1.3 | Distributions of the priors in Example 1.5.1 . . . . .                 | 29 |
| 1.4 | Distributions of the priors in Example 1.5.2 . . . . .                 | 31 |
| 1.5 | Distributions of the priors in Example 1.5.3 . . . . .                 | 32 |

# List of Figures

|     |                                                                                  |    |
|-----|----------------------------------------------------------------------------------|----|
| 2.1 | Comparison with Different Menu Preference Models . . . . .                       | 53 |
| 3.1 | Data Sets: (a) $\geq$ -Rationalizable and (b) not $\geq$ -Rationalizable . . . . | 66 |
| 3.2 | An Example of Observation Errors . . . . .                                       | 87 |

# Chapter 1

## Conditional Maximum Likelihood Updating

## 1.1 Introduction

In decision theory, the term ambiguity refers either to an event with unknown probability or to information that has multiple probabilistic interpretations. Since the seminal work of Ellsberg (1961), a variety of models have been proposed to rationalize decision makers' (henceforth DM) choices over bets on ambiguous events.<sup>1</sup> In contrast, relatively few papers focus on updating ambiguous information. The growing literature on mechanism and information design under ambiguity has highlighted the importance of the latter topic.

In many applications (e.g., Bose and Renou (2014) and Beauchêne, Li, and Li (2019)), the DM's prior over the payoff-relevant state space is unambiguous but the information he receives is ambiguous. Our focus in this chapter is on such applications. As in most existing models of choice under uncertainty, our DM is a max-min expected utility maximizer. That is, when the DM has a set of priors, he evaluates each ambiguous prospect according to its minimal expected utility over all possible priors. In this chapter, we offer an alternative to full-Bayesian updating (henceforth FB) where the set of posteriors is the set of Bayes' updates of the priors.

One consequence of FB is that a DM's set of posteriors may be a superset of his set of priors. This *dilation* of beliefs may occur even when the DM has a single prior; that is, even if there is no payoff-relevant ambiguity (Wasserman and Kadane, 1990). Shishkin and Ortoleva (2019) (henceforth SO19) offer experimental evidence indicating that ambiguity averse DMs do not dilate after receiving ambiguous information. This finding is inconsistent with FB and other well-known updating rules such as the maximum likelihood rule (henceforth ML). Our new updating rule, CML, does not create dilation and is consistent with experimental evidence presented in recent papers.

---

<sup>1</sup>See, for instance, Gilboa and Schmeidler (1989), Schmeidler (1989), Maccheroni, Marinacci, and Rustichini (2006), Chew and Sagi (2008), Gul and Pesendorfer (2014), etc.

To see how our updating rule works, consider the following example: there is a coin and an urn. The urn contains 100 balls; each ball is either red or blue. It is not known how many balls there are of either color. The coin is tossed, and a ball is drawn from the urn. Consider a bet that yields \$20 to the DM if the coin shows tails and \$0 if the coin shows heads. Before taking the bet, the DM receives a message about the outcome of the toss. The message is either “tails” or “heads”. The message matches the outcome of the coin toss if a red ball is drawn and does not match the outcome if a blue ball is drawn. After receiving the message, the DM is asked if he would be willing to give up the bet in exchange for a objective lottery that yields \$20 with probability  $1/2$  and \$0 with probability  $1/2$ .

Suppose that the DM receives the signal “tails”. Given the signal, the bet depends on the color of the ball that is drawn: if the ball is red, the DM has won; if it is blue, he has lost. Hence, it could be argued that once the signal is observed, the bet becomes ambiguous and therefore, an ambiguity averse DM might strictly prefer the objective lottery to the bet. Full-Bayesian updating is consistent with this view: with FB, an ambiguity averse DM would strictly prefer the objective lottery to the bet after receiving either signal “tails” or signal “heads”.

Since the bet is equivalent to the objective lottery before the signal, the above argument suggests that ambiguity averse DMs lower their value of the bet after observing the signal. SO19, however, find that ambiguity averse DMs typically do not change their value of the bet after observing the signal. Our model provides the following rationale for the behavior of such DMs: if the outcome of the coin toss is tails, then signal “tails” is most likely when the number of red balls in the urn is maximal. Symmetrically, if the outcome of the coin toss is heads, then the signal “tails” is most likely when the number of blue balls in the urn is maximal. Therefore, updating based on the conditionally most likely event, i.e., CML, preserves the initial

symmetry and implies that the DM is indifferent between the bet and the objective lottery even after observing the signal.

Our model has a fundamental state space  $S$  and a signal space  $\Theta$ . We call  $S \times \Theta$  the extended state space. A function mapping from the extended state space to payoffs is an extended act. A function mapping from the state space to payoffs is an act. An extended evaluation function,  $V$ , describes the DM's ex ante preferences over extended acts. An evaluation function,  $V_\theta$ , describes the DM's ex post preferences over acts after observing signal  $\theta$ . Section 1.2.1 provides the necessary formalism and an axiomatic characterization of max-min (extended) evaluation functions.

An updating rule specifies for each extended evaluation function  $V$  and each signal  $\theta$  an evaluation function  $V_\theta$ . CML is defined as follows. Consider an extended evaluation function  $V$ , which admits a max-min representation with a set of priors  $\mathcal{P}$  over  $S \times \Theta$ . Throughout the chapter, we assume that  $\mathcal{P}$  has no ambiguity over  $S$ , i.e., each prior  $p \in \mathcal{P}$  has the same marginal distribution over  $S$ . For any signal  $\theta$ , the evaluation function  $V_\theta$  specified by CML admits a max-min representation with the singleton posterior set  $\{\mu_\theta\}$  over  $S$ , where for each  $s \in S$ ,  $\mu_\theta(s)$  is given by

$$\mu_\theta(s) = \frac{\max_{p \in \mathcal{P}} p(s, \theta)}{\sum_{s' \in S} \max_{p \in \mathcal{P}} p(s', \theta)}.$$

That is, with CML, a DM who has prior set  $\mathcal{P}$  updates his belief over  $S$  to  $\mu_\theta$  after observing  $\theta$ . The maximization in the formula captures the conditional maximum likelihood of the signal on each state.

In Section 1.2.3, we introduce a new axiom, *increased sensitivity after updating* (henceforth ISU), which distinguishes CML from existing updating rules. The idea of the axiom can be illustrated through the following example. Consider a state space containing three states  $\{s_1, s_2, s_3\}$  with prior  $(1/3, 1/3, 1/3)$ . Let event  $\{s_1, s_2\}$  be realized. The Bayes' posterior is given by  $(1/2, 1/2)$ . Assume that the DM is an

expected utility maximizer. If his payoff at state  $s_1$  increases by  $\epsilon > 0$ , his ex ante expected payoff increases by  $\epsilon/3$ , and his ex post expected payoff increases by  $\epsilon/2$ . Clearly, he is more sensitive to payoff changes on  $s_1$  after event  $\{s_1, s_2\}$  is realized. Axiom ISU is motivated by this observation.

We state axiom ISU in our framework. Consider two extended acts  $f^*$  and  $g^*$ . Suppose that  $g^*(s, \theta) > f^*(s, \theta)$  for some  $(s, \theta) \in S \times \Theta$ , and  $g^*(s', \theta') = f^*(s', \theta')$  for all  $(s', \theta')$  different from  $(s, \theta)$ . Let  $f$  and  $g$  be two acts satisfying  $f(\hat{s}) = f^*(\theta, \hat{s})$  and  $g(\hat{s}) = g^*(\theta, \hat{s})$  for each  $\hat{s} \in S$ . An updating rule satisfies ISU if for any extended evaluation function  $V$  and any signal  $\theta$ ,

$$V(f^*) = V_\theta(f) \text{ implies } V(g^*) \leq V_\theta(g),$$

where  $V_\theta$  is the evaluation function specified by the updating rule. We interpret this axiom as follows. If the extended act changes from  $f^*$  to  $g^*$ , the DM has a payoff increase at  $(s, \theta)$ . The payoff increase should affect the DM more after  $\theta$  is observed, since it rules out event  $S \times (\Theta \setminus \{\theta\})$  and increases the chance that  $(s, \theta)$  occurs. Therefore, the difference between the evaluations of  $f^*$  and  $g^*$  at the ex post stage should be higher than that at the ex ante stage. Since  $f$  has the same value as  $f^*$ ,  $g$  must have a higher value than  $g^*$ . The condition  $V(f^*) = V_\theta(f)$  further rules out the possibility that different sensitivities to payoff changes are caused by different ex ante and ex post utility levels. Axiom ISU is satisfied by CML but violated by other updating rules.

In Section 1.3, we characterize CML. We show that an updating rule is CML if and only if it satisfies axiom ISU, axiom independence of irrelevant signals, and axiom mixture independence (Theorem 1.3.1). The latter two axioms are also satisfied by FB and ML. When the DM's extended evaluation function is not observed, we provide a characterization of CML rationalizable belief profiles. Specifically, a belief profile

contains the DM’s ex ante belief over  $S$  and ex post belief over  $S$  after observing each signal. We provide a necessary and sufficient condition for the existence of an information structure under which the belief profile is consistent with CML updating.

In Section 1.4, we consider three applications of CML. First, we show that CML is divisible. That is, the DM’s posterior is not affected by the order in which independent signals arrive. This makes CML suitable for the analysis of Wald-type problems under ambiguity. Second, we relate CML to experimental evidence involving signals of unknown accuracy. We show that a DM who updates according to CML under-reacts to ambiguous information, as observed by Liang (2021) (henceforth L19). We also investigate the learning behavior of a DM who updates signals of unknown accuracy with CML.

Finally, we apply CML to a Kamenica and Gentzkow (2011)-style information design problem. We consider a designer who wants to persuade an agent to take an action that affects the designer’s payoff. The designer can choose any ambiguous or unambiguous information structure. We show that the designer can almost achieve her maximal payoff when facing a CML agent. That is, for any  $\epsilon > 0$ , there exists an ambiguous information structure such that the designer’s payoff is no less than her maximal payoff minus  $\epsilon$ .

Section 1.5 contains our comparisons of CML with existing alternatives. We show that within the max-min expected utility framework, no updating rule satisfies axiom ISU if we allow multiple priors over the payoff-relevant state space. Hence, we show that the ISU assumption is appropriate only when ambiguity arises from information.

**Related Literature.** A couple of experimental studies directly test how subjects react to ambiguous information, e.g., Epstein, Halevy, et al. (2019), Kellner, Le Quement, and Riener (2019), etc. Our study is mostly related to the two experimental papers: SO19 and L19. In both papers, the state space is finite. Both papers test

how DMs react to ambiguous information when there is no ambiguity over the state space. Our theory can partially accommodate their findings.

Another stream of literature studies ambiguous information in game-theoretical frameworks, including Blume and Board (2014), Bose and Renou (2014), Kellner and Le Quement (2017), Beauchêne, Li, and Li (2019), Kellner and Le Quement (2018), etc. In this chapter, we also apply CML to study the information design problem. We characterize the set of payoffs that can be achieved by an information designer when the agent updates according to CML.

This chapter contributes to the literature of updating under ambiguity. The two most studied updating rules under ambiguity are FB and ML. FB is analyzed by Wasserman and Kadane (1990) and Jaffray (1992) and axiomatized by Pires (2002). ML is introduced by Dempster (1967) and Shafer (1976) and axiomatized by Gilboa and Schmeidler (1993) and Cheng (2021). Cheng (2021) also provides a new updating rule, the relative maximum likelihood rule, under which a DM's posterior set is a convex combination of FB and ML posterior sets. CML differs from these rules in two respects. First, CML satisfies axiom ISU. Second, with CML, the DM's prior over the state space is always updated to a single posterior and thus is not dilated.

Hanany and Klibanoff (2007) introduce the dynamic consistent updating rule, featuring a DM who only updates a subset of his priors using Bayes' rule to maintain the optimality of the ex ante optimal choice. The posterior set of a DM who updates according to Hanany and Klibanoff (2007) depends on the ex ante act as well as the choice menu, which violates consequentialism.<sup>2</sup> CML differs from this rule in two respects. First, CML satisfies consequentialism. Second, the CML posterior of a DM may not be contained in his FB posterior set (Example 1).

A new updating rule, named the proxy rule, is proposed and axiomatized by Gul and Pesendorfer (forthcoming). Gul and Pesendorfer (forthcoming) introduce

---

<sup>2</sup>Consequentialism says that the ex post evaluation of a choice does not depend on its payoffs on states not contained in the realized event.

the axiom “not all news is bad news” and show that the axiom is satisfied by the proxy rule but not FB or ML. A common feature of the proxy rule and CML is that information does not render unambiguous events ambiguous. The key difference between CML and the proxy rule is that CML does not satisfy “not all news is bad news” while the proxy rule does not satisfy ISU.<sup>3</sup>

The rest of the chapter is organized as follows. We present the model in Section 1.2. We characterize CML in Section 1.3. Section 1.4 contains applications of CML. We discuss CML in Section 1.5. All omitted proofs in this chapter are in Appendix A.

## 1.2 Model

For an arbitrarily nonempty set  $H$ , let  $\Delta(H)$  denote the set of finitely supported probability distributions over  $H$ , i.e.,  $d \in \Delta(H)$  if and only if there exists a nonempty and finite subset  $H' \subseteq H$  such that  $d(H') = 1$ . If  $H$  is finite, let  $\Delta^o(H)$  be the relative interior of  $\Delta(H)$ , i.e.,  $d \in \Delta^o(H)$  if and only if  $d(H') > 0$  for each nonempty  $H' \subseteq H$ . Any probability distribution over  $H' \subseteq H$  is viewed as a probability distribution over  $H$  that has support  $H'$ , and any probability distribution over  $H$  that has support  $H'$  is also viewed as a probability distribution over  $H'$ . For any two sets of probability distributions  $\mathcal{D}$  and  $\mathcal{D}'$  over  $H$  and any  $\alpha \in [0, 1]$ , let  $\alpha\mathcal{D} + (1 - \alpha)\mathcal{D}'$  be the  $\alpha$ -convex combination of the two sets:  $\alpha\mathcal{D} + (1 - \alpha)\mathcal{D}' = \{\alpha d + (1 - \alpha)d' : d \in \mathcal{D}, d' \in \mathcal{D}'\}$ .

For any function  $f : H \rightarrow \mathbb{R}$  and any  $d \in \Delta(H)$ , define  $\mathbb{E}_d(f) = \sum_{h \in H} d(h)f(h)$  as the  $d$ -evaluation of  $f$ . For any nonempty  $\mathcal{D} \subseteq \Delta(H)$ , define  $\mathbb{E}_{\mathcal{D}}(f) = \inf_{d \in \mathcal{D}} \mathbb{E}_d(f)$  as the  $\mathcal{D}$ -evaluation of  $f$ . For any two functions  $f$  and  $g$ , we write  $f \geq g$  if  $f(h) \geq g(h)$  for all  $h \in H$ . For any subset  $H' \subseteq H$ , we write  $f =^{H'} g$  if  $f(h) = g(h)$  for all  $h \in H'$ . We let  $fH'g$  denote the function that agrees with  $f$  on  $H'$  and agrees with  $g$  on  $H \setminus H'$ . For any  $\alpha \in [0, 1]$ ,  $\alpha f + (1 - \alpha)g$  denotes the function satisfying

---

<sup>3</sup>See Section 1.5.2 for a detailed discussion.

that  $(\alpha f + (1 - \alpha)g)(h) = \alpha f(h) + (1 - \alpha)g(h)$  for all  $h \in H$ . For any partition  $\Pi = \{H^1, \dots, H^n\}$  of  $H$ ,  $f$  is said to be measurable with respect to  $\Pi$  if  $f(h) = f(h')$  whenever  $h$  and  $h'$  are in the same block of the partition.

Let  $H = H_1 \times H_2$ . For any  $d \in \Delta(H)$  and any  $h_2 \in H_2$ , if  $d(h_1, h_2) > 0$  for some  $h_1 \in H_1$ , then  $d|h_2 \in \Delta(H_1)$  denotes the conditional probability of  $d$  on  $h_2$ :  $d|h_2(h'_1) = d(h'_1, h_2) / (\sum_{\hat{h}_1 \in H_1} d(\hat{h}_1, h_2))$  for all  $h'_1 \in H_1$ . For any  $\mathcal{D} \subseteq \Delta(H)$  and any  $h_2 \in H_2$ , define the set of conditional distributions of  $\mathcal{D}$  on  $h_2$  as  $\mathcal{D}|h_2 = \{d|h_2 : d \in \mathcal{D}, \exists h_1 \in H_1 \text{ s.t. } d(h_1, h_2) > 0\}$ . For any function  $f : H \rightarrow \mathbb{R}$ , let  $f|h_2$  denote the function mapping from  $H_1$  to  $\mathbb{R}$  satisfying that  $f|h_2(h_1) = f(h_1, h_2), \forall h_1 \in H_1$ . For convenience, we write  $h$  for the singleton set  $\{h\}$  throughout the chapter when there is no confusion.

### 1.2.1 Evaluation Functions and Extended Evaluation Functions

In this section, we introduce evaluation functions and extended evaluation functions, both of which are assumed to have max-min representations. Each updating rule is defined as a function that maps extended evaluation functions and signals to evaluation functions.

Let  $S$  be the state space. We assume that  $S$  contains at least three states and is finite. Let  $\Theta$  be the set of signals.  $\Theta$  is countably infinite. Let  $\mathbb{K}$  be the payoff space, which is an interval of  $\mathbb{R}$  containing a nonempty interior. An act is a function  $f : S \rightarrow \mathbb{K}$ , and an extended act is a function  $f^* : S \times \Theta \rightarrow \mathbb{K}$ . We use  $x \in \mathbb{K}$  to denote the constant act as well as the constant extended act that equals  $x$  everywhere. Let  $\mathcal{F}$  denote the set of all acts and  $\mathcal{F}^*$  the set of all extended acts.

The DM's priors over  $S \times \Theta$  can be identified through his choices over extended acts. The DM's choices over extended acts are characterized by an extended evalua-

tion function. An extended evaluation function is a map  $V : \mathcal{F}^* \rightarrow \mathbb{K}$  satisfying that for any  $x \in \mathbb{K}$  and any  $f^*, g^*, h^* \in \mathcal{F}^*$ :

- (i) **(Identity)**:  $V(x) = x$ .
- (ii) **(Continuity)**: the sets  $\{\alpha \in [0, 1] : V(\alpha f^* + (1 - \alpha)g^*) \geq V(h^*)\}$  and  $\{\alpha \in [0, 1] : V(\alpha f^* + (1 - \alpha)g^*) \leq V(h^*)\}$  are closed.
- (iii) **(Certainty Independence)**:  $\forall \alpha \in (0, 1)$ ,  $V(f^*) > V(g^*)$  if and only if  $V(\alpha f^* + (1 - \alpha)x) > V(\alpha g^* + (1 - \alpha)x)$ .
- (iv) **(Monotonicity)**:  $f^* \geq g^*$  implies  $V(f^*) \geq V(g^*)$ .
- (v) **(Uncertainty Aversion)**:  $V(f^*) = V(g^*)$  implies  $V(\frac{1}{2}f^* + \frac{1}{2}g^*) \geq V(f^*)$ .
- (vi) **(Finite Support)**: There exists a nonempty and finite subset  $\Theta' \subseteq \Theta$  such that  $V(f^*) = V(g^*)$  whenever  $f^* =^{S \times \Theta'} g^*$ .
- (vii) **(No Ambiguity over State Space)**: If  $f^*$  and  $g^*$  are measurable with respect to the partition  $\{s \times \Theta\}_{s \in S}$  and  $V(f^*) = V(g^*)$ , then  $V(\frac{1}{2}f^* + \frac{1}{2}g^*) = V(f^*)$ .

Let  $\mathcal{V}$  be the set of all extended evaluation functions. Conditions (i)-(v) are standard for a max-min expected utility representation.<sup>4</sup> Together with condition (vi), there exists a unique nonempty, convex and closed set  $\mathcal{P} \subseteq \Delta(S \times \Theta')$  such that  $V(f^*) = \mathbb{E}_{\mathcal{P}}(f^*)$  for each extended act  $f^*$ .  $\mathcal{P}$  is said to represent  $V$ . Conditions (vii) ensures that each prior  $p$  in  $\mathcal{P}$  has the same marginal distribution over  $S$ , i.e., for each  $p, p' \in \mathcal{P}$  and each  $s \in S$ ,  $p(s \times \Theta) = p'(s \times \Theta)$ . Such a set of priors over  $S \times \Theta$  is said to be simple. We discuss more general prior sets in Section 1.5.1.

For any  $\mathcal{P}$  that is simple, we can decompose it to a prior  $\mu$  over  $S$  and a set of conditional probabilities  $\{c^t(\cdot|s)_{s \in S}\}_{t \in T} \subseteq (\Delta(\Theta))^S$  satisfying that:

- (i)  $\mu(s) = p(s \times \Theta)$  for each  $s \in S$  and each  $p \in \mathcal{P}$ ,
  - (ii)  $\forall t \in T$ , there exists  $p \in \mathcal{P}$  such that  $\mu(s)c^t(\theta|s) = p(s, \theta)$ ,  $\forall s \in S$  and  $\forall \theta \in \Theta$ ,
- and

---

<sup>4</sup>See Gilboa and Schmeidler (1989) for more details.

(iii)  $\forall p \in \mathcal{P}$ , there exists  $t \in T$  such that  $p(s, \theta) = \mu(s)c^t(\theta|s)$ ,  $\forall s \in S$  and  $\forall \theta \in \Theta$ .

We write  $\mathcal{P} = (\mu, \{c^t(\cdot|s)_{s \in S}\}_{t \in T})$  if the above conditions hold. Based on the reformulation, a DM who has a simple prior set has no ambiguity over the state space but may have multiple probabilistic interpretations over the signals, where each  $t \in T$  denotes one possible interpretation.

After observing some signal  $\theta$ , the DM's ex post beliefs can be revealed from his choices over acts. The DM's choices over acts are described by an evaluation function. An evaluation function is a map  $U : \mathcal{F} \rightarrow \mathbb{K}$  satisfying identify, continuity, certainty independence, monotonicity and uncertainty aversion. These conditions ensure that  $U$  admits a max-min representation by a unique nonempty, convex and closed set of probability distributions  $\mathcal{Q} \subseteq \Delta(S)$ . When  $\mathcal{Q} = \{\mu^*\}$ , we say that  $U$  is represented by  $\mu^*$ .  $\mathcal{Q}$  is the DM's set of beliefs over the state space after observing signal  $\theta$ . Let  $\mathcal{U}$  be the set of all evaluation functions.

We proceed to define updating rules. We only consider signals that happen with non-zero probabilities.<sup>5</sup> For this purpose, we define non-null signals for an arbitrary extended evaluation function  $V$ .  $(s, \theta) \in S \times \Theta$  is said to be  $V$  null if  $V(f^*) = V(g^*)$  for any  $f^*, g^* \in \mathcal{F}^*$  satisfying  $f^* =_{(S \times \Theta) \setminus (s, \theta)} g^*$ . If  $(s, \theta)$  is not  $V$  null, it is  $V$  non-null. It can be easily shown that if  $V$  is represented by  $\mathcal{P}$ ,  $(s, \theta)$  is  $V$  null if and only if for all  $p \in \mathcal{P}$ ,  $p(s, \theta) = 0$ . A signal  $\theta$  is said to be  $V$  non-null if there exists  $s \in S$  such that  $(s, \theta)$  is  $V$  non-null. The set of all  $V$  non-null signals is denoted as  $\Theta_V$ .  $V$  non-null states are defined similarly. In the examples and applications, we omit states and signals that are not non-null.

Define  $\mathcal{G} \subseteq \mathcal{V} \times \Theta$  such that  $(V, \theta) \in \mathcal{G}$  if and only if  $\theta \in \Theta_V$ . That is, a tuple  $(V, \theta)$  is contained in  $\mathcal{G}$  if and only if  $\theta$  is  $V$  non-null. An updating rule is a function

$$\Gamma : \mathcal{G} \rightarrow \mathcal{U}.$$

---

<sup>5</sup>We do not model how DMs react to unexpected information in this chapter. For theories of updating events with zero probability, see, for example, Ortoleva (2012).

For simplicity, for any  $V \in \mathcal{V}$  and  $\theta \in \Theta_V$ , we write  $V_\theta$  for  $\Gamma(V, \theta)$ . To interpret, the extended evaluation function  $V$  characterizes the DM's prior beliefs over the produce of the state and signal spaces. Given  $V$ , an updating rule maps each observed signal to the DM's ex post evaluation function over acts, which reflects the DM's ex post beliefs over the state space after observing the signal.

### 1.2.2 CML Updating

We introduce CML in this section. Let the DM's prior set over  $S \times \Theta$  be a simple set  $\mathcal{P}$ . Consider a signal  $\theta$  satisfying that  $p(s, \theta) > 0$  for some  $p \in \mathcal{P}$  and  $s \in S$ . When signal  $\theta$  is observed, CML updating leads to a singleton posterior set  $\{\mu_\theta\} \subseteq \Delta(S)$ , where for each  $s \in S$

$$\mu_\theta(s) = \frac{\max_{p \in \mathcal{P}} p(s, \theta)}{\sum_{s' \in S} \max_{p \in \mathcal{P}} p(s', \theta)}. \quad (1.1)$$

Equivalently, if  $\mathcal{P} = (\mu, \{c^t(\cdot|s)_{s \in S}\}_{t \in T})$ , then for each  $s \in S$ , we have

$$\mu_\theta(s) = \frac{\mu(s) \max_{t \in T} c^t(\theta|s)}{\sum_{s' \in S} \mu(s') \max_{t \in T} c^t(\theta|s')}. \quad (1.2)$$

We interpret the above formulation as the DM uses the signal's maximal conditional probability on each state to update his belief. The DM's CML posterior set is always a singleton and thus is not dilated. With CML, information does render unambiguous events ambiguous.

We note that the CML posterior may not be the Bayes' posterior of any one of the DM's priors over  $S \times \Theta$ , i.e.,  $\mu_\theta$  may not be contained in  $\mathcal{P}|\theta$ . Example 1.2.1 illustrates this point.

**Example 1.2.1** Let  $S = \{s, s', s''\}$  and  $\Theta = \{\theta, \theta'\}$ . The DM's prior set over  $S \times \Theta$  is a simple set  $\mathcal{P} = \{\alpha p^1 + (1 - \alpha)p^2 : \alpha \in [0, 1]\}$ , where  $p^1$  and  $p^2$  are shown in Table 1.1.

Table 1.1: Distributions of the priors in Example 1.2.1

| $p^1$ | $\theta$ | $\theta'$ | $p^2$ | $\theta$ | $\theta'$ |
|-------|----------|-----------|-------|----------|-----------|
| $s$   | 4/15     | 1/15      | $s$   | 4/15     | 1/15      |
| $s'$  | 4/15     | 1/15      | $s'$  | 1/15     | 4/15      |
| $s''$ | 1/15     | 4/15      | $s''$ | 4/15     | 1/15      |

Since the maximal probabilities of  $(s, \theta)$ ,  $(s', \theta)$  and  $(s'', \theta)$  are all equal to 4/15, the DM's CML posterior is  $(1/3, 1/3, 1/3)$  when signal  $\theta$  is observed. Note that  $\mathcal{P}|\theta$  consists of all convex combinations of  $p^1|\theta$  and  $p^2|\theta$ , where  $p^1|\theta = (4/9, 4/9, 1/9)$  and  $p^2|\theta = (4/9, 1/9, 4/9)$ . Obviously,  $(1/3, 1/3, 1/3)$  is not a convex combination of  $p^1|\theta$  and  $p^2|\theta$ . That is,  $(1/3, 1/3, 1/3)$  is not contained in  $\mathcal{P}|\theta$ . Since the DM updates his belief based on the conditionally most likely scenario for each state, correlations among different states' conditional probabilities are neglected. Hence, the DM's posterior can be outside of  $\mathcal{P}|\theta$ .

CML is formally defined as follows.

**Definition 1.2.2** *An updating rule is CML if for any  $V \in \mathcal{V}$  represented by  $\mathcal{P}$  and any  $\theta \in \Theta_V$ , the evaluation function  $V_\theta$  specified by the updating rule is represented by the CML posterior  $\mu_\theta$  given by condition (1.1).*

### 1.2.3 Axiom ISU

In this section, we introduce the axiom ISU, which distinguishes CML from existing updating rules. To start with, we introduce two existing updating rules: FB and ML. Let  $V$  be the DM's extended evaluation function, which is represented by  $\mathcal{P}$ . Let a  $V$  non-null signal  $\theta$  be observed.

(**FB.**) A DM who follows FB updates  $\mathcal{P}$  to the posterior set  $\mathcal{P}|\theta$ . That is, given signal  $\theta$ , he updates each prior  $p$  in  $\mathcal{P}$  following Bayes' rule. With FB, the DM's ex post evaluation function  $V_\theta$  is represented by the closure of  $\mathcal{P}|\theta$ .

(**ML.**) A DM who follows ML first selects priors  $\mathcal{P}^{\theta,M}$  from  $\mathcal{P}$  that maximize the probability of  $S \times \theta$  :

$$\mathcal{P}^{\theta,M} = \{p \in \mathcal{P} : p(S \times \theta) \geq p'(S \times \theta), \forall p' \in \mathcal{P}\}.$$

Then, he updates each prior in  $\mathcal{P}^{\theta,M}$  following Bayes' rule, which leads to the posterior set  $\mathcal{P}^{\theta,M}|\theta$ . With ML, the DM's ex post evaluation function  $V_\theta$  is represented by  $\mathcal{P}^{\theta,M}|\theta$ .

We motivate axiom ISU in an unambiguous environment, in which the DM updates according to Bayes' rule and is an expected utility maximizer. Let  $S = \{s, s'\}$  and  $\Theta = \{\theta, \theta'\}$ . The DM has a single prior  $p$  over  $S \times \Theta$  satisfying that  $p(s, \theta) = p(s', \theta) = p(s, \theta') = p(s', \theta') = 1/4$ . Consider an arbitrary extended act  $f^*$ . If the payoff of  $f^*$  at  $(s, \theta)$  increases by  $\epsilon > 0$ , its evaluation increases by  $\epsilon/4$ . When  $\theta$  is observed, the DM's posterior over  $\{s, s'\}$  is  $(1/2, 1/2)$ . Hence, the DM's ex post evaluation of the act  $f^*|\theta$  increases by  $\epsilon/2$  if the payoff of  $f^*$  at  $(s, \theta)$  increases by  $\epsilon$ .

The key observation from the above example is that the DM becomes more sensitive to payoff changes at  $(s, \theta)$  after signal  $\theta$  is observed, since the realization of  $\theta$  rules out event  $S \times (\Theta \setminus \theta)$ , which makes payoff changes at  $(s, \theta)$  more likely to occur. This argument does not rely on whether the information is ambiguous or not. Based on this observation, we introduce axiom ISU.

**Definition 1.2.3** *An updating rule satisfies axiom ISU if for any extended evaluation function  $V$ , any  $\theta \in \Theta_V$  and any  $f^*, g^* \in \mathcal{F}^*$  with  $g^*(s, \theta) > f^*(s, \theta)$  and*

$$g^* =_{(S \times \Theta) \setminus (s, \theta)} f^*,$$

$$V(f^*) = V_\theta(f^*|\theta) \text{ implies } V(g^*) \leq V_\theta(g^*|\theta).$$

To understand the definition of axiom ISU, note that  $V(f^*) = V_\theta(f^*|\theta)$ . Since the increase of payoff at  $(s, \theta)$  affects the DM more after  $\theta$  is observed, we have  $V(g^*) \leq V_\theta(g^*|\theta)$ . However, the next example shows that axiom ISU is violated by both FB and ML.

**Example 1.2.4** Let  $S = \{s, s'\}$  and  $\Theta = \{\theta, \theta'\}$ . The DM's prior set over  $S \times \Theta$  is a simple set  $\mathcal{P} = \{\alpha p^1 + (1 - \alpha)p^2 : \alpha \in [0, 1]\}$ , where  $p^1$  and  $p^2$  are shown in Table 1.2.

Table 1.2: Distributions of the priors and extended acts in Example 1.2.4

| $p^1$ | $\theta$ | $\theta'$ | $p^2$ | $\theta$ | $\theta'$ | $f^*$ | $\theta$ | $\theta'$ | $g^*$ | $\theta$ | $\theta'$ |
|-------|----------|-----------|-------|----------|-----------|-------|----------|-----------|-------|----------|-----------|
| $s$   | 1/8      | 3/8       | $s$   | 1/3      | 1/6       | $s$   | 3        | 6         | $s$   | 15/4     | 6         |
| $s'$  | 3/8      | 1/8       | $s'$  | 1/8      | 3/8       | $s'$  | 2        | 0         | $s'$  | 2        | 0         |

Let  $V$  be the DM's extended evaluation function, represented by  $\mathcal{P}$ . Consider extended acts  $f^*$  and  $g^*$ , as shown in Table 1.2. We can verify that

$$V(f^*) = \mathbb{E}_{\{p^1, p^2\}}(f^*) = \mathbb{E}_{p^2}(f^*) = 9/4,$$

and

$$V(g^*) = \mathbb{E}_{\{p^1, p^2\}}(g^*) = \mathbb{E}_{p^2}(g^*) = 5/2.$$

Let signal  $\theta$  be observed.

First, assume that the DM updates according to FB. The DM's posterior set over  $S$  consists of all convex combinations of  $p^1|\theta = (1/4, 3/4)$  and  $p^2|\theta = (8/11, 3/11)$ . Hence, a DM who follows FB evaluates  $f^*|\theta$  as  $\mathbb{E}_{\{p^1|\theta, p^2|\theta\}}(f^*|\theta)$  and evaluates  $g^*|\theta$  as

$\mathbb{E}_{\{p^1|\theta, p^2|\theta\}}(g^*|\theta)$ . We have

$$\mathbb{E}_{\{p^1|\theta, p^2|\theta\}}(f^*|\theta) = \mathbb{E}_{p^1|\theta}(f^*|\theta) = 9/4,$$

and

$$\mathbb{E}_{\{p^1|\theta, p^2|\theta\}}(g^*|\theta) = \mathbb{E}_{p^1|\theta}(g^*|\theta) = 39/16.$$

Next, assume that the DM follows ML. The DM selects the prior that maximizes the probability of  $S \times \theta$ , which is  $p^1$ . Thus, his posterior set is  $\{p^1|\theta\}$ . Again, his ex post evaluations of  $f^*|\theta$  and  $g^*|\theta$  are given by

$$\mathbb{E}_{p^1|\theta}(f^*|\theta) = 9/4 \text{ and } \mathbb{E}_{p^1|\theta}(g^*|\theta) = 39/16.$$

Since  $\mathbb{E}_{p^2}(f^*) = \mathbb{E}_{p^1|\theta}(f^*|\theta)$  and  $\mathbb{E}_{p^2}(g^*) > \mathbb{E}_{p^1|\theta}(g^*|\theta)$ , both ML and FB violate axiom ISU. In contrast, CML satisfies axiom ISU, which is implied by the following proposition.

**Proposition 1.2.5** *Let  $V \in \mathcal{V}$ ,  $\theta \in \Theta_V$  and  $V_\theta = \Gamma(V, \theta)$ , where  $\Gamma$  is CML. For any extended acts  $f^*$  and  $g^*$  and any  $(s, \theta) \in S \times \Theta$ , if  $g^*(s, \theta) > f^*(s, \theta)$  and  $g^* =_{(S \times \Theta) \setminus (s, \theta)} f^*$ , then*

$$V(g^*) - V(f^*) \leq V_\theta(g^*|\theta) - V_\theta(f^*|\theta).$$

Proposition 1.2.5 says that CML satisfies a stronger version of axiom ISU. If a DM updates according to CML, he is always more sensitive to payoff changes at  $(s, \theta)$  after  $\theta$  is observed, no matter whether his ex ante evaluation equals his ex post evaluation or not.

## 1.3 Characterization

In this section, we provide an axiomatic foundation for CML. We provide three axioms to fully characterize CML, two of which are also satisfied by both FB and ML. The only axiom that is violated by FB and ML is axiom ISU. After that, we consider the situation in which a DM's extended evaluation function is not observable. Instead, we observe the DM's prior belief  $\mu$  over the state space and posterior belief after observing each signal. We provide a sufficient and necessary condition under which the DM's belief profile can be rationalized by CML, i.e., there exists a simple set of priors over  $S \times \Theta$  that has marginal distribution  $\mu$  over  $S$  such that CML updating leads to the desired posterior of the DM under each signal.

### 1.3.1 Axiomatic Foundation

For any two extended evaluation functions  $V$  and  $V'$  and any  $E \subseteq S \times \Theta$ ,  $V$  and  $V'$  are said to agree on  $E$ , denoted by  $V \approx^E V'$ , if for any extended act  $f^*$  and any payoff  $x$ ,  $V(f^*Ex) = V'(f^*Ex)$ . To interpret, consider two DMs who have extended evaluation functions  $V$  and  $V'$  respectively. If  $V$  and  $V'$  agree on  $E$ , then the two DMs agree on how to trade off payoffs on  $E$ . Therefore, when the payoff of an extended act on the complement of  $E$  is constant, the two DMs have the same evaluation of the extended act.

**Independence of Irrelevant Signals (Axiom IIS).** For any  $\{V, V'\} \subseteq \mathcal{V}$  and any  $\theta \in \Theta_V$ , if  $V \approx^{S \times \theta} V'$ , then  $V_\theta = V'_\theta$ .

Axiom IIS says that if two DMs agree on how to trade off payoffs on  $S \times \theta$  ex ante, then when  $\theta$  is realized, the two DMs have the same ex post evaluation function. Equivalently, when  $V$  is represented by  $\mathcal{P}$  and  $V'$  by  $\mathcal{P}'$ , axiom IIS says that if  $\mathcal{P}$  and  $\mathcal{P}'$  induce the same set of probability distributions over the algebra generated by all subsets of  $S \times \theta$ , then the posterior beliefs of the two DMs are the same after  $\theta$

is observed. That is, if the two DMs agree on how probabilities are distributed over  $S \times \theta$  ex ante, then they have the same posterior beliefs over  $S$  given  $\theta$ . Clearly, this axiom is satisfied by FB and ML.

For any  $\alpha \in [0, 1]$  and any two extended evaluation functions  $V$  and  $V'$ , we define the  $\alpha$ -mixture of  $V$  and  $V'$  as  $\alpha V + (1 - \alpha)V'$  such that for all  $f^* \in \mathcal{F}^*$ ,

$$\alpha V + (1 - \alpha)V'(f^*) = \alpha V(f^*) + (1 - \alpha)V'(f^*).$$

Note that  $\alpha V + (1 - \alpha)V'$  is also an extended evaluation function. With this notation, we state our second axiom.

**Mixture Independence (Axiom MI).** For any  $\alpha \in (0, 1]$ , any  $\{V, V'\} \subseteq \mathcal{V}$  and any  $\theta \in \Theta_V \setminus \Theta_{V'}$ ,  $V_\theta = (\alpha V + (1 - \alpha)V')_\theta$ .

Consider two DMs whose extended evaluation functions are  $V$  and  $\alpha V + (1 - \alpha)V'$  separately where  $V, V'$  and  $\theta$  satisfy the primitive condition of axiom MI. Let  $V$  be represented by  $\mathcal{P}$  and  $\alpha V + (1 - \alpha)V'$  by  $\mathcal{P}^\alpha$ . The probabilities of each event in  $S \times \theta$  assigned by  $\mathcal{P}^\alpha$  shrink uniformly by a factor of  $\alpha$  compared to the probabilities assigned by  $\mathcal{P}$ . Axiom MI says that the ratio  $\alpha$  does not affect the DM's posterior beliefs as long as the signal  $\theta$  is observed. Therefore, the two DMs have the same ex post evaluation function given  $\theta$ . FB and ML also satisfy this axiom. The last axiom is axiom ISU.

**Increased Sensitivity after Updating (Axiom ISU).** For any  $V \in \mathcal{V}$ , any  $\theta \in \Theta_V$  and any  $\{f^*, g^*\} \subseteq \mathcal{F}^*$  satisfying  $f^*(s, \theta) < g^*(s, \theta)$  and  $f^* =^{(S \times \Theta) \setminus (s, \theta)} g^*$ ,  $V(f^*) = V_\theta(f^*|\theta)$  implies  $V(g^*) \leq V_\theta(g^*|\theta)$ .

The three axioms fully characterize CML.

**Theorem 1.3.1** *An updating rule is CML if and only if it satisfies axioms IIS, MI and ISU.*

### 1.3.2 CML Rationalizability

In this section, we study the situation in which the DM's extended evaluation function cannot be observed. Let the DM's ex ante evaluation function be  $U$ .  $U$  describes the DM's choices over acts before any information. Let  $U$  be represented by  $\mu \in \Delta(S)$ . We assume throughout this section that  $\mu \in \Delta^o(S)$ .<sup>6</sup> The DM's choices over acts under signal  $\theta$  is characterized by the ex post evaluation function  $U_\theta$ . Let  $\{U_\theta\}_{\theta \in \Theta^*}$  be the DM's set of ex post evaluation functions, where  $\Theta^*$  is a finite set of signals.  $U_\theta$  is represented by  $\mu_\theta \in \Delta(S)$ .<sup>7</sup> Given the profile  $(\mu, \{\mu_\theta\}_{\theta \in \Theta^*})$ , we want to provide a sufficient and necessary condition under which the ex post beliefs of the DM are updated according to CML under some set of interpretations of signals.

**Definition 1.3.2**  $(\mu, \{\mu_\theta\}_{\theta \in \Theta^*})$  is CML rationalizable if there exists a finite set of conditional probabilities  $\{c^t(\cdot|s)_{s \in S}\}_{t \in T}$  over  $\Theta^*$  such that for each  $\theta \in \Theta^*$ ,

$$\sum_{s' \in S} \mu(s') \max_{t \in T} c^t(\theta|s') > 0, \quad (1.3)$$

and for each  $s \in S$ ,

$$\mu_\theta(s) = \frac{\mu(s) \max_{t \in T} c^t(\theta|s)}{\sum_{s' \in S} \mu(s') \max_{t \in T} c^t(\theta|s')}. \quad (1.4)$$

Condition (1.3) requires that all observed signals are not unexpected. Condition (1.4) says that the DM updates according to CML. The following example shows that not all belief profiles are CML rationalizable.

**Example 1.3.3** Let  $S = \{s, s', s''\}$  and  $\Theta^* = \{\theta, \theta'\}$ . Suppose that the DM's prior over  $S$  is  $\mu = (1/3, 1/3, 1/3)$ . His posterior is  $\mu_\theta = (1/6, 1/2, 1/3)$  under signal  $\theta$  and  $\mu_{\theta'} = (1/6, 1/2, 1/3)$  under signal  $\theta'$ . We argue that  $(\mu, \{\mu_\theta, \mu_{\theta'}\})$  is not CML

---

<sup>6</sup>If  $\mu(s) = 0$ , we can simply delete state  $s$  and define the state space to be  $S \setminus \{s\}$ . If any posterior of the DM assigns positive probability to  $s$ , we can directly reject the hypothesis that the DM updates with CML.

<sup>7</sup>Again, if  $U_\theta$  is represented by a non-singleton set  $\mathcal{Q}_\theta$ , we can reject the hypothesis that the DM updates with CML.

rationalizable. The reason is that the posterior probabilities of state  $s$  are too small compared to the posterior probabilities of state  $s'$ . When signal  $\theta$  is realized, the conditional likelihoods ratio used for updating between  $s$  and  $s'$  is 1 : 3. Since the conditional maximum likelihood of  $\theta$  on  $s'$  is at most 1, the conditional maximum likelihood of  $\theta$  on  $s$  is at most 1/3. By a similar argument, the conditional maximum likelihood of  $\theta'$  on  $s$  is at most 1/3. Obviously, this is impossible, since it indicates that the summation of the conditional probabilities of  $\theta$  and  $\theta'$  on  $s$  is strictly less than one.

The next theorem completely characterizes the set of CML rationalizable belief profiles. The characterization exactly rules out the situation in Example 1.3.3: conditional on any state  $s$ , the summation of the probabilities of all signals cannot be strictly less than one.

**Theorem 1.3.4**  $(\mu, \{\mu_\theta\}_{\theta \in \Theta^*})$  is CML rationalizable if and only if for each  $s \in S$ ,

$$\sum_{\theta \in \Theta} \frac{\mu_\theta(s)}{\mu(s)} \left( \max_{s' \in S'} \frac{\mu_\theta(s')}{\mu(s')} \right)^{-1} \geq 1.$$

## 1.4 Application

In this section, we give three applications of CML. First, we show that CML is divisible, i.e., CML updating is unaffected by the order in which independent signals arrive. Second, we apply CML to study how DMs update with signals of unknown accuracy. We show that CML predicts under-reaction to such signals and illustrate how DMs learn with such signals. Finally, we apply CML to study an information design problem where a designer can introduce ambiguous information structures to affect the action taken by an agent. We show that if the agent follows CML, the designer can almost achieve the maximal payoff.

### 1.4.1 Divisibility

An important question in the literature of belief updating is whether DMs' beliefs are affected by the order independent signals arrive. An updating rule is divisible, or path-independent, if the posterior set of the DM is unaffected by the order of independent signals. Cripps (2018) gives the representations of divisible rules when there is no ambiguity. In this section, we show that CML satisfies divisibility.

For simplicity, let  $S$ ,  $\Theta$  and  $\Theta'$  be finite.  $\Theta$  and  $\Theta'$  are two different sets of signals. Let the DM's prior over  $S$  be  $\mu$ . The DM believes that the conditional distributions of  $\Theta$  and  $\Theta'$  on  $S$  have multiple possibilities, which are given by  $\{c^t(\cdot|s)_{s \in S}\}_{t \in T}$  and  $\{c^{t'}(\cdot|s)_{s \in S}\}_{t' \in T'}$ . The two signal sets  $\Theta$  and  $\Theta'$  are independent. Hence, the DM believes that the set of possible conditional distributions over  $\Theta \times \Theta'$  is given by  $\{c^{(t,t')}(\cdot|s)_{s \in S}\}_{(t,t') \in T \times T'}$  where for each  $(t, t') \in T \times T'$ ,  $c^{(t,t')}(\theta, \theta'|s) = c^t(\theta|s) \cdot c^{t'}(\theta'|s)$  for each  $\theta \in \Theta$  and  $\theta' \in \Theta'$ .

Fix the two independent signal sets. We argue that CML updating is path-independent. First, assume that signals  $\theta$  and  $\theta'$  arrive simultaneously. The DM's posterior  $\mu_{\theta, \theta'}$  satisfies that for any  $s \in S$ ,

$$\mu_{\theta, \theta'}(s) = \frac{\mu(s) \max_{(t,t') \in T \times T'} c^{(t,t')}(\theta, \theta'|s)}{\sum_{s' \in S} \mu(s') \max_{(t,t') \in T \times T'} c^{(t,t')}(\theta, \theta'|s')}.$$

Since  $c^{(t,t')}(\theta, \theta'|s) = c^t(\theta|s) \cdot c^{t'}(\theta'|s)$ , we have for any  $s \in S$ ,

$$\mu_{\theta, \theta'}(s) = \frac{\mu(s) \max_{t \in T} c^t(\theta|s) \max_{t' \in T'} c^{t'}(\theta'|s)}{\sum_{s' \in S} \mu(s') \max_{t \in T} c^t(\theta|s') \max_{t' \in T'} c^{t'}(\theta'|s')}.$$

Based on the updating formula, the conditional likelihoods used for updating is the multiplication of the conditional maximum likelihoods of the two signals on the each state.

Next, consider the case where signal  $\theta$  is observed first. The DM updates his prior  $\mu$  to  $\mu_\theta$  such that for each  $s \in S$ ,

$$\mu_\theta(s) = \frac{\mu(s) \max_{t \in T} c^t(\theta|s)}{\sum_{s' \in S} \mu(s') \max_{t \in T} c^t(\theta|s')}.$$

Then, he observes signal  $\theta'$  and updates  $\mu_\theta$  to  $(\mu_\theta)_{\theta'}$  such that for each  $s \in S$ ,

$$(\mu_\theta)_{\theta'}(s) = \frac{\mu_\theta(s) \max_{t' \in T'} c^{t'}(\theta'|s)}{\sum_{s' \in S} \mu_\theta(s') \max_{t' \in T'} c^{t'}(\theta'|s')}.$$

Obviously,  $\mu_{\theta, \theta'}$  and  $(\mu_\theta)_{\theta'}$  are the same. Similarly, if signal  $\theta'$  arrives before  $\theta$ , the DM's posterior remains to be  $\mu_{\theta, \theta'}$ . In conclusion, CML is divisible.

### 1.4.2 Information with Ambiguous Accuracy

SO19 and L19 study how DMs react to information of ambiguous accuracy. Some of their experimental findings are inconsistent with FB and ML under the max-min expected utility framework. For example, the experimental evidence of SO19 suggests that ambiguity averse DMs may not dilate their prior when processing ambiguous information. In addition to the non-dilation property, we show that CML can accommodate another experimental finding by L19: under-reaction to ambiguous information. We then discuss how DMs learn in the long run with CML when information has ambiguous accuracy.

We introduce the framework. Let  $S = \{s_1, s_2\}$  be the state space and  $\Theta = \{\theta_1, \theta_2\}$  the signal space. There are two possible levels of accuracy of the signals:  $H$  and  $L$ . We require that  $H, L \in (0, 1)$ ,  $H > L$  and  $H + L \geq 1$ . Hence, the DM believes that the conditional probabilities of the signals have two possibilities  $c^H$  and  $c^L$  where  $c^H(\theta_1|s_1) = c^H(\theta_2|s_2) = H$  and  $c^L(\theta_1|s_1) = c^L(\theta_2|s_2) = L$ .

The first observation is that when  $H = 1 - L$ , the DM will maintain his prior over  $S$  upon receiving any signal with CML. Since given any signal  $\theta_i$ , the conditional

maximum likelihood of  $\theta_i$  on  $s_i$  is  $H$  and the conditional maximum likelihood of  $\theta_i$  on  $s_j$  ( $j \neq i$ ) is  $1 - L$ . The two conditional likelihoods cancel out. This coincides with the experimental finding of SO19 that ambiguous averse DMs do not adjust their value of bets after such signals.

Next, we introduce the definition of under-reaction to unambiguous information by L19. Consider acts  $f$  and  $g$  that satisfy  $f(s_1) = g(s_2) = 1$  and  $f(s_2) = g(s_1) = 0$ . Consider another conditional distribution  $c(\cdot|s)_{s \in S}$  where

$$c(\theta_1|s_1) = c(\theta_2|s_2) = \frac{H + L}{2}.$$

Let  $v_f^a$  be the DM's ex post evaluation of  $f$  after observing  $\theta_1$  when he believes the conditional distributions of the signals are given by either  $c^H$  or  $c^L$ . Let  $v_f^u$  be the DM's ex post evaluation of  $f$  after observing  $\theta_1$  when he believes the conditional distributions of the signals are given by  $c$ .  $v_g^a$  and  $v_g^n$  are defined similarly. A DM *under-reacts* to ambiguous information if  $v_f^a < v_f^u$  and  $v_g^a > v_g^u$ . To interpret, observing signal  $\theta_1$  is good news for  $f$  and bad news for  $g$ . Hence, after observing  $\theta_1$ , the DM increases his evaluation of  $f$  and decreases his evaluation of  $g$ . If he under-reacts to ambiguous information, he increases less for his evaluation of  $f$  and decreases less for his evaluation of  $g$  when the information is ambiguous, i.e., there are multiple possible conditional distributions of the signals.

CML predicts under-reaction to ambiguous information when  $H + L > 1$ . To see this, let the DM's prior over  $\{s_1, s_2\}$  be  $\mu = (\alpha, 1 - \alpha)$ , where  $\alpha \in (0, 1)$ . Assume that signal  $\theta_1$  is observed. Given the set of interpretations  $\{c^H, c^L\}$ , the DM's CML posterior  $\mu_{\theta_1}$  satisfies

$$\frac{\mu_{\theta_1}(s_1)}{\mu_{\theta_1}(s_2)} = \frac{\alpha H}{(1 - \alpha)(1 - L)}.$$

Given the unambiguous interpretation  $\{c\}$ , the DM's CML posterior  $\mu_{\theta_1}^*$  satisfies

$$\frac{\mu_{\theta_1}^*(s_1)}{\mu_{\theta_1}^*(s_2)} = \frac{\alpha \frac{H+L}{2}}{(1-\alpha)(1-\frac{H+L}{2})}.$$

To see that the DM always under-reacts to ambiguous information, note that

$$\begin{aligned} \frac{\mu_{\theta_1}(s_1)}{\mu_{\theta_1}(s_2)} - \frac{\mu_{\theta_1}^*(s_1)}{\mu_{\theta_1}^*(s_2)} &= \frac{H}{1-L} - \frac{(H+L)/2}{1-(H+L)/2} \\ &= \frac{(H-L)(1-H-L)}{(1-L)(2-H-L)} \\ &< 0. \end{aligned}$$

Hence, the DM updates more when the information is unambiguous. The same argument applies for signal  $\theta_2$ .

We end this section by discussing the implications of CML on learning. Consider the situation in which the DM believes the conditional probabilities of the signals are given by either  $c^H$  or  $c^L$  where  $H+L > 1$ . Suppose that the true signal generating process is given by  $c$ . In this case, the information is asymmetrically informative. Fix the true state of the world. At each period, a signal from  $\{\theta_1, \theta_2\}$  is generated according to  $c$  independently. CML has simple predictions on learning in such an environment: as the number of signals observed by the DM increases, for any  $\epsilon > 0$ , the probability that the DM's posterior assigns probability more than  $1 - \epsilon$  to the true state converges to 1. That is, the DM will learn the true state of the world with the informative signals if she updates according to CML.

To see why the DM would learn the true state of the world, note that the DM always updates according to the conditional likelihood ratio  $\frac{H}{1-L}$  if  $\theta_1$  is observed and according to the ratio  $\frac{1-L}{H}$  if  $\theta_2$  is observed. Hence, if the proportion of signals  $\theta_1$  ( $\theta_2$ ) is strictly greater than the proportion of signals  $\theta_2$  ( $\theta_1$ ) and the number of the signals is large enough, the DM's belief would converge to state  $s_1$  ( $s_2$ ). Since the underlying

conditional probabilities are given by  $c$  which is informative, the probability that the signal with the same index as the true state has a higher proportion converges to 1. The DM learns the true state almost surely if he updates according to CML.

### 1.4.3 Information Design

The problem of information design is first studied by Kamenica and Gentzkow (2011), in which information structures are unambiguous. Beauchêne, Li, and Li (2019) study the case where the information designer can use ambiguous information structures, and the agent follows FB. In this section, we consider ambiguous information structures and assume that the agent follows CML.

Let a finite set  $S$  be the state space. A designer and an agent have a common prior  $\mu \in \Delta^o(S)$ . Let a finite set  $A$  be the set of all actions. The designer can design information structures for the agent. After observing a signal, the agent chooses an action. The utilities of the designer and the agent depend on both the true state and the action taken by the agent. Let  $v : S \times A \rightarrow \mathbb{R}$  be the designer's utility function. Let  $u : S \times A \rightarrow \mathbb{R}$  be the agent's utility function.

Since the agent follows CML, the posterior set of the agent is always a singleton. For any  $\mu^* \in \Delta(S)$ , let  $\mathcal{BR}(\mu^*)$  be the set of actions that maximize the expected utility of the agent under  $\mu^*$ , i.e.,

$$\mathcal{BR}(\mu^*) = \{a \in A : \sum_{s \in S} \mu^*(s)u(s, a) \geq \sum_{s \in S} \mu^*(s)u(s, a'), \forall a' \in A\}.$$

Let  $A^* \subseteq A$  be the set of all optimal actions of the agent, i.e.,

$$A^* = \bigcup_{\mu^* \in \Delta(S)} \mathcal{BR}(\mu^*).$$

We impose the following assumption on  $A^*$ .

**Assumption 1.4.1** For each  $a \in A^*$ , there exists  $\mu^* \in \Delta^o(S)$  such that  $a \in \mathcal{BR}(\mu^*)$ .

The assumption says that the agent's optimal actions can be achieved by posteriors in the interior of  $\Delta(S)$ . Note that this assumption is generically true.

We next define ambiguous information structures. An ambiguous information structure is a tuple  $(\Theta, \{c^t(\cdot|s)_{s \in S}\}_{t \in T})$  where  $\Theta$  is a finite set of signals and  $\{c^t(\cdot|s)_{s \in S}\}_{t \in T} \subseteq (\Delta(\Theta))^S$  is a finite set of signal generating systems. We require that for each  $\theta \in \Theta$ ,  $\max_{t \in T} c^t(\theta|s) > 0$  for some  $s \in S$ . Let  $\mu_\theta$  be the CML posterior of  $\mu$  under this information structure when signal  $\theta$  is observed, i.e.,  $\mu_\theta$  satisfies condition (1.2).

Consider an ambiguous information structure  $(\Theta, \{c^t(\cdot|s)_{s \in S}\}_{t \in T})$  and a set of actions  $\{a^\theta\}_{\theta \in \Theta}$ , we say that  $G = (\Theta, \{c^t(\cdot|s)_{s \in S}\}_{t \in T}, \{a^\theta\}_{\theta \in \Theta})$  is implementable if

- (i) For each  $\theta \in \Theta$ ,  $a^\theta \in \mathcal{BR}(\mu_\theta)$ .
- (ii) For any  $t, t' \in T$ ,

$$\sum_{s \in S} \sum_{\theta \in \Theta} \mu(s) c^t(\theta|s) v(s, a^\theta) = \sum_{s \in S} \sum_{\theta \in \Theta} \mu(s) c^{t'}(\theta|s) v(s, a^\theta).$$

Condition (i) says that  $a^\theta$  is the agent's optimal action when observing signal  $\theta$ . Condition (ii) ensures the commitment power of the designer: she has no incentive to pick any particular signal generating system since under each signal generation system, her expected payoff is the same. If  $G = (\Theta, \{c^t(\cdot|s)_{s \in S}\}_{t \in T}, \{a^\theta\}_{\theta \in \Theta})$  is implementable, let  $v^*(G)$  denote the designer's expected payoff under  $G$ , i.e.,

$$v^*(G) = \sum_{s \in S} \sum_{\theta \in \Theta} \mu(s) c^t(\theta|s) v(s, a^\theta), \forall t \in T.$$

Ideally, the designer's maximal payoff is given by

$$\sum_{s \in S} \left( \mu(s) \max_{a \in A^*} v(s, a) \right).$$

That is, under each state of the world, the agent takes the designer's optimal action from the feasible action set  $A^*$ . The next theorem states that the maximal payoff of the designer can be almost achieved with suitable information structures.

**Theorem 1.4.2** *Suppose that assumption 1.4.1 holds. For any  $\epsilon > 0$ , there exists an ambiguous information structure  $(\Theta, \{c^t(\cdot|s)_{s \in S}\}_{t \in T})$  and a set of actions  $\{a^\theta\}_{\theta \in \Theta}$  such that  $G = (\Theta, \{c^t(\cdot|s)_{s \in S}\}_{t \in T}, \{a^\theta\}_{\theta \in \Theta})$  is implementable and*

$$v^*(G) > \sum_{s \in S} \left( \mu(s) \max_{a \in A^*} v(s, a) \right) - \epsilon.$$

We give a simple example to illustrate the idea of Theorem 1.4.2. Let  $S = \{s_1, s_2\}$  and  $A^* = \{a_1, a_2\}$ .  $a_1$  is the designer's optimal action on state  $s_1$ .  $a_2$  is the designer's optimal action on  $s_2$ . The common prior of the designer and the agent is  $\mu = (1/2, 1/2)$ . The agent's optimal action is  $a_1$  if his posterior is  $\mu_1 = (1/3, 2/3)$ ; the agent's optimal action is  $a_2$  if his posterior is  $\mu_2 = (2/3, 1/3)$ .

Consider the following ambiguous information structure. There are 20000 signals  $\Theta = \{\theta^{j,l}\}_{j \in \{1, \dots, 1000\}, l \in \{1, 2\}}$  and 20000 signal generating systems  $\{c^{j,l}\}_{j \in \{1, \dots, 1000\}, l \in \{1, 2\}}$ . Our target is that for  $l \in \{1, 2\}$ , when signal  $\theta^{j,l}$  is realized, the agent's CML posterior is  $\mu_l$ . Thus, the agent chooses action  $a_l$  at signal  $\theta^{j,l}$  for any  $j$ . Moreover, we want the agent's posterior at signal  $\theta^{j,l}$  to be induced by  $c^{j,l}$  for each  $j$  and each  $l$ . We illustrate the construction of  $c^{j,1}$  for some  $j$  as an example. Let

$$c^{j,1}(\theta^{j,1}|s_1) = 0.01, \text{ and } c^{j,1}(\theta^{j,1}|s_2) = 0.02.$$

The two conditional probabilities indeed induce the posterior  $\mu_1 = (1/3, 2/3)$  given the prior  $\mu = (1/2, 1/2)$ . We want to allocate the rest of the conditional probabilities correctly, i.e., we want the rest of the conditional probabilities on each state to be

allocated to the signals that induce the designer's optimal action on the state. Let

$$c^{j,1}(\theta^{j',1}|s_1) = \frac{1 - 0.01}{10000 - 1}, \forall j' \neq j,$$

$$c^{j,1}(\theta^{j^*,2}|s_2) = \frac{1 - 0.02}{10000}, \forall j^*,$$

$$c^{j,1}(\theta^{j^*,2}|s_1) = 0, \forall j^*, \text{ and } c^{j,1}(\theta^{j',1}|s_2) = 0, \forall j' \neq j.$$

With  $c^{j,1}$ , conditional on  $s_1$ , the agent chooses action  $a_1$  with probability 1. Condition on  $s_2$ , the agent chooses action  $a_2$  with probability 0.98. Other signal generating systems can be constructed similarly. We need many signals to ensure the conditional probability of  $\theta^{j,l}$  to be very small under  $c^{j',l'}$  if  $(j, l) \neq (j', l')$  so that only  $c^{j,l}$  plays a role for controlling the agent's posterior at signal  $\theta^{j,l}$ . Note that the only payoff loss under  $c^{j,1}$  is when  $(s_2, \theta^{j,1})$  is realized. However, we can let the payoff loss be smaller by adding more signals and shrinking the conditional probability  $c^{j,1}(\theta^{j,1}|s_2)$ .

We note that the optimal payoff is usually not achievable. For example, assume that  $a \in A^*$  is the the designer's unique optimal action at state  $s$  and is not an optimal action of the designer at any other state. Assume further that the agent chooses  $a$  only if the probability of state  $s'$  ( $s' \neq s$ ) is greater than half. In this case, the designer must incur a payoff loss in order to induce action  $a$ , where the payoff loss comes from the situation in which  $a$  is chosen at state  $s'$ .

## 1.5 Discussion

### 1.5.1 A More General Framework: Violation of Axiom ISU

A natural extension of our framework is to allow the DM to have multiple priors over the state space. However, we show that axiom ISU will be violated by any updating rule. The only assumption we impose on the updating rule is that if  $(s, \theta)$  has

zero prior probability, then  $s$  has zero posterior probability after signal  $\theta$  is observed. Consider the following example.

**Example 1.5.1** Let  $S = \{s_1, s_2, s_3, s_4\}$  and  $\theta = \{\theta_1, \theta_2\}$ . The DM's prior set over  $S \times \Theta$  is  $\mathcal{P} = \{\alpha p^1 + (1 - \alpha)p^2 : \alpha \in [0, 1]\}$ , where  $p^1$  and  $p^2$  are shown in Table 1.3. Note that  $\mathcal{P}$  is not simple since it does not induce a unique prior over  $S$ .

Table 1.3: Distributions of the priors in Example 1.5.1

| $p^1$ | $\theta_1$ | $\theta_2$ | $p^2$ | $\theta_1$ | $\theta_2$ |
|-------|------------|------------|-------|------------|------------|
| $s_1$ | 0.1        | 0          | $s_1$ | 0.6        | 0          |
| $s_2$ | 0.7        | 0          | $s_2$ | 0.1        | 0          |
| $s_3$ | 0          | 0.1        | $s_3$ | 0          | 0.2        |
| $s_4$ | 0          | 0.1        | $s_4$ | 0          | 0.1        |

Let  $\theta_1$  be observed. Since only  $(s_1, \theta_1), (s_2, \theta_1), (s_3, \theta_2)$  and  $(s_4, \theta_2)$  have non-zero prior probabilities, the DM's posterior set is given by some  $\mathcal{Q} \subseteq \Delta(\{s_1, s_2\})$ . Assume without loss of generality that the payoff space is  $\mathbb{R}$ . Consider two extended acts  $f^*$  and  $g^*$  :

$$f^*(s_1, \theta_1) = 0, f^*(s_2, \theta_1) = 1, f^*(s_3, \theta_2) = 10, f^*(s_4, \theta_2) = x,$$

$$g^*(s_1, \theta_1) = 1, g^*(s_2, \theta_1) = 0, g^*(s_3, \theta_2) = -10, g^*(s_4, \theta_2) = y,$$

where  $x$  and  $y$  will be determined later on. The  $\mathcal{P}$ -evaluations of  $f^*$  and  $g^*$  are given by

$$\mathbb{E}_{\{p^1, p^2\}}(f^*) = \mathbb{E}_{p^1}(f^*) = 1.7 + 0.1x,$$

$$\mathbb{E}_{\{p^1, p^2\}}(g^*) = \mathbb{E}_{p^2}(g^*) = -1.4 + 0.1y.$$

After  $\theta_1$  is observed, the ex post evaluations are given by

$$\mathbb{E}_{\mathcal{Q}}(f^*|\theta) = \min_{\mu^* \in \mathcal{Q}} \mu^*(s_2),$$

$$\mathbb{E}_{\mathcal{Q}}(g^*|\theta) = \min_{\mu^* \in \mathcal{Q}} \mu^*(s_1).$$

Let  $\epsilon > 0$  be small enough. Increasing the payoff of  $f^*$  on  $(s_2, \theta_1)$  by  $\epsilon$  increases its  $\mathcal{P}$ -evaluation by  $p^1(s_2, \theta_1)\epsilon$  and increases the  $\mathcal{Q}$ -evaluation of  $f^*|\theta_1$  by  $\min_{\mu^* \in \mathcal{Q}} \mu^*(s_2)\epsilon$ . Similarly, increasing the payoff of  $g^*$  on  $(s_1, \theta_1)$  by  $\epsilon$  increases its  $\mathcal{P}$ -evaluation by  $p^2(s_1, \theta_1)\epsilon$  and increases the  $\mathcal{Q}$ -evaluation of  $g^*|\theta_1$  by  $\min_{\mu^* \in \mathcal{Q}} \mu^*(s_1)\epsilon$ . We can pick  $x$  and  $y$  such that

$$\mathbb{E}_{p^1}(f^*) = \mathbb{E}_{\mathcal{Q}}(f^*|\theta),$$

$$\mathbb{E}_{p^2}(g^*) = \mathbb{E}_{\mathcal{Q}}(g^*|\theta).$$

By axiom ISU, we conclude that

$$\min_{\mu^* \in \mathcal{Q}} \mu^*(s_2) \geq p^1(s_2, \theta_1) = 0.7,$$

and

$$\min_{\mu^* \in \mathcal{Q}} \mu^*(s_1) \geq p^2(s_1, \theta_1) = 0.6,$$

which is impossible.

Example 1.5.1 shows that axiom ISU is generally violated under the max-min expected utility framework. This gives a justification for our framework where ambiguity only comes from signals, in which ISU assumption works.

### 1.5.2 Comparison with the Proxy Rule

In this section, we compare CML with the proxy rule by Gul and Pesendorfer (forthcoming). To start with, we introduce the proxy rule. For simplicity, assume that  $S \times \Theta$  is finite. Proxy rule works for totally monotone capacities. Totally monotone capacities have the following equivalent multi-prior characterization. For any nonempty  $E \subseteq S \times \Theta$ , let  $\mathcal{P}_E$  be the set of all probability distributions that have

support  $E$ , i.e.,  $\mathcal{P}_E = \{p \in \Delta(S \times \Theta) : p(E) = 1\}$ . A totally monotone capacity has a multiple-prior representation given by

$$\sum_{E \subseteq S \times \Theta : E \neq \emptyset} \alpha_E \mathcal{P}_E, \quad (1.5)$$

where  $\alpha_E \in [0, 1]$  for each nonempty  $E \subseteq S \times \Theta$  and  $\sum_{E \subseteq S \times \Theta : E \neq \emptyset} \alpha_E = 1$ . Let  $|E|$  denote the cardinality of event  $E$ . With the proxy rule, when  $\theta$  is observed, the DM's set of posteriors is given by

$$\sum_{E \subseteq S \times \Theta : E \neq \emptyset} \frac{\alpha_E \cdot |E \cap (S \times \theta)| \cdot |E|^{-1}}{\sum_{E' \subseteq S \times \Theta : E' \neq \emptyset} (\alpha_{E'} \cdot |E' \cap (S \times \theta)| \cdot |E'|^{-1})} \mathcal{P}_{E \cap (S \times \theta)}.$$

A key property of the proxy rule is that “not all news is bad news”: given an information structure and an extended act  $f^*$ , the DM's ex post evaluation of  $f^*|\theta$  should be weakly higher than his evaluation of  $f^*$  under some signal  $\theta$ . The following example illustrates that CML violates “not all news is bad news”.

**Example 1.5.2** Let  $S = \{s, s'\}$  and  $\Theta = \{\theta, \theta'\}$ . The DM's prior set is a simple set  $\mathcal{P}$  which consists of all convex combinations of  $p^1, p^2, p^3$  and  $p^4$ , as shown in Table 1.4.

Table 1.4: Distributions of the priors in Example 1.5.2

|       |          |           |       |          |           |
|-------|----------|-----------|-------|----------|-----------|
| $p^1$ | $\theta$ | $\theta'$ | $p^2$ | $\theta$ | $\theta'$ |
| $s$   | 9/20     | 1/20      | $s$   | 1/20     | 9/20      |
| $s'$  | 2/5      | 1/10      | $s'$  | 1/10     | 2/5       |

|       |          |           |       |          |           |
|-------|----------|-----------|-------|----------|-----------|
| $p^3$ | $\theta$ | $\theta'$ | $p^4$ | $\theta$ | $\theta'$ |
| $s$   | 9/20     | 1/20      | $s$   | 1/20     | 9/20      |
| $s'$  | 1/10     | 2/5       | $s'$  | 2/5      | 1/10      |

Note that  $\mathcal{P}$  admits a totally monotone capacity since it satisfies condition (1.5):

$$\mathcal{P} = \frac{1}{20}\mathcal{P}_{\{(s,\theta)\}} + \frac{1}{20}\mathcal{P}_{\{(s,\theta')\}} + \frac{2}{5}\mathcal{P}_{\{(s,\theta),(s,\theta')\}} + \frac{1}{10}\mathcal{P}_{\{(s',\theta)\}} + \frac{1}{10}\mathcal{P}_{\{(s',\theta')\}} + \frac{3}{10}\mathcal{P}_{\{(s',\theta),(s',\theta')\}}.$$

To see that CML violates “not all news is bad news”, consider an extended act  $f^*$  such that  $f^*(s, \theta) = f^*(s, \theta') = 0$  and  $f^*(s', \theta) = f^*(s', \theta') = 1$ . The DM’s prior over  $S$  is  $\mu = (1/2, 1/2)$  according to  $\mathcal{P}$ . When  $\theta$  is realized, the DM’s CML posterior over  $S$  is  $(\frac{9}{17}, \frac{8}{17})$ . When  $\theta'$  is realized, the DM’s CML posterior is  $(\frac{9}{17}, \frac{8}{17})$ . Obviously, the DM lowers his evaluation of  $f^*$  after each signal. As a result, CML violates “not all news is bad news”. The next example shows that the proxy rule violates axiom ISU.

**Example 1.5.3** Let  $S = \{s, s'\}$  and  $\Theta = \{\theta, \theta'\}$ . The DM’s prior set is a simple set  $\mathcal{P} = \{\alpha p^1 + (1 - \alpha)p^2 : \alpha \in [0, 1]\}$ , where  $p^1$  and  $p^2$  are shown in Table 1.5.

Table 1.5: Distributions of the priors in Example 1.5.3

| $p^1$ | $\theta$ | $\theta'$ | $p^2$ | $\theta$ | $\theta'$ |
|-------|----------|-----------|-------|----------|-----------|
| $s$   | 1/3      | 1/6       | $s$   | 1/3      | 1/6       |
| $s'$  | 0        | 1/2       | $s'$  | 1/2      | 0         |

$\mathcal{P}$  admits a totally monotone capacity since

$$\mathcal{P} = \frac{1}{3}\mathcal{P}_{\{(s,\theta)\}} + \frac{1}{6}\mathcal{P}_{\{(s,\theta')\}} + \frac{1}{2}\mathcal{P}_{\{(s',\theta),(s',\theta')\}}.$$

The DM’s prior over  $\{s, s'\}$  is  $\mu = (1/2, 1/2)$  according to  $\mathcal{P}$ . With the proxy rule, when  $\theta$  is observed, the DM’s posterior over  $S$  is  $\mu^* = (4/7, 3/7)$ . Consider an extended act  $f^* : f^*(s, \theta) = f^*(s', \theta') = 1$ ,  $f^*(s', \theta) = 0$  and  $f^*(s, \theta') = 10/7$ . We can verify that the evaluation  $\mathbb{E}_{\mathcal{P}}(f^*)$  is  $4/7$ . With the proxy rule, the ex post evaluation of  $f^*|\theta$  is given by  $\mathbb{E}_{\mu^*}(f^*|\theta)$ , which is again equal to  $4/7$ . Consider another extended act  $g^*$  where  $g^*(s, \theta) = g^*(s', \theta') = 1$ ,  $g^*(s', \theta) = 1/2$  and  $g^*(s, \theta') = 10/7$ .  $g^*$  differs

from  $f^*$  only at  $(s', \theta)$ , where  $g^*$  yields a higher payoff. We have  $\mathbb{E}_{\mathcal{P}}(g^*) = 23/28 > \mathbb{E}_{\mu^*}(g^*|\theta) = 11/14$ . Hence, axiom ISU is violated by the proxy rule.

## Chapter 2

# Optimism in Choices over Menus

## 2.1 Introduction

<sup>1</sup>Understanding dynamic choices of a DM plays an important role in the analysis of various economic problems. When today’s choice affects future choices, the current choice of a DM reflects not only her current preference but also how she anticipates her future self to behave. Evidence from a large strand of literature suggests that DMs are typically optimistic towards their future choices (Larwood and Whittaker, 1977). Such optimism leads to contingent plans that are usually overturned. For instance, health-club members often choose to sign monthly or annual contracts with the gym instead of paying for per-visit passes, which turns out to cost them more than \$300 on average (DellaVigna and Malmendier, 2006); subjects may overestimate their effort input initially and tend to work less than planned in real effort tasks (Augenblick, Niederle, and Sprenger, 2015; Augenblick and Rabin, 2018; Fedyk, 2018).

In this chapter, we model the optimism of DMs through their ex-ante choices over menus. We consider a DM who has multiple guesses of her future preferences. For any given menu, different future preferences correspond to possibly different future choices. An optimistic DM always anticipates the future preference that generates the best future choice to be realized and hence evaluates the menu according to the best possible future choice.

Our model provides a clear connection between optimism and naivete of the DM. For instance, in the naive quasi-hyperbolic discounting model introduced by O’Donoghue and Rabin (1999, 2001), the DM discounts future utility by  $\delta \in (0, 1]$  and has an additional present bias parameter  $\beta \in (0, 1]$  capturing her taste for immediate gratification. The DM exhibits naivete when she underestimates her future present bias, i.e., she anticipates her future self to have present bias parameter  $\hat{\beta} \geq \beta$ . To see how this fits in our model, consider a DM who believes that her future present

---

<sup>1</sup>This chapter is based on the paper “Optimism in Choices over Menus” joint with Mu Zhang. The paper has been presented in 2021 North American Winter Meeting of the Econometric Society.

bias lies within the interval  $[\underline{\beta}, \bar{\beta}] \subseteq (0, 1]$  and is optimistic about her future choices. Our model predicts that she optimistically anticipates her future present bias to be  $\bar{\beta}$  since this is the most aligned future preference with her current one. Therefore, when the DM's actual present bias  $\beta$  is contained in the set  $[\underline{\beta}, \bar{\beta}]$ , she exhibits naive quasi-hyperbolic discounting. More recently, Ahn, Iijima, Le Yaouanq, and Sarver (2019) develop a behavioral notion of naivete by considering a DM who always prefers a menu to her ex-post choices from the menu. The DM's favor towards menus reflects her optimism.

We formally present our model in Section 2.2. Let  $X$  be a finite outcome space. A preference of the DM is represented by a von Neumann-Morgenstern expected utility function over the set of lotteries on  $X$ . A menu is a nonempty and compact subset of lotteries. Following Kreps (1979), the primitive of our model is a menu preference  $\succsim$ .<sup>2</sup> A DM has a current preference  $u$  and a set of anticipated future preferences  $\mathcal{V}$ . Given a menu  $A$ , the DM first considers the subset  $c(\mathcal{V}, A)$  of  $A$  which contains all *rationalizable future choices*, i.e., choices in  $A$  that optimize some preference in  $\mathcal{V}$ :

$$c(\mathcal{V}, A) = \{p \in A : \exists v \in \mathcal{V} \text{ s.t. } v(p) = \max_{q \in A} v(q)\}.$$

The DM then evaluates  $A$  according to the  $u$ -optimal choice in  $c(\mathcal{V}, A)$ . The induced menu preference, represented by the tuple  $(u, \mathcal{V})$ , is called an *optimistic menu preference* (OMP). With an OMP, menu  $A$  is preferred to menu  $B$  if and only if

$$\max_{p \in c(\mathcal{V}, A)} u(p) \geq \max_{q \in c(\mathcal{V}, B)} u(q).$$

Theorem 2.3.1 characterizes OMPs with seven axioms, among which four axioms are the same as or natural weakening of standard axioms in the literature of menu

---

<sup>2</sup>We will explicitly use the term “menu preference” to denote the DM's preference over menus throughout this chapter. Otherwise, a “preference” refers to an expected utility of the DM over lotteries.

preferences. (i) The menu preference is complete and transitive and (ii) satisfies independence of degenerate decisions and (iii) weak continuity. (iv) Allowing randomization over choices in any given menu does not change the utility of the menu.

The remaining three axioms are driven by optimism of the DM. For any given menu, the DM optimistically anticipates the best rationalizable choice to be chosen in the future. First, any choice in a given menu that is strictly better than the menu itself cannot be rationalized by any future preference of the DM. Hence, deleting such choices should not change the utility of the menu. This is *axiom independence of irrationalizable choices*. Second, consider two menus  $A$  and  $B$  as well as their union  $A \cup B$ . The best rationalizable future choice in menu  $A \cup B$  must also be rationalizable in either  $A$  or  $B$ . Therefore,  $A \cup B$  cannot be strictly better than  $A$  and  $B$  simultaneously. This is *axiom positive set-betweenness*. Finally, consider a randomized menu with probability  $\alpha$  to be  $A$  and probability  $1 - \alpha$  to be  $B$ . The DM would prefer this uncertainty to be resolved early since she can then tailor her anticipated future choice to the realized menu in order to be as optimistic as possible. As a result, the randomized menu whose uncertainty is resolved in the future cannot be strictly better than  $A$  and  $B$  simultaneously. This is *axiom preference for early resolution of uncertainty*. Together, the seven axioms are sufficient and necessary for a menu preference to be an OMP.

We present the uniqueness result in Theorem 2.4.3 and comparative statics of our model in Theorem 2.4.6. For both theorems, we adopt the notion “ $u$ -alignment” following Ahn, Iijima, Le Yaouanq, and Sarver (2019) for each preference  $u$ . A preference  $v$  is said to be *more  $u$ -aligned* than a preference  $v'$  if for any menu  $A$  and any choice  $q$  that is optimal in  $A$  under  $v'$ , there exists a choice  $p$  that is optimal in  $A$  under  $v$  such that  $u(p)$  is weakly greater than  $u(q)$ . Hence, if the DM’s current preference is  $u$  and she anticipates that her future preference set contains both  $v$  and  $v'$ , then she always ignores  $v'$  since future choices rationalized by  $v$  in any menu are

always weakly better than those rationalized by  $v'$ .  $u$ -alignment generates a pre-order  $\succeq_u$  over the set of preferences. Our uniqueness result says that if  $(u, \mathcal{V})$  and  $(u', \mathcal{V}')$  represent the same OMP, then  $u$  and  $u'$  represent the same preference, and the sets of  $\succeq_u$ -undominated preferences in  $\mathcal{V}$  and  $\mathcal{V}'$  are the same. For the comparative statics, consider two DMs with OMPs. DM1 is said to be *more optimistic* than DM2 if whenever DM2 prefers a menu to a singleton menu, so does DM1. We show by Theorem 2.4.6 that DM1 is more optimistic than DM2 if and only if they share the same current preference  $u$  and for each DM2's future preference  $v_2$ , there exists some DM1's future preference  $v_1$  such that  $v_1 \succeq_u v_2$ .

We apply our model to investigate intertemporal choices in Section 2.5. We demonstrate how our model accommodates naive quasi-hyperbolic discounting. We then show that our model generates the disjunction effect in choices over menus through a simple example: a DM can make the same choice in different scenarios due to different reasons but find no reason to make that choice when she is uncertain about which scenario is going to occur.

**Related Literature.** This chapter provides a richer framework to accommodate recent experimental findings on DMs' optimism. Breig, Gibson, and Shrader (2020) show that present bias is not the only source of procrastination through experiments. Their results suggest that procrastination can also be induced by excessively optimistic beliefs about future demands on an individual's time. This is accommodated by our model since in our model, DMs can exhibit optimism not only about her future present bias parameters but also about other parameters of future preferences.

This chapter is closely related to Strotz (1955) and Dekel and Lipman (2012) (henceforth DL12). Strotz (1955) considers a DM who anticipates her future preference set to be a singleton. DL12 and this chapter extend Strotz (1955) towards different directions by considering multiple future preferences: DL12 study a random version of the Strotz model and demonstrates how it relates to the costly self-control

model introduced by Gul and Pesendorfer (2001); our model takes a non-probabilistic approach and captures the optimism of the DM. Our model predicts that the DM prefers early resolution of uncertainty while the model of DL12 features indifference to the timing of resolution of uncertainty. As we will show, the intersection of our model and the model of DL12 is exactly the Strotz model.

In another related paper by Chandrasekher (2018) (henceforth C18), he considers a finite alternative space and studies the planner-doer model introduced by Thaler and Shefrin (1981). In his model, the doer (future self) has a unique preference, and the planner (current self) can restrict the feasible set of the doer in each menu using informal commitments. C18 gives a characterization for the menu preference of the planner. Our model coincides with the model of C18 when the alternative space is finite.<sup>3</sup> However, since we study menus of lotteries, our model requires different identification strategies and has additional implications on how uncertainty affects the DM's choices.

The remaining part of this chapter is organized as follows. We introduce the model in Section 2.2 and characterize it in Section 2.3. In Section 2.4, we discuss the uniqueness of our model and comparative statics. The applications of our model are given in Section 2.5. In Section 2.6, we discuss the connections between our model and other existing models of menu preferences. All omitted proofs are in Appendix B.

## 2.2 Model

Let  $X$  be a finite and nonempty outcome space.  $\Delta(X)$  denotes the set of probability distributions over  $X$  and is endowed with the Euclidean topology, which is induced by the Euclidean metric  $d$ . Elements in  $\Delta(X)$  are called lotteries. A menu is a nonempty closed subset of  $\Delta(X)$ .  $\mathcal{M}$  denotes the set of all menus and is endowed

---

<sup>3</sup>We formally show this in Appendix B.

with the Hausdorff topology, which can be induced by the Hausdorff metric  $d_h$ .<sup>4</sup> A menu preference is a binary relation  $\succsim$  defined over  $\mathcal{M}$ . As is standard in the literature, we use  $\sim$  and  $\succ$  to denote the symmetric part and the asymmetric part of the menu preference.

Let  $\mathbb{R}^X$  be the DM's set of possible preferences over lotteries. Each element  $u \in \mathbb{R}^X$  is an expected utility of the DM: For each lottery  $p \in \Delta(X)$ , the utility of  $p$  under  $u$  is given by  $u(p) = \sum_{x \in X} u_x p_x$ . Let  $o \in \mathbb{R}^X$  denote the origin. Following Ergin and Sarver (2010), we consider a normalized space of preferences

$$\mathcal{U} = \{u \in \mathbb{R}^X : \sum_{x \in X} u_x^2 = 1, \sum_{x \in X} u_x = 0\} \cup \{o\}.$$

Since  $\mathcal{U}$  is bounded and closed, it is compact. It is commonly known that we can uniquely identify each preference with elements in  $\mathcal{U}$ , up to positive affine transformations. We thus work with  $\mathcal{U}$  instead of  $\mathbb{R}^X$ .

For any nonempty  $\mathcal{V} \subseteq \mathcal{U}$  and any  $A \in \mathcal{M}$ , let  $c(\mathcal{V}, A)$  denote the set of choices in  $A$  that are rationalized by  $\mathcal{V}$ , i.e.,  $c(\mathcal{V}, A) = \{p \in A : \exists v \in \mathcal{V} \text{ s.t. } v(p) = \max_{q \in A} v(q)\}$ . To interpret,  $c(\mathcal{V}, A)$  contains the DM's all possible future choices in  $A$  given her set of future preferences  $\mathcal{V}$ .

**Definition 2.2.1** *A menu preference  $\succsim$  is an optimistic menu preference (OMP) if there is a tuple  $(u, \mathcal{V})$ , where  $u \in \mathcal{U}$  and  $\mathcal{V} \subseteq \mathcal{U} \setminus \{o\}$  is nonempty and closed, such that  $A \succsim B$  if and only if*

$$\max_{p \in c(\mathcal{V}, A)} u(p) \geq \max_{q \in c(\mathcal{V}, B)} u(q).$$

*When the above holds, the menu preference  $\succsim$  is said to be represented by  $(u, \mathcal{V})$ .*

---

<sup>4</sup>The Hausdorff metric  $d_h$  is defined as

$$d_h(A, B) := \max \left\{ \max_{p \in A} \min_{q' \in B} d(p, q'), \max_{q \in B} \min_{p' \in A} d(p', q) \right\}, \forall A, B \in \mathcal{M}.$$

Note that in Definition 1, we require  $\mathcal{V} \subseteq \mathcal{U} \setminus \{o\}$ . This is without loss of generality. If  $o \in \mathcal{V}$ , then for any menu  $A$ ,  $c(\mathcal{V}, A) = A$ , and thus  $\max_{p \in c(\mathcal{V}, A)} u(p) = \max_{q \in A} u(q)$ . This is equivalent to the case where  $u \in \mathcal{V}$ . Hence, we can rule out  $o$  from the set of possible future preferences.

If  $\succsim$  is represented by  $(u, \mathcal{V})$ , we interpret  $u$  as the DM's current preference and  $\mathcal{V}$  as the set of the DM's anticipated future preferences. For a given menu  $A$ , the DM considers all future rationalizable choices  $c(\mathcal{V}, A)$  and optimistically anticipates the best one to be chosen.

## 2.3 Axiomatization

We characterize OMPs in this section. To start with, we define some notations. We write  $p$  instead of  $\{p\}$  to denote the singleton menu, and use the terms “lotteries” and “singleton menus” interchangeably when there is no confusion. For a given menu  $A$ , let  $\text{conv}(A)$  be the convex hull of  $A$ , i.e.,  $p \in \text{conv}(A)$  if and only if  $p$  is a convex combination of elements in  $A$ . Note that  $\text{conv}(A)$  is also a well-defined menu. For each menu  $A$ , define  $A^\downarrow := \{p \in A : A \succsim p\}$  as the set of lotteries in  $A$  that are weakly worse than  $A$ . For any two menus  $A, B$  and any  $\alpha \in [0, 1]$ , let  $\alpha A + (1 - \alpha)B$  be the menu defined as

$$\alpha A + (1 - \alpha)B := \{\alpha p + (1 - \alpha)q : p \in A, q \in B\}.$$

$\alpha A + (1 - \alpha)B$  is the  $\alpha$ -mixture of menus  $A$  and  $B$ . The mixture menu can be interpreted as a randomization over the two menus such that the uncertainty is resolved in the future. The realized menu is  $A$  with probability  $\alpha$  and  $B$  with probability  $1 - \alpha$ .

The first axiom guarantees the rationality of the DM.

**Axiom 1—Weak Order (WO):**  $\succsim$  is complete and transitive.

The next axiom follows Ergin and Sarver (2010), which states that the preference ranking between two menus remains unchanged if they are mixed with the same singleton menu.

**Axiom 2—Independence of Degenerate Decisions (IDD):** For any  $A, B \in \mathcal{M}$  and any  $p \in \Delta(X)$ , if  $A \succ B$ , then for any  $\alpha \in (0, 1)$ ,

$$\alpha A + (1 - \alpha)p \succ \alpha B + (1 - \alpha)p.$$

For a given menu  $A$ , its convex hull  $\text{conv}(A)$  can be considered as the menu that allows the DM to randomize her choices in  $A$ . The next axiom states that allowing randomization over choices in any given menu does not change the attractiveness of the menu.

**Axiom 3—Indifference to Randomization (IR):** For any  $A \in \mathcal{M}$ ,  $A \sim \text{conv}(A)$ .

Since the DM's future preferences are all linear, when we allow for randomization, her future self will only randomize among choices that are optimal without randomization. Since the DM's current preference is also linear, such randomization in the future does not change the DM's current optimal utility. Hence, each menu is indifferent to its convex hull.

**Axiom 4—Weak Continuity (WC):** For any  $A, B, C \in \mathcal{M}$ ,  $\{B' \in \mathcal{M} : A \succ B'\}$  is open in  $\mathcal{M}$  and  $\{\alpha \in [0, 1] : \alpha B + (1 - \alpha)C \succ A\}$  is open in  $[0, 1]$ .

Axiom WC contains two parts. First, it says that the set of menus that are strictly worse than menu  $A$  is an open set. However, the set of menus that are strictly better than menu  $A$  is not open. To see why, consider menu  $A = \{p, q\}$ . Assume that the DM has only one possible future preference. Assume further that  $p$  is indifferent

with  $q$  under the DM's future preference, but is strictly better than  $q$  under the DM's current preference. Therefore, an optimistic DM would evaluate  $A$  according to  $p$ . However, when  $p$  is slightly perturbed to  $p'$  such that  $p'$  is strictly worse than  $q$  under the future preference, the DM realizes that her future self will choose  $q$  from the perturbed menu  $\{p', q\}$  and thus evaluates the menu according to  $q$ . Hence, the utility of a menu might discontinuously decrease when there is an infinitesimal perturbation. The second part of axiom WC fills the above gap by a weaker continuity condition.

The remaining three axioms are driven by optimism of the DM. The first of them is axiom independence of irrationalizable choices.

**Axiom 5—Independence of Irrationalizable Choices (IIC):** For any  $A, B \in \mathcal{M}$ , if  $A^\downarrow \subseteq B \subseteq A$ , then  $A \sim B$ .

Since any choice in a given menu that is strictly better than the menu itself cannot be rationalized by any future preference of the DM, deleting such choices should not change the utility of the menu.

**Axiom 6—Positive Set-Betweenness (PSB):** For any  $A, B \in \mathcal{M}$ , if  $A \succsim B$ , then  $A \succsim A \cup B$ .

Axiom PSB is introduced by Dekel, Lipman, and Rustichini (2009). To see how it is implied by optimism, consider two menus  $A, B$  and their union  $A \cup B$ . The best choice in  $A \cup B$  that can be rationalized by some future preference must also be a rationalizable choice in either  $A$  or  $B$ . Therefore  $A \cup B$  cannot be strictly better than  $A$  and  $B$  simultaneously.

**Axiom 7—Preference for Early Resolution of Uncertainty (PERU):** For any  $A, B \in \mathcal{M}$ ,  $A \succsim B$  implies  $\forall \alpha \in (0, 1)$ ,

$$A \succsim \alpha A + (1 - \alpha)B.$$

Axiom PERU is also called axiom aversion to contingent planning in Ergin and Sarver (2010). To interpret the axiom, consider a randomized menu with probability  $\alpha$  to be  $A$  and  $1 - \alpha$  to be  $B$ . Recall that  $\alpha A + (1 - \alpha)B$  can be interpreted as the randomized menu where the uncertainty is resolved after the future preference is realized. Hence, the DM's optimistic anticipation of her future choice does not depend on the realization of  $A$  or  $B$ . By comparison, if the uncertainty is resolved at the current stage, then the DM has more freedom to tailor her anticipated future choice to the realized menu in order to be as optimistic as possible. As a result,  $\alpha A + (1 - \alpha)B$  cannot be strictly better than  $A$  and  $B$  simultaneously.

We proceed to state our characterization theorem. For a given OMP  $\succsim$ ,  $\mathcal{V}$  is said to be the *maximal set of future preferences* of  $\succsim$  if there exists  $u \in \mathcal{U}$  such that (i)  $(u, \mathcal{V})$  represents  $\succsim$  and (ii)  $\mathcal{V}' \subseteq \mathcal{V}$  for any  $(u', \mathcal{V}')$  that also represents  $\succsim$ . Let  $\mathcal{V}(\succsim)$  denote the maximal set of future preferences of  $\succsim$  if such a set exists. Our main theorem asserts that the seven axioms fully characterize OMPs, and that  $\mathcal{V}(\succsim)$  always exists.

**Theorem 2.3.1** *A menu preference  $\succsim$  is an OMP if and only if it satisfies axioms WO, IDD, IR, WC, IIC, PSB and PERU. In addition, the current preference  $u \in \mathcal{U}$  is unique and the maximal set of future preferences of an OMP always exists.*

We sketch the proof of the sufficiency part. Let  $\mathcal{M}^F$  denote the set of finite menus. For any  $p \in A$ , let  $N(p, A)$  denote the set of preferences in  $\mathcal{U}$  that rationalize  $p$  in  $A$ , i.e.,  $N(p, A) = \{v \in \mathcal{U} : v(p) \geq v(q), \forall q \in A\}$ . We write  $A \succ^* B$  if  $p \succ q$  for all  $p \in A$  and  $q \in B$ .

By axioms WO, IDD and WC, we can identify the current preference  $u \in \mathcal{U}$  by restricting the menu preference to singleton menus. We then identify the maximal set of future preferences. Let  $\mathcal{T}$  be the set of tuples  $(A, p)$  such that  $A \in \mathcal{M}^F$ ,  $p \succ^* A$  and  $p \succ A \cup p$ . For any  $(A, p) \in \mathcal{T}$ ,  $N(p, A \cup p)$  contains no future preference of the DM since otherwise,  $p$  is rationalized in  $A \cup p$  by some future preference which leads

to  $p \sim A \cup p$ . The maximal set of future preferences is thus given by

$$\mathcal{V} = (\mathcal{U} \setminus \{o\}) \setminus \left( \bigcup_{(A,p) \in \mathcal{T}} N(p, A \cup p) \right).$$

The compactness of  $\mathcal{V}$  is ensured by axiom WC.

We verify that the tuple  $(u, \mathcal{V})$  indeed represents the menu preference  $\succsim$ . First, we focus on finite menus. The key observation is that for any  $(A, p)$  and  $(B, q)$  such that  $p \succ^* A$ ,  $q \succ^* B$  and  $N(p, A \cup p) = N(q, B \cup q)$ , it holds that  $p \succ A \cup p$  if and only if  $q \succ B \cup q$ , i.e.,  $(A, p) \in \mathcal{T}$  if and only if  $(B, q) \in \mathcal{T}$ . We prove this by showing that  $p \sim A \cup p$  and  $q \succ B \cup q$  lead to a contradiction. For illustrative purposes, we focus on the case where  $q \succ p \succ^* A \cup B$  and  $A \cup p \succ B \cup q$ . By axiom PERU, for any  $\alpha \in (0, 1)$ ,

$$A \cup p \succsim \alpha A \cup p + (1 - \alpha) B \cup q.$$

By taking  $\alpha$  close to 1, axiom WC implies that the utility of  $\alpha A \cup p + (1 - \alpha) B \cup q$  is close to the utility of  $A \cup p$ , and thus close to the utility of  $p$ . By axioms IR, PSB and IIC, we show that there exists  $r \in \alpha A \cup p + (1 - \alpha) B \cup q$  such that (i)  $r \sim \alpha A \cup p + (1 - \alpha) B \cup q$ , (ii)  $r$  is an extreme point<sup>5</sup> of the mixture menu, and (iii)  $p \succsim r$ . By condition (i),  $u(r)$  is close to  $u(p)$ . Therefore,  $r \in \alpha p + (1 - \alpha) B \cup q$ . Condition (iii) further implies that  $r \in \alpha p + (1 - \alpha) B$ . However, the unique extreme point of  $\alpha A \cup p + (1 - \alpha) B \cup q$  in  $\alpha p + (1 - \alpha) B \cup q$  is  $\alpha p + (1 - \alpha) q$ , which is a contradiction.

By a similar argument, we can show that if  $q \succ^* B$  and  $N(q, B \cup q) \cap \mathcal{V} = \emptyset$ , then  $q \succ B \cup q$ . We then prove that for any  $A \in \mathcal{M}^F$  and any  $p \in A$ , (i)  $p \succ A$  implies  $N(p, A) \cap \mathcal{V} = \emptyset$ , and (ii) there is some  $q \in A$  such that  $A \sim q$  and  $N(q, A) \cap \mathcal{V} \neq \emptyset$ .

---

<sup>5</sup> $r$  is an extreme point of a menu  $C$  if  $r$  is not a convex combination of any two different points in  $C$ .

This means the tuple  $(u, \mathcal{V})$  represents the menu preference  $\succsim$  over finite menus  $\mathcal{M}^F$ . Finally, we extend the representation from finite menus  $\mathcal{M}^F$  to all menus  $\mathcal{M}$ .

## 2.4 Uniqueness and Comparative Statics

Throughout this section, any menu preference is assumed to be an OMP. Let  $\mathcal{R}(\succsim)$  be the set of all tuples  $(u, \mathcal{V})$  representing the menu preference  $\succsim$ . We first give a characterization for  $\mathcal{R}(\succsim)$ .

By Theorem 2.3.1, if both  $(u, \mathcal{V})$  and  $(u', \mathcal{V}')$  represent  $\succsim$ ,  $u$  must be the same as  $u'$ . Thus, we only investigate the uniqueness of the set of future preferences. When the current preference  $u$  equals to  $o$ , we know that  $(o, \mathcal{V})$  represents  $\succsim$  for any nonempty and closed set  $\mathcal{V}$ . To avoid triviality, we only consider the case where  $u \neq o$ . Define  $\mathcal{W}_u := \{w \in \mathcal{U} \setminus \{o\} : u \cdot w = 0\}$ .  $\mathcal{W}_u$  contains all preferences over lotteries that are orthogonal to  $u$ . To proceed, we have the following definitions.

**Definition 2.4.1** *For any  $u, v, w \in \mathcal{U}$ , any  $\eta \in [-1, 1]$  and any  $\theta \in [0, 1]$ ,  $(\eta; \theta, w)$  is a  $u$ -decomposition of  $v$  if*

$$w \in \mathcal{W}_u \text{ and } v = \eta u + \theta w.$$

$\mathcal{D}_u(v)$  denotes the set of  $u$ -decompositions of  $v$ .

**Definition 2.4.2** *For any  $u, v, v' \in \mathcal{U} \setminus \{o\}$ ,  $v$  is more  $u$ -aligned than  $v'$ , denoted by  $v \succeq_u v'$ , if there exist  $(\eta; \theta, w) \in \mathcal{D}_u(v)$  and  $(\eta'; \theta', w) \in \mathcal{D}_u(v')$  such that  $\eta \geq \eta'$ .*

Our definition of  $u$ -alignment is equivalent to the definition by Ahn, Iijima, Le Yaouanq, and Sarver (2019), which adapts the technology developed by DL12. To understand the definitions, consider two preferences over lotteries  $v_1$  and  $v_2$  such that  $v_1 \succeq_u v_2$ . For any menu  $A$  and any choice  $q$  that is rationalized by  $v_2$  in  $A$ , the  $u$ -decompositions of  $v_1$  and  $v_2$  indicate that we can find a  $v_1$ -rationalizable choice  $p$

in  $A$  that is  $u$ -better than  $q$ . Consequently, for an optimistic DM, she ignores the future preference  $v_2$  if she anticipates  $v_1$  to be one of her possible future preferences. That is, only  $\succeq_u$ -undominated future preferences matter for the DM's preference over menus. This idea is formalized by the following uniqueness theorem.

**Theorem 2.4.3** *For any menu preference  $\succsim$  that is an OMP, if  $\{(u, \mathcal{V}), (u, \hat{\mathcal{V}})\} \subseteq \mathcal{R}(\succsim)$  where  $u \neq o$ , then for any  $v \in \mathcal{V}$ , there exists  $v' \in \hat{\mathcal{V}}$  such that  $v' \succeq_u v$ , and vice versa. That is, the sets of  $\succeq_u$ -undominated preferences in  $\mathcal{V}$  and  $\hat{\mathcal{V}}$  are identical.*

By Theorem 2.4.3, we can also identify the minimal set of future preferences of the DM, which is the set of  $\succeq_u$ -undominated preferences in  $\mathcal{V}(\succsim)$ . Denote this set as  $\mathcal{V}^\uparrow(\succsim)$ . The following straightforward corollary fully characterizes the DM's possible future preference sets.

**Corollary 2.4.4** *For any OMP  $\succsim$  that is represented by  $(u, \mathcal{V}(\succsim))$  where  $u \neq o$ ,  $(u, \mathcal{V}) \in \mathcal{R}(\succsim)$  if and only if  $\mathcal{V}^\uparrow(\succsim) \subseteq \mathcal{V} \subseteq \mathcal{V}(\succsim)$ .*

The uniqueness theorem motivates us to investigate the following comparative statics. Recall that the DM optimistically evaluates each menu according to the optimal choice under the current preference among rationalizable choices in the future. That is, her optimism favors non-singleton menus against singleton menus. Accordingly, DM1 is said to be *more optimistic* than DM2 if whenever DM2 prefers a menu to a singleton menu, so does DM1.

**Definition 2.4.5** *Menu preference  $\succsim_1$  is said to be more optimistic than menu preference  $\succsim_2$  if for any  $p \in \Delta(X)$  and any  $A \in \mathcal{M}$ ,  $A \succsim_2 p$  implies  $A \succsim_1 p$ .*

Suppose that DM1's menu preference  $\succsim_1$  is represented by  $(u, \mathcal{V}_1)$  and DM2's menu preference  $\succsim_2$  is represented by  $(u, \mathcal{V}_2)$ . If  $\mathcal{V}_1$  is *more  $u$ -aligned than  $\mathcal{V}_2$* , i.e., for each  $v' \in \mathcal{V}_2$ , there exists  $v \in \mathcal{V}_1$  such that  $v \succeq_u v'$ , then DM1 must evaluate each menu based on a  $u$ -better choice than DM2. This implies that  $\succsim_1$  is more optimistic than  $\succsim_2$ . The following theorem asserts that the reverse also holds.

**Theorem 2.4.6** *Consider two menu preferences  $\succsim_1$  and  $\succsim_2$  that are OMPs. Suppose that there exist lotteries  $p$  and  $q$  such that  $p \succ_1 q$ . The following statements are equivalent.*

- (i)  $\succsim_1$  is more optimistic than  $\succsim_2$ .
- (ii) For any  $(u_1, \mathcal{V}_1) \in \mathcal{R}(\succsim_1)$  and any  $(u_2, \mathcal{V}_2) \in \mathcal{R}(\succsim_2)$ ,  $u_1 = u_2$ , and for any  $v \in \mathcal{V}_2$ , there exists  $v_1 \in \mathcal{V}_1$  such that  $v_1 \succeq_u v$ .

We note that by Corollary 2.4.4, condition (ii) in Theorem 2.4.6 is equivalent to  $u_1 = u_2$  and  $\mathcal{V}(\succsim_2) \subseteq \mathcal{V}(\succsim_1)$ .

## 2.5 Application

In this section, we apply our model to investigate intertemporal choices. We show how our model accommodates naive quasi-hyperbolic discounting. After that, we demonstrate through a simple example that our model leads to the disjunction effect in the DM's choices over menus.

### 2.5.1 Naive Quasi-Hyperbolic Discounting

The naive quasi-hyperbolic discounting model has been studied in O'Donoghue and Rabin (1999, 2001). We show how to connect OMP with this model. Consider three time periods: period 1, period 2 and period 3. A DM's period-1 preference is characterized by a tuple  $(\theta_a, \theta_b)$ , where  $\theta_a \in (0, 1)$  is the DM's discount rate between periods 1 and 2, and  $\theta_b \in (0, 1)$  is the DM's discount rate between periods 2 and 3. Hence, when period  $t$ 's payoff is  $x_t$  for each  $t \in \{1, 2, 3\}$ , the DM's total payoff is given by  $u(x_1, x_2, x_3) = x_1 + \theta_a x_2 + \theta_a \theta_b x_3$ .

The DM's preference in period 2 is determined by her period-2 discount rate  $\theta$ . The DM's belief about her period-2 discount rates is given by a compact set  $\Theta \subseteq [0, 1]$ . Each element of  $\Theta$  represents one possible rate at which her

period-2 self discounts future payoffs. Her set of future preferences is thus  $\mathcal{V}_\Theta = \{v_\theta(x_1, x_2, x_3) = x_2 + \theta x_3, \theta \in \Theta\}$ . The DM's choices made in period 1 are fully determined by the tuple  $(\theta_a, \theta_b, \Theta)$ . The following proposition characterizes an optimistic DM's choice behavior for an arbitrary tuple  $(\theta_a, \theta_b, \Theta)$ .

**Proposition 2.5.1** *Consider a DM with an OMP whose period-1 preference and belief are given by  $(\theta_a, \theta_b, \Theta)$ . The DM can be characterized by*

- (i)  $(\theta_a, \theta_b, \{\theta_b\})$  when  $\theta_b \in \Theta$ ;
- (ii)  $(\theta_a, \theta_b, \{\max_{\theta \in \Theta} \theta\})$  when  $\theta_b > \max_{\theta \in \Theta} \theta$ ;
- (iii)  $(\theta_a, \theta_b, \{\min_{\theta \in \Theta} \theta\})$  when  $\theta_b < \min_{\theta \in \Theta} \theta$ ;
- (iv)  $(\theta_a, \theta_b, \{\underline{\theta}, \bar{\theta}\})$  when  $\min_{\theta \in \Theta} \theta \leq \theta_b \leq \max_{\theta \in \Theta} \theta$  and  $\theta_b \notin \Theta$ , where

$$\underline{\theta} = \max_{\theta \in \Theta: \theta < \theta_b} \theta \text{ and } \bar{\theta} = \min_{\theta \in \Theta: \theta > \theta_b} \theta.$$

Proposition 2.5.1 is a direct corollary of Theorem 2.4.3. Given the DM's period-1 preference  $u$ , her period-1 choices are only affected by her future preferences that are most aligned with  $u$  in  $\mathcal{V}_\Theta$ . Note that the DM's period-1 discount rate between period 2 and period 3 is  $\theta_b$ . Thus, any period-2 preference  $v_\theta$  is  $\succeq_u$ -undominated in  $\mathcal{V}_\Theta$  if and only if there exists no  $\theta' \in \Theta$  such that  $\theta'$  is "closer" to  $\theta_b$  than  $\theta$ . That is, the DM's current choices can only be affected by the discount rates in  $\Theta$  that are closest to  $\theta_b$  from above and from below.

Several observations of the proposition should be noted. First, when  $\theta_b > \max_{\theta \in \Theta} \theta$  and  $\theta_a \in \Theta$ , the DM exhibits naive quasi-hyperbolic discounting. To see this, define  $\delta := \theta_b$ ,  $\beta := \theta_a / \theta_b$ , and  $\hat{\beta} := \left( \max_{\theta \in \Theta} \theta \right) / \theta_b$ . In period 1, the DM discounts period-2 and period-3's payoffs by discount rates  $\beta\delta$  and  $\beta\delta^2$  respectively, and she optimistically anticipates her period-2 self's discount rate to be  $\hat{\beta}\delta$ .

Second, when  $\theta_b \in \Theta$ , the DM makes choices in period 1 as if she is fully naive. She has no concern about dynamic inconsistency because she anticipates her future self to share the same preference as her current self.

### 2.5.2 Disjunction Effect

In this section, we use a simple example to show that the DM's optimism can be mitigated by uncertainty that is resolved in the future. We provide a novel connection between this prediction and the disjunction effect (Tversky and Shafir, 1992) in choices over menus and demonstrate how to apply it to economic settings such as gym membership.

Consider a DM whose preference over menus is an OMP. The DM decides whether and when to buy a durable good in two periods. For simplicity, assume the discount factor is 1 and the DM's utility is additive across different periods. From the perspective of period 1, the DM's flow payoff of consuming the good in each period is  $g > 0$ . The price of the good in period 1 is  $\frac{19}{10}g$ . Hence the DM can secure herself the total payoff  $\frac{1}{10}g$  if she purchases the good in period 1. The DM anticipates her future preference to change: her period-2 self receives either  $\frac{1}{2}g$  or  $2g$  flow payoff from the good.

Consider case 1 where the period-2 price is  $\frac{4}{5}g$  for sure. An optimistic DM will delay the purchase decision in period 1 since when her period-2 self has flow payoff  $2g$ , she will buy the good in period 2 for sure, which yields her a higher total payoff  $\frac{1}{5}g$  than the payoff  $\frac{1}{10}g$  from purchasing the good in period 1.

Next, consider case 2 where the price in period 2 has probability  $\frac{1}{4}$  to be  $\frac{1}{2}g$  and probability  $\frac{3}{4}$  to be  $\frac{3}{2}g$ . The DM will also delay the purchase decision, since when her period-2 self has flow payoff  $\frac{1}{2}g$ , she will buy the good only if its price is  $\frac{1}{2}g$  in period 2. In this case, the DM's period-1 total payoff from delaying the purchase decision is  $\frac{1}{8}g$ , which is again higher than  $\frac{1}{10}g$ .

Finally, consider case 3, which is a randomization of the above two cases with equal probability: period-2 price of the good has probability  $\frac{1}{8}$  to be  $\frac{3}{2}g$ , probability  $\frac{3}{8}$  to be  $\frac{1}{2}g$  and probability  $\frac{1}{2}$  to be  $\frac{4}{5}g$ . By delaying the purchase decision in period 1, the DM's optimal total payoff is  $\frac{1}{16}g$ , which happens when her period-2 self has flow payoff  $\frac{1}{2}g$ . Since this payoff is strictly lower than  $\frac{1}{10}g$ , the DM would immediately purchase the good in period 1.

This simple example can be interpreted as the disjunction effect (Tversky and Shafir, 1992) in choices over menus.<sup>6</sup> That is, a DM can make the same choice in different scenarios due to different reasons but find no reason to make that choice if she is uncertain about which scenario will finally occur. We elaborate the reasons for the DM to delay her purchase in cases 1 and 2 as follows.

In case 1, the DM delays the purchase decision because the price of the good will decrease for sure and she knows that she will purchase the good in period 2 when her period-2 self has a high flow payoff. In case 2, the DM delays her purchase of the good because of the *value of information*. In particular, if her period-2 self has a low flow payoff, she will buy the good only when the period-2 price is low. Therefore, she optimistically believes that by waiting, she can benefit from the information as the price of the good will be realized in period 2.

However, in case 3, since the DM does not know whether case 1 or case 2 occurs, both reasons for delaying the purchase decision are weakened: The expected price of the good in period 2 is no longer attractive, and she does not gain from the information as much as case 2. Hence, she strictly prefers to buy the good immediately.

We end this section by briefly discussing how the results can be applied to other economic settings. For instance, if a health-club member does not anticipate future

---

<sup>6</sup>In Tversky and Shafir (1992), the disjunction effect is viewed as a violation of the Savage's sure-thing principle, where the DM prefers  $x$  to  $y$  conditional on knowing that event  $A$  occurs or knowing that event  $A$  does not occur, but reverses her preference if she does not know whether  $A$  occurs or not. See also Croson (1999) and Hristova and Grinberg (2008) for evidence of the disjunction effect in Prison's Dilemma games.

price of per-visit passes to decrease and believes that she will not make more informed decisions by waiting, then she may sign an annual contract with the gym even if she is not present-biased. This provides an alternative explanation for the findings of DellaVigna and Malmendier (2006).

## 2.6 Discussion of Axioms

In this section, we strengthen some of our axioms to show the connections between our model and other existing models of menu preferences. First, we introduce axiom continuity, which is a stronger version of axiom WC.

**Axiom Continuity (C):** For any  $A \in \mathcal{M}$ ,  $\{B \in \mathcal{M} : B \succ A\}$  and  $\{C \in \mathcal{M} : A \succ C\}$  are open.

Axiom C says that any infinitesimal perturbation of a menu cannot make its utility change drastically. The axiom is used in many menu preference models including Gul and Pesendorfer (2001), Dekel, Lipman, and Rustichini (2001), Dekel, Lipman, and Rustichini (2009), Stovall (2010), etc. The next theorem shows that our model reduces to two trivial cases when it satisfies axiom C.

**Theorem 2.6.1** *An OMP  $\succsim$  satisfies axiom C if and only if it has one of the following two representations:*

- (i)  $\exists u \in \mathcal{U}$  such that  $\forall A, B \in \mathcal{M}$ ,  $A \succsim B \Leftrightarrow \max_{p \in A} u(p) \geq \max_{q \in B} u(q)$ ,
- (ii)  $\exists u \in \mathcal{U}$  such that  $\forall A, B \in \mathcal{M}$ ,  $A \succsim B \Leftrightarrow \min_{p \in A} u(p) \geq \min_{q \in B} u(q)$ .

Next, we introduce axiom independence, which strengthens axiom IDD.

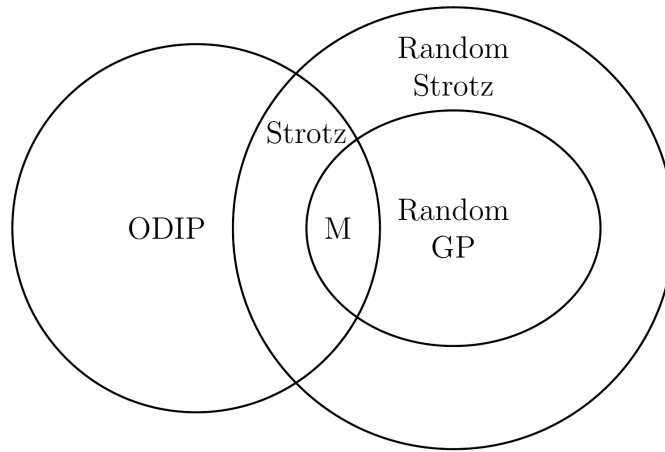
**Axiom Independence (I):** For any  $A, B, C \in \mathcal{M}$  and any  $\alpha \in (0, 1)$ ,  $A \succ B$  if and only if  $\alpha A + (1 - \alpha)C \succ \alpha B + (1 - \alpha)C$ .

Axiom I says that the DM's preference over two menus preserves if they are mixed with the same *menu*. Axiom I is used for characterizing linear menu preferences such as Gul and Pesendorfer (2001), Dekel, Lipman, and Rustichini (2001), etc. The axiom is also satisfied by the Strotz model and the random Strotz model by DL12. Our next theorem shows that our model reduces to the Strotz model if it satisfies axiom I.

**Theorem 2.6.2** *An OMP  $\succsim$  satisfies axiom I if and only if  $\succsim$  can be represented by  $(u, \mathcal{V})$  such that  $\mathcal{V}$  is a singleton.*

We summarize the relation between our model and other models in Figure 2.1, where M refers to the representation given by Theorem 2.6.1. The random GP model allows for random temptation utilities and is a generalization of the temptation and self-control model of GP01. DL12 show that the random GP model is a special class of the random Strotz model. Since the random GP model satisfies axiom continuity, it intersects with our model at the two trivial cases given by Theorem 2.6.1. Although not shown in Figure 2.1, our model also intersects with the model of GP01 at the two trivial cases.

Figure 2.1: Comparison with Different Menu Preference Models



## Chapter 3

# A Theory of Revealed Indirect Preference

## 3.1 Introduction

<sup>1</sup>This chapter explores the structure of indirect preference. Given a set of alternatives  $X$ , we refer to nonempty subsets of  $X$  as menus. A preference over menus constitutes an *indirect preference* so long as there is a preference over the alternatives in  $X$  such that menu  $A$  is preferred to another menu  $B$  whenever  $A$  contains an object that is preferred to every object in  $B$ .

The study of indirect preferences has a long history in economic theory. Indeed, a basic question in consumer theory concerns the recovery of the direct utility function (defined on bundles of  $\ell$  goods) from the indirect utility over the prices of those goods. In this case, plainly, a vector of prices (with income held fixed at some value) corresponds to a linear budget set, which is just a specific type of menu from the consumption space  $X = \mathbb{R}_+^\ell$ . It is well-known that the crucial property that makes it possible for a function defined over prices to be a bona fide indirect utility function is for it to be quasiconvex in prices (see Krishna and Sonnenschein (1990) and Jackson (1986)).

This question could be posed in a more general form that is not specific to the consumer theory context. Kreps (1979) considers *all* possible menus drawn from a set of alternatives  $X$  and shows that a preference over these menus constitutes an indirect preference if and only if it satisfies the following property: if an agent prefers menu  $A$  to  $B$ , then he is indifferent between  $A$  and  $A \cup B$ .<sup>2</sup> Tyson (2018) extends this result by characterizing indirect preference defined over a given *subcollection* of menus; note that Tyson’s result is also related to the classical result on indirect utility functions since the latter induces a preference defined over all linear budget

---

<sup>1</sup>This chapter is based on the paper “A Theory of Revealed Indirect Preference” joint with Gaoji Hu, Jiangtao Li and John K.-H. Quah. The paper has been presented in 2019 North American Winter Meeting of the Econometric Society.

<sup>2</sup>Kreps (1979) uses this as a benchmark for the axiomatization of preference for flexibility. Other models of menu preferences include (among others) Dekel, Lipman, and Rustichini (2001), Gul and Pesendorfer (2001), and Dekel and Lipman (2012).

sets, which forms a subcollection of menus among all possible menus drawn from the space of consumption bundles  $X = \mathbb{R}_+^\ell$ .

In this chapter, we consider an observer who has access to a finite collection of observations,  $\mathcal{M} := \{(A^t, B^t)\}_{t \in T}$ . At each observation  $t$ , the observer knows that the agent weakly prefers menu  $A^t$  to  $B^t$ ; and for a (possibly empty) subcollection  $S$  of those observations, the observer knows that  $A^t$  is strictly preferred to  $B^t$  for each  $t \in S$ . We say that the data set  $\mathcal{M}$  can be *rationalized* if the observed preference between each pair of menus is part of an indirect preference, i.e., there exists a preference over the space of alternatives  $X$  such that  $A^t$  is (strictly) preferred to  $B^t$  if and only if  $A^t$  contains an alternative that is (strictly) better than each alternative in  $B^t$ .

Our fundamental result establishes that a data set  $\mathcal{M}$  can be rationalized if and only if it satisfies the *never-covered property*; this is an intuitive property that could be defined via an iterative procedure. As a simple example of what it entails, suppose the observer knows that the agent strictly prefers menu  $A$  to  $B$  and weakly prefers  $A'$  to  $B'$ . Since there is an alternative in  $A$  that strictly dominates everything in  $B$ , a necessary condition for rationalization is that  $B$  does not contain  $A$ , i.e.,  $A \setminus B \neq \emptyset$ . But this is not all. We also need to check if  $A' \subseteq B$ ; if this holds, then clearly there is some element in  $A$  that dominates every element in  $B \cup B'$  and so  $A \setminus (B \cup B')$  must also be nonempty. In other words, some element must remain in  $A$  after all of  $A$ 's revealed dominated elements have been iteratively excluded; this property turns out to be both necessary and sufficient for rationalization.

Note that the setting of our result is different from that of Kreps (1979), Tyson (2018), or the characterization results on indirect utility over prices. In those papers, it is assumed that the observer knows the *complete* ranking over the collection of menus under consideration. On the other hand, our result allows the observer to have only an incomplete ranking over menus; for example, the observer need not

know the agent's preference between menus  $A^t$  and  $A^{t'}$ . The never-covered property reduces to the properties found in the earlier papers if the set of observations leads to a complete ranking over the collection of menus being considered.

Another paper that is closely related to this chapter is that of Fishburn (1976). That paper studies a different rationalizability problem (which is closely connected to the issue of coarse rationalizability discussed below), but its basic result could also be interpreted as a result on preference over menus; interpreted in that way, it provides a characterization of indirect preference (through a property called the *partial congruence axiom*) in the special case where  $A^t$  is *strictly* preferred to  $B^t$  for each  $t \in T$ . Our never-covered property is a generalization of the partial congruence axiom that allows for the possibility that the observer only knows that one menu is weakly preferred to another. This generalization is nontrivial and it is also crucial in certain applications, including the application to coarse rationalizability discussed below.

In some applications, it may not suffice to have a preference on  $X$  that rationalizes a data set  $\mathcal{M}$ ; it may also be desirable to have the preference be the extension of some given preorder. For example, in the case of consumer theory, it would be natural to require any rationalizing preference to be increasing in the product order on the consumption space  $X = \mathbb{R}_+^\ell$ . In cases where the space of alternatives  $X$ , and the menus defined on it, contain infinitely many elements, it is also natural to assume that there is a topology on  $X$  and to require preferences over  $X$  to be continuous, which guarantees the existence of optimal elements when menus are compact. We show that our basic result can be extended to incorporate these features. Last but not least, we show that the never-covered property can be verified via an efficient algorithm, which facilitates the empirical application of our results.

This chapter also discusses four applications of our theory.

(i) First, we revisit the question of characterizing indirect utility over prices. Instead of assuming that the entire indirect utility function is known, we assume that the observer only knows the consumer's preference for a finite set of price pairs, i.e.,  $p^t$  is preferred to  $q^t$  (for  $t = 1, 2, \dots, T$ ), with income normalized at 1. We show that there is an increasing, continuous and concave utility function that rationalizes the observed price preferences provided the latter satisfies a generalization of the quasi-convex property (on indirect utility functions). In the case in which  $p^t$  is strictly preferred to  $q^t$  for all  $t$ , the condition says the following: for any nonempty subset  $T'$  of observations, there exists some  $t' \in T'$  such that no  $q$  in the convex hull of  $\{q^s\}_{s \in T'}$  is weakly less than  $p^{t'}$ .

(ii) Consider a data set with  $T$  observations; at each observation  $t$ , it is observed that an agent chooses  $x^t$  from the menu  $C^t$ . Various versions of Afriat's theorem (1967)<sup>3</sup> answer the following question: what necessary and sufficient conditions on  $\mathcal{O} = \{(x^t, C^t)\}_{t \in T}$  guarantee the existence of a utility function  $U$  defined on  $X$  such that  $x^t \in \arg \max_{x \in C^t} U(x)$  for all observations  $t$ . Afriat's Theorem assumes that an observer knows precisely the choice made by the agent, but there are many scenarios where the observation is *coarse*, in the sense that the agent's choice cannot be precisely pinned down to  $x^t$ , but is known simply to come from a subset of  $C^t$ . For instance, suppose that goods are categorized into several categories, and the observer only knows the consumer's total spending on each category of goods at each observation. In this case, the observer will be able to infer that the consumer is choosing from a set  $A^t \subseteq C^t$  but not the precise choice itself. Another possibility is that the observations are known to be imperfect; even though  $y^t$  is the recorded choice, one may then wish to have a more permissive check of rationalization where the true choice is allowed to deviate from  $y^t$  in a limited way.

---

<sup>3</sup>See, for example, Varian (1982), Forges and Minelli (2009), Reny (2015), and Nishimura, OK, and Quah (2017). For a textbook treatment of the Afriat's theorem, see Kreps (2013) and Chambers and Echenique (2016).

These scenarios are formally captured by specifying an observation as  $(A^t, C^t)$ , where  $A^t \subseteq C^t$  contains the agent's choice. The corresponding rationalization condition requires that there be a utility function  $U$  and  $x^t$  such that  $x^t \in A^t$  and  $x^t \in \arg \max_{x \in C^t} U(x)$  for all observations  $t$ . Notice that this is equivalent to the condition that (considered as menus)  $A^t$  is weakly preferred to  $C^t$  for all  $t$ . It follows that the never-covered property could be used to characterize rationalizable data sets and our algorithm provides a practical way of checking for this property in empirical applications.<sup>4</sup>

The last two applications of our theory provide characterizations of choice models outside the classical paradigm.

(iii) In the multiple preferences model (Aizerman and Malishevski, 1981; Salant and Rubinstein, 2008), an agent's choice from a menu  $C$  could be the optimal choice for any one of a *set* of preferences. The question we pose is the following: suppose that for a finite collection of menus  $C^t$  we observe the agent's choices from each menu, which we denote by  $A^t$ ; when can we find a collection of preferences  $\{\succsim_i\}_{i \in I}$  such that  $Z^t = A^t$ , where  $Z^t$  consists of all elements of  $C^t$  that are optimal according to some preference  $\succsim'$  in  $\{\succsim_i\}_{i \in I}$ ? It turns out that this problem can be reformulated as a problem of rationalizing menu preferences and thus could be solved with a version of the never-covered property. Our result generalizes the finding in Aizerman and Malishevski (1981), who address this issue in the case where the choice set from *every* possible menu is observed.

(iv) The second model we consider is the minimax regret model (Wald, 1950; Savage, 1951). In this model, an agent has multiple utility functions over alternatives, drawn from a set  $\mathcal{U}$ . For a given menu  $C$ , the regret of alternative  $x$  under one of

---

<sup>4</sup>Fishburn (1976) considers a related problem where he requires  $\arg \max_{x \in C^t} U(x) \subseteq A^t$  for all observations  $t$ . This is equivalent to requiring the menu  $A^t$  to be strictly preferred to  $C^t \setminus A^t$  for all  $t$ . We think that our formulation of rationalizability is more appropriate in empirical applications; unlike Fishburn's formulation, we allow for the possibility that an alternative in  $C^t \setminus A^t$  is also optimal for the agent and it is the natural generalization of the rationalizability notion in Afriat's Theorem, which does allow for the optimality of alternatives in  $C^t \setminus \{x^t\}$ .

the agent's utility functions  $u \in \mathcal{U}$  is given by  $u(x) - \max_{y \in A} u(y)$ . The agent evaluates alternative  $x$  according to its maximal regret  $\max_{u \in \mathcal{U}} (u(x) - \max_{y \in A} u(y))$ , and chooses alternatives from the choice set that minimize the maximal regret. Using the never-covered property once again, we could find necessary and sufficient conditions on a data set under which the set of choices at each observed menu coincides exactly with the set of model consistent choices, for an appropriately chosen  $\mathcal{U}$ .

The rest of the chapter is organized as follows. In Section 3.2, we set out the notations that are used throughout the chapter. In Section 3.3, we introduce the main theorem, discuss several of its special cases, and also formulate the algorithm to test the never-covered property. The four applications of our theory are presented (respectively) in Sections 3.4, 3.5, 3.6, and 3.7. All omitted proofs are in Appendix C.

## 3.2 Preliminaries

We work with a fixed nonempty set  $X$ , which can be viewed as the universal set of alternatives. Let  $\mathcal{X}$  denote the collection of nonempty subsets of  $X$ . We refer to elements of  $\mathcal{X}$  as menus. Generic elements of  $X$  are denoted by  $x, y, z$ , etc, while generic elements of  $\mathcal{X}$  are denoted by  $A, B, C$ , etc.

A preorder  $\succeq$  on  $X$  is a binary relation on  $X$  that is reflexive and transitive.<sup>5</sup> We use  $\succ$  to denote the asymmetric part of  $\succeq$ . For a given preorder  $\succeq$  on  $X$  and a set  $A \in \mathcal{X}$ , we define  $A^\downarrow$  to be the decreasing closure of  $A$  with respect to the preorder  $\succeq$ , i.e.,

$$A^\downarrow := \{x \in X : y \succeq x \text{ for some } y \in A\},$$

---

<sup>5</sup>Terminology: a binary relation  $R$  on  $X$  is a nonempty subset of  $X \times X$ , but as usual, we write  $xRy$  instead of  $(x, y) \in R$ . We say that  $R$  is reflexive if  $xRx$  for each  $x \in X$ , transitive if  $xRyRz$  implies  $xRz$  for each  $x, y, z \in X$ , complete if either  $xRy$  or  $yRx$  holds for any  $x, y \in X$ , and anti-symmetric if for any  $x, y \in X$  such that  $x \neq y$ ,  $xRy$  and  $yRx$  do not hold simultaneously. The asymmetric part of  $R$  is defined as the binary relation  $P$  on  $X$  such that  $xPy$  if and only if  $xRy$  but not  $yRx$ .

and define  $A^{\Downarrow}$  to be the strictly decreasing closure of  $A$  with respect to the preorder  $\succeq$ , i.e.,

$$A^{\Downarrow} := \{x \in X : y \succ x \text{ for some } y \in A\}.$$

We write  $\max(A; \succeq) := A \setminus A^{\Downarrow}$  to denote the set of  $\succeq$ -undominated alternatives in  $A$ .

A *preference*  $\succsim$  on  $X$  is a complete preorder on  $X$ . We use  $\succ$  to denote the asymmetric part of  $\succsim$ . If  $\succeq$  is anti-symmetric, then we call  $\succ$  a *strict preference*. We say that  $x$  is a  $\succsim$ -maximal element in  $A$  if  $x \in A$  and  $x \succsim y$  for all  $y \in A$ , and write  $\max(A; \succsim)$  to denote the set of  $\succsim$ -maximal alternatives in  $A$ . When convenient, we write  $x \succsim A$  if  $x \succsim y$  for each  $y \in A$  and  $x \succ A$  if  $x \succ y$  for each  $y \in A$ .

Oftentimes, it is useful to study preferences restricted to a particular class. We say that the preference  $\succsim$  extends the preorder  $\succeq$  if

$$x \succsim y \text{ whenever } x \succeq y, \text{ and } x \succ y \text{ whenever } x \succ y.$$

We can think of  $\succeq$  as an exogenously given dominance relation on  $X$ , and view the statement  $x \succeq y$  as saying that  $x$  is an objectively better alternative than  $y$ , which the individual's preference  $\succsim$  should respect.

The general choice environment we have defined here broadly follows that in Nishimura, OK, and Quah (2017). As a basic example of this environment, we note that in consumer theory, the consumption space with  $n$  commodities is typically  $X = \mathbb{R}_+^n$  and it is common to assume that a consumer would always strictly prefer to have more of any good. In this case, the consumer's preference would extend the coordinatewise ordering  $\geq$  on  $\mathbb{R}_+^n$ .<sup>6</sup> Bear in mind that this setup allows for an arbitrary preference without a preorder restriction, since this can be thought of as the case in which the preference extends the trivial preorder  $\succeq$  where  $x \succeq y$  if and only if  $x = y$ .

---

<sup>6</sup>Formally,  $x \geq y$  if and only if  $x_i \geq y_i$  for each  $i \in \{1, 2, \dots, n\}$ , and  $x > y$  if and only if  $x \geq y$  and  $x_i > y_i$  for some  $i \in \{1, 2, \dots, n\}$ . We write  $x \gg y$  if and only if  $x_i > y_i$  for each  $i \in \{1, 2, \dots, n\}$ .

### 3.3 Rationalizability of menu preferences

In this section, we study the conditions under which a finite list of observed menu preference pairs collected from an agent is consistent with some (unobserved) preference on the underlying alternatives. The data collected by the observer is formally represented as  $\mathcal{M} := \{(A^t, B^t)\}_{t \in T}$ , where  $T$  is a nonempty finite index set and  $A^t$  and  $B^t$  are menus. For each  $t$ , the observer either knows that the agent weakly prefers  $A^t$  to  $B^t$  or that the agent strictly prefers  $A^t$  to  $B^t$ .<sup>7</sup> Let  $W$  be the collection of observations where  $A^t$  is weakly preferred to  $B^t$ , and let  $S$  be the collection of observations where  $A^t$  is strictly preferred to  $B^t$ . By definition,  $\{W, S\}$  is a partition of  $T$ . For any nonempty  $T' \subseteq T$ , we let

$$A(T') := \bigcup_{t \in T'} A^t \text{ and } B(T') := \bigcup_{t \in T'} B^t.$$

The following definition specifies precisely what it means for  $\mathcal{M}$  to be rationalized.

**Definition 3.3.1** *A set of menu preference pairs  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  is rationalized by a preference  $\succsim$  on  $X$  if for any  $t \in T$ , there exists  $x^t \in A^t$  such that*

- (i)  $x^t \succsim B^t$  and
- (ii)  $x^t \succ B^t$  if  $t \in S$ .

*We say that  $\mathcal{M}$  is rationalizable in this case. We say that a preference  $\succsim$   $\supseteq$ -rationalizes  $\mathcal{M}$  if  $\succsim$  rationalizes  $\mathcal{M}$  and extends  $\supseteq$ ; in this case,  $\mathcal{M}$  is  $\supseteq$ -rationalizable.<sup>8</sup>*

---

<sup>7</sup>This formulation includes the case in which the observer knows that the agent is indifferent between two menus (say)  $A$  and  $B$ , because this case could be considered as two observations, with the agent weakly preferring  $A$  to  $B$  in one observation and  $B$  to  $A$  in the other observation.

<sup>8</sup>Definition 3.3.1 is one of several possible formulations of the rationalizability of a set of menu preference pairs. For example, one could require that for any  $t \in T$ , (1) for any  $y \in B^t$ , there exists  $x \in A^t$  such that  $x \succeq y$  and (2) there exists  $x^t \in A^t$  such that  $x^t \succ B^t$  if  $t \in S$ . To see the (subtle) differences between these two formulations, suppose that the preorder is  $\succeq$ . The first data set has only one observation  $(B, B)$  where  $B = (0, 1)$  and the relation is weak. The second data set has only one observation  $(A, B)$  where  $A$  is the set of rational numbers in  $(0, 1)$ ,  $B = (0, 1)$ , and the relation is weak. Definition 3.3.1 would classify both data sets as not  $\succeq$ -rationalizable, while the alternative formulation would classify both as  $\succeq$ -rationalizable. That being said, in all the applications that we study (and in most economic environments), either  $X$  is finite or  $X$  is infinite but the menus are compact. In these cases, these two formulations are equivalent.

Our objective in this section is to characterize those sets of menu preference pairs  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  which can be  $\succeq$ -rationalized. Readers familiar with the revealed preference theory will notice that the issue is fairly straightforward when  $A^t$  is a singleton for each  $t$ . In that case, the problem is (in its essentials) within the scope of the well-known theorems of Afriat (1967) and Richter (1966) and their extensions; in the manner of those theorems, some version of a no-cycling condition on the revealed preference relations defined on  $\{x^t\}_{t \in T}$  (which will be stated formally later in Definition 3.3.4) is both necessary and sufficient for the existence of a preference that extends  $\succeq$  and satisfies (1) and (2) in Definition 3.3.1.<sup>9</sup> Of course,  $A^t$  is typically not a singleton. Thus, we could understand the issue before us in the following way: we have to formulate a property on  $\mathcal{M}$  that guarantees *the existence of a selection*  $x^t$  from  $A^t$ , such that the resulting (notional) set of observations  $\{(\{x^t\}, B^t)\}_{t \in T}$  satisfies the required no-cycling condition. The next subsection provides the property guaranteeing that such a selection exists.

### 3.3.1 The Never-Covered Property

Suppose that the data set  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  is  $\succeq$ -rationalized by some preference relation  $\succsim$ . Consider an arbitrary observation  $t$ . By the definition of  $\succeq$ -rationalizability,  $A^t$  contains an alternative  $x$  with  $x \succsim B^t$ . Since  $\succsim$  extends  $\succeq$ ,  $x \notin B^{t\downarrow\downarrow}$ . Moreover, if  $t \in S$ , then  $A^t$  contains an alternative  $x$  with  $x \succ B^t$  and so  $x \notin B^{t\downarrow}$ . Thus  $A^t$  cannot be covered by  $B^{t\downarrow\downarrow}$  if  $t \in W$ , and cannot be covered by  $B^{t\downarrow}$  if  $t \in S$ .

This argument could be generalized to more than one observation. For any nonempty subset  $T' \subseteq T$ , notice that (i) if  $x$  satisfies  $x \succsim B(T')$  then  $x \notin B(T')^{\downarrow\downarrow}$ ;

---

<sup>9</sup>To be precise, Nishimura, OK, and Quah (2017) already provides a condition on  $\{(\{x^t\}, B^t)\}_{t \in T}$  that is necessary and sufficient for the existence of a preference  $\succsim$  that extends  $\succeq$  and satisfies  $x^t \succsim B^t$  for each  $t \in T$ . In our case, we potentially have observations where we require  $x^t \succ B^t$ ; thus a modification of the condition in Nishimura, OK, and Quah (2017) is required to accommodate these cases, but this extension of their result is fairly straightforward.

and (ii) if  $x$  satisfies  $x \succ B(T' \cap S)$  then  $x \notin B(T' \cap S)^\downarrow$ . Thus, if  $A(T')$  contains an alternative  $\hat{x}$  satisfying both conditions, then  $A(T')$  cannot be covered by  $B(T')^{\downarrow\downarrow} \cup B(T' \cap S)^\downarrow$ . Indeed, we can find such an alternative  $\hat{x}$  in  $A(T')$ : for each  $t \in T'$ , pick  $x^t \in A^t$  such that  $x^t \succsim B^t$  if  $t \in W$  and  $x^t \succ B^t$  if  $t \in S$ ; then let  $\hat{x} = \max(\{x^t\}_{t \in T'}; \succsim) \in A(T')$ .

We are now ready to introduce the procedure that we call *the iterated exclusion of dominated observations*. Given a nonempty subset of observations  $T'$ , let  $\Phi^0(T') := \emptyset$ , and let  $\Phi^1(T')$  be the collection of observations  $t$  such that  $A^t$  is covered by  $B(T')^{\downarrow\downarrow} \cup B(T' \cap S)^\downarrow$ , i.e.,

$$\Phi^1(T') := \{t \in T' : A^t \subseteq B(T')^{\downarrow\downarrow} \cup B(T' \cap S)^\downarrow\}.$$

We know there is  $\hat{x} \in A(T')$  such that  $\hat{x} \succsim B(T')$  and  $\hat{x} \succ B(T' \cap S)$ . Then,  $\hat{x} \succ A^t$  for each  $t \in \Phi^1(T')$  and, since  $A^t$  is preferred to  $B^t$ , we obtain that  $\hat{x} \succ B^{t\downarrow}$  for each  $t \in \Phi^1(T')$ . Thus,  $\hat{x} \succ B((T' \cap S) \cup \Phi^1(T'))^\downarrow$ . Obviously,  $A(T')$  cannot be covered by  $B(T')^{\downarrow\downarrow} \cup B((T' \cap S) \cup \Phi^1(T'))^\downarrow$ . Let

$$\Phi^2(T') := \{t \in T' : A^t \subseteq B(T')^{\downarrow\downarrow} \cup B((T' \cap S) \cup \Phi^1(T'))^\downarrow\}.$$

Note that  $\Phi^1(T') \subseteq \Phi^2(T')$ . By a similar argument, we know that if  $t \in \Phi^2(T')$  then  $\hat{x} \succ B^{t\downarrow}$  and consequently,  $A(T')$  cannot be covered by  $B(T')^{\downarrow\downarrow} \cup B((T' \cap S) \cup \Phi^2(T'))^\downarrow$ .

We may repeat this argument for  $m = 2, 3, \dots$ , where

$$\Phi^{m+1}(T') := \{t \in T' : A^t \subseteq B(T')^{\downarrow\downarrow} \cup B((T' \cap S) \cup \Phi^m(T'))^\downarrow\}.$$

Since  $\Phi^m(T')$  is an increasing sequence in  $m$  in the set inclusion sense and the data set is finite, the procedure stops at  $m^*$  when  $\Phi^{m^*}(T') = \Phi^{m^*+1}(T')$ . Let  $\Phi(T') := \Phi^{m^*}(T')$ ; we refer to  $\Phi(T')$  as the set of *revealed dominated observations* (or simply *dominated*

observations) in  $T'$ . Since  $B(T')^{\downarrow\downarrow} \cup B((T' \cap S) \cup \Phi(T'))^{\downarrow}$  cannot contain  $\hat{x}$ , it does not cover  $A(T')$ . In other words,  $\Phi(T')$  is a strict subset of  $T'$ .

**Definition 3.3.2**  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  satisfies the never-covered property under  $\succeq$  if, for any nonempty  $T' \subseteq T$ , the set of revealed dominated observations  $\Phi(T')$  satisfies  $\Phi(T') \neq T'$ .

We have shown that the never-covered property under  $\succeq$  is a necessary condition for a data set to be  $\succeq$ -rationalizable. The main result of this chapter, Theorem 3.3.5, shows that it is also sufficient. The example below illustrates how we can use the never-covered property under  $\succeq$  to test the  $\succeq$ -rationalizability of a set of menu preference pairs.

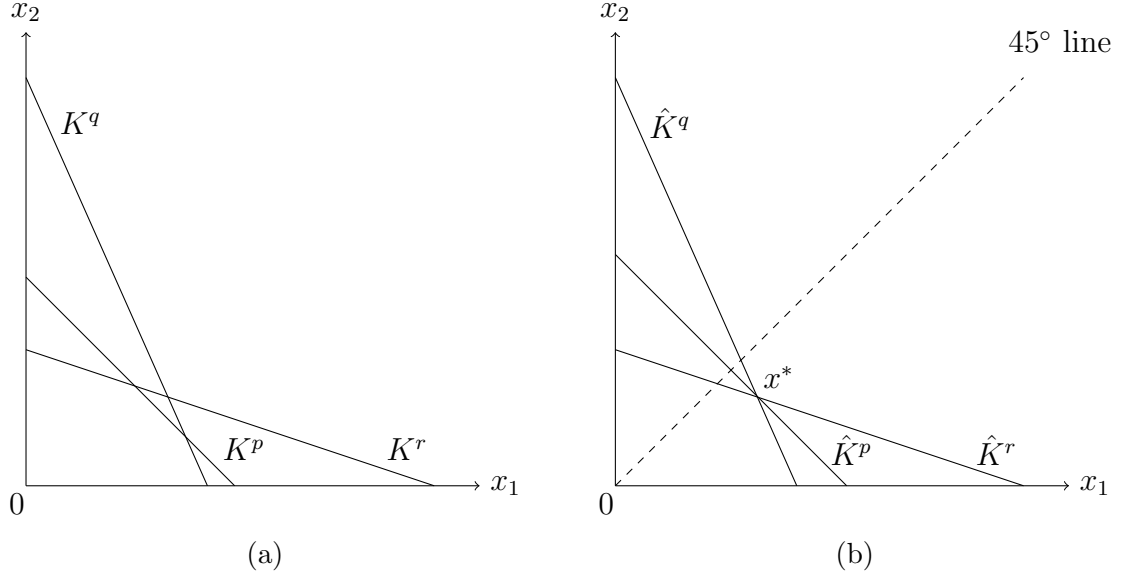
**Example 3.3.3** Consider the classic model of consumer demand with two goods. We take the preorder to be the coordinatewise ordering  $\geq$  on  $X = \mathbb{R}_+^2$ . Figure 3.1(a) depicts linear budget sets,  $K^p$ ,  $K^q$ , and  $K^r$ . Figure 3.1(b) depicts linear budget sets,  $\hat{K}^p$ ,  $\hat{K}^q$ , and  $\hat{K}^r$ .<sup>10</sup>

Suppose that  $\mathcal{M}$  consists of two observations,  $(K^p, K^q)$  and  $(K^q, K^r)$ , where both relations are weak. We claim that  $\mathcal{M}$  is not  $\succeq$ -rationalizable. Suppose to the contrary that  $\mathcal{M}$  is  $\succeq$ -rationalizable, say, by a preference relation  $\succsim$ . Then, there exists at least one bundle  $\hat{x}$  contained in  $K^p \cup K^q$  such that  $\hat{x} \succsim K^q \cup K^r$ . Since  $K^q \cup K^r = K^p \cup K^q \cup K^r$ , the bundle  $\hat{x}$  is  $\succsim$ -maximal in  $K^p \cup K^q \cup K^r$ . However, any  $\succsim$ -maximal alternative in  $K^p \cup K^q \cup K^r$  cannot be in  $K^p$  since (i)  $\succsim$  extends  $\geq$ ; and (ii)  $K^p$  is covered by  $(K^q \cup K^r)^o$ .<sup>11</sup> Furthermore,  $\hat{x}$  cannot be contained in  $K^q$  either; otherwise, there would be some  $y \in K^p$  such that  $y \succsim \hat{x}$  and thus  $y$  is  $\succsim$ -maximal in  $K^p \cup K^q \cup K^r$ , which is a contradiction. Notice that our argument corresponds precisely to a violation of the never-covered property under  $\geq$  for  $T' = T$ . Since  $B(T)^{\downarrow\downarrow} \cup B(T' \cap S)^{\downarrow} = (K^q \cup K^r)^o$  covers  $K^p$ ,  $\Phi^1(T)$  contains the observation

<sup>10</sup>The 45 degree line in Figure 3.1(b) will be used in Example 3.3.16.

<sup>11</sup>For any set  $K$ , we use  $K^o$  to denote its interior.

Figure 3.1: Data Sets: (a)  $\geq$ -Rationalizable and (b) not  $\geq$ -Rationalizable



$(K^p, K^q)$ . Since  $K^q \subseteq B(\Phi^1(T))$ ,  $K^q \subseteq B(T)^{\downarrow\downarrow} \cup B((T \cap S) \cup \Phi^1(T))^{\downarrow}$  and  $\Phi^2(T) = T$ . Thus,  $\Phi(T) = T$ .

On the other hand, the data set  $\hat{\mathcal{M}} = \{(\hat{K}^p, \hat{K}^q), (\hat{K}^q, \hat{K}^r)\}$ , where both relations are weak, is  $\geq$ -rationalizable. The optimal bundle in each set must be  $x^*$ , and it is easy to check that the set  $\Phi(T)$  is empty.

### 3.3.2 The Basic Result

Our proof of the sufficiency of the never-covered property under  $\supseteq$  in guaranteeing the  $\supseteq$ -rationalizability of  $\mathcal{M}$  proceeds by explicitly providing a way of selecting  $x^t$  in  $A^t$  for each  $t$  such that there exists a preference  $\succsim$  on  $X$  that extends  $\supseteq$  and satisfies (i)  $x^t \succsim B^t$  for all  $t \in T$  and (ii)  $x^t \succ B^t$  for  $t \in S$  (see Definition 3.3.1). Suppose that we have selected  $x^t$  from  $A^t$  for each  $t$  in some way. How do we check whether there exists a preference with the required conditions? This can be characterized by a no-cycling property which we now explain.

Let  $Y = \{x^t\}_{t \in T}$ . For  $x^t$  and  $x^{t'}$  in  $Y$ , we say that  $x^t$  is *revealed preferred* to  $x^{t'}$  and denote it by  $x^t R x^{t'}$  if  $x^{t'} \in B^{t\downarrow}$ , and we say that  $x^t$  is *revealed strictly preferred*

to  $x^{t'}$  and denote it by  $x^t P x^{t'}$  if either (i)  $x^{t'} \in B^{t\downarrow}$  or (ii)  $t \in S$  and  $x^{t'} \in B^{t\downarrow}$ . The following is a no-cycling condition on the binary relations  $R$  and  $P$ .

**Definition 3.3.4** *Given a data set  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$ , a selection  $x^t$  from  $A^t$  for each  $t \in T$  is a no-cycling selection under  $\succeq$  if the revealed preference relations  $R$  and  $P$  obey the following no-cycling property: there does not exist  $x^{t_1}, x^{t_2}, \dots, x^{t_n}$  in  $\{x^t\}_{t \in T}$  such that*

$$x^{t_1} R x^{t_2} R \dots R x^{t_n} \text{ and } x^{t_n} P x^{t_1}. \quad (3.1)$$

To see why this could be a plausible characterization, note that it is plainly a *necessary* condition. Indeed, suppose the preference  $\succsim$  extends  $\succeq$  and, with this preference,  $x^t$  satisfies (1)  $x^t \succsim B^t$  for all  $t \in T$  and (2)  $x^t \succ B^t$  for  $t \in S$ . If  $x^t$  is revealed preferred to  $x^{t'}$ , then by definition,  $x^t \succsim y \succeq x^{t'}$  for some  $y \in B^t$ ; since  $\succsim$  extends  $\succeq$  and  $\succsim$  is transitive, we obtain  $x^t \succsim x^{t'}$ . If  $x^t$  is revealed strictly preferred to  $x^{t'}$ , then we have either (i)  $x^t \succsim y \succ x^{t'}$  for some  $y \in B^t$  or (ii)  $t \in S$  and  $x^t \succ y \succeq x^{t'}$  for some  $y \in B^t$ ; in both cases, we may conclude that  $x^t \succ x^{t'}$ . Given the transitivity of  $\succsim$ , we plainly cannot have  $x^{t_1}, x^{t_2}, \dots, x^{t_n}$  in  $\{x^t\}_{t \in T}$  satisfying (3.1).

We have just shown that if a data set of menu preference pairs is  $\succeq$ -rationalizable, then it admits a no-cycling selection under  $\succeq$ . Theorem 3.3.5 below states that the converse is also true and that both are equivalent to the never-covered property under  $\succeq$ .

**Theorem 3.3.5** *Given a set of menu preference pairs  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  and a preorder  $\succeq$ , the following statements are equivalent:*

- (i)  $\mathcal{M}$  is  $\succeq$ -rationalizable.
- (ii)  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$ .
- (iii)  $\mathcal{M}$  admits a no-cycling selection under  $\succeq$ .

We end this subsection by providing a different formulation of the never-covered property under  $\succeq$ . The never-covered property under  $\succeq$  implies that for any nonempty

$T' \subseteq T$ , there exists some  $\Phi \subsetneq T'$  such that for any  $t \in T' \setminus \Phi$ ,  $A^t \not\subseteq B(T')^{\downarrow\downarrow} \cup B((T' \cap S) \cup \Phi)^\downarrow$ . To see this, we could simply set  $\Phi = \Phi(T')$ . The following proposition shows that the converse is also true.

**Proposition 3.3.6** *Consider the data set  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$ .*

(i) *For any nonempty  $T' \subseteq T$ , if there exists  $\Phi \subsetneq T'$  such that for any  $t \in T' \setminus \Phi$ ,*

$$A^t \not\subseteq B(T')^{\downarrow\downarrow} \cup B((T' \cap S) \cup \Phi)^\downarrow,$$

*then  $\Phi(T') \subseteq \Phi$ .*

(ii) *If for any nonempty  $T' \subseteq T$ , there exists  $\Phi \subsetneq T'$  such that for any  $t \in T' \setminus \Phi$ ,*

$$A^t \not\subseteq B(T')^{\downarrow\downarrow} \cup B((T' \cap S) \cup \Phi)^\downarrow,$$

*then  $\mathcal{M}$  satisfies the never-covered property under  $\supseteq$ .*

### 3.3.3 Algorithm

Given a subset  $T'$ , it is straightforward to check whether  $\Phi(T') = T'$ . Thus, Theorem 3.3.5 provides us with a way of checking if a set of menu preference pairs is  $\supseteq$ -rationalizable: we need to check whether  $\Phi(T') \neq T'$  for all  $T' \subseteq T$ . This may not seem promising as an empirical procedure, since for a data set with  $n$  observations, we would have to go through all  $2^n - 1$  nonempty subsets of  $T$  to guarantee the  $\supseteq$ -rationalizability of that set. In this subsection, we provide a simple algorithm to check whether the never-covered property under  $\supseteq$  holds. This algorithm requires us to check whether  $\Phi(T') \neq T'$  for at most  $n$  subsets of  $T$ . Thus, the never-covered property under  $\supseteq$  can be checked in an efficient manner.

Following the convention in the computer science literature, we use  $k'$  to denote the updated value of a variable  $k$ .

ALGORITHM. Set  $T^0 := T$ . Set  $k := 1$ .

START. Derive  $T^k := \Phi(T^{k-1})$ . Consider the following three mutually exclusive cases:

(i)  $T^k = \emptyset$ .

Stop and output  $\succeq$ -Rationalizable.

(ii)  $\emptyset \neq T^k \subsetneq T^{k-1}$ .

Go to START with  $k' = k + 1$ .

(iii)  $\emptyset \neq T^k = T^{k-1}$ .

Stop and output *Not*  $\succeq$ -Rationalizable.

Note that the ALGORITHM is effectively checking whether  $\Phi(T^k) = T^k$  for an endogenous sequence of subsets of  $T$ . We emphasize that, for a data set with  $n$  observations, the ALGORITHM necessarily terminates within  $n$  steps, and we only need to check at most  $n$  subsets of  $T$ .

Proposition 3.3.7 below provides the justification for this ALGORITHM.

**Proposition 3.3.7** *The set of menu preference pairs  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  is  $\succeq$ -rationalizable if and only if the ALGORITHM outputs  $\succeq$ -Rationalizable.*

### 3.3.4 Nice Rationalization when $\succeq$ is Trivial

The analysis in the previous subsections holds for any preorder  $\succeq$  and any partition  $\{W, S\}$  of  $T$ . In this subsection, we focus on the important special case in which the rationalizing preference is not required to extend any given preorder or, put another way, the preorder  $\succeq$  is simply the trivial preorder where  $x \succeq y$  if and only if  $x = y$ .

Consider a set of menu preference pairs  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  where  $S$  is nonempty.<sup>12</sup> Since  $\succeq$  is trivial,  $A^\downarrow = A$  and  $A^{\downarrow\downarrow} = \emptyset$  for all  $A$ . Thus, the procedure of iterated

---

<sup>12</sup>If  $\succeq$  is trivial, then any data set such that  $S = \emptyset$  is rationalizable (by the preference relation that the DM is indifferent among all alternatives).

exclusion of dominated observations reduces to the following: for any nonempty  $T' \subseteq T$ ,

$$\begin{aligned}\Phi^1(T') &= \{t \in T' : A^t \subseteq B(T' \cap S)\} \text{ and} \\ \Phi^{m+1}(T') &= \{t \in T' : A^t \subseteq B((T' \cap S) \cup \Phi^m(T'))\}, \text{ for } m = 1, 2, \dots\end{aligned}$$

This iteration must stop at some point, i.e., there is  $m^*$  such that  $\Phi^{m^*}(T') = \Phi^{m^*+1}(T')$ . The set of dominated observations is  $\Phi(T') := \Phi^{m^*}(T')$ . By definition,  $\mathcal{M}$  satisfies the never-covered property under the trivial preorder if  $\Phi(T') \neq T'$  for all nonempty  $T' \subseteq T$ . For the sake of brevity, we would simply refer to  $\mathcal{M}$  as *satisfying the never-covered property*, if it satisfies the never-covered property under the trivial preorder.

The following example is a simple illustration of the application of Theorem 3.3.5 and the ALGORITHM to this setting.

**Example 3.3.8** Let  $X = \{x, y, z, r, w\}$  and suppose that  $\succeq$  is trivial. The data set  $\mathcal{M}$  consists of the following observations:

$$\begin{aligned}A^1 &= \{x, y\}, B^1 = \{z, r, w\}; \\ A^2 &= \{y, z\}, B^2 = \{x, r, w\}; \\ A^3 &= \{x, r\}, B^3 = \{y, z, w\}; \text{ and} \\ A^4 &= \{r, w\}, B^4 = \{z\}.\end{aligned}$$

where  $A^1$  is strictly preferred to  $B^1$  and the other relations are weak. In this case,  $T = \{1, 2, 3, 4\}$  and  $S = \{1\}$ . Since  $A^4 \subseteq B^1$ , we obtain  $\Phi^1(T) = \{4\}$ . Since  $B^1 \cup B^4 = B^1$ ,  $\Phi^2(T) = \{4\}$ , and thus  $\Phi(T) = \{4\}$ . The ALGORITHM now directs us to calculate  $\Phi(\{4\})$ ; this set is empty and so the ALGORITHM concludes that  $\mathcal{M}$

is rationalizable. And indeed it is: for example, the preference  $x \sim y \succ r \sim w \sim z$  rationalizes  $\mathcal{M}$ .

The following definition imposes a stronger notion of rationalizability than the one provided in Definition 3.3.1 because the preference  $\succsim$  is required to have an optimum in every menu in the list of observations.

**Definition 3.3.9** *A set of menu preference pairs  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  is nicely rationalized by a preference  $\succsim$  on  $X$  if for all  $t \in T$ ,  $\max(A^t; \succsim)$  and  $\max(B^t; \succsim)$  are nonempty, and for any  $x^t \in \max(A^t; \succsim)$  and  $y^t \in \max(B^t; \succsim)$ , we have (i)  $x^t \succsim y^t$  and (ii)  $x^t \succ y^t$  if  $t \in S$ . In this case, we say that  $\mathcal{M}$  is nicely rationalizable. A preference  $\succsim$  nicely  $\supseteq$ -rationalizes  $\mathcal{M}$  if  $\succsim$  nicely rationalizes  $\mathcal{M}$  and extends  $\supseteq$ ; in this case, we say that  $\mathcal{M}$  is nicely  $\supseteq$ -rationalizable.*

In general, it is possible for a preference to rationalize a data set without it being a nice rationalization. Of course, this cannot happen when  $X$  is finite, since the existence of an optimum given a preference is then guaranteed. The next result says that it cannot happen when the preorder is trivial either, in the sense that every data set that is rationalizable (by some preference) is also nicely rationalizable (by a possibly different preference).

**Theorem 3.3.10** *The following statements on the data set  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  are equivalent:*

- (i)  $\mathcal{M}$  is rationalizable.
- (ii)  $\mathcal{M}$  satisfies the never-covered property.
- (iii)  $\mathcal{M}$  is nicely rationalizable.

**Example 3.3.11** As an illustration of Theorem 3.3.10, consider the case in which the data set consists of just one observation:  $A^1 = \{1\}$ ,  $B^1 = (0, 1)$ , with  $A^1$  strictly preferred to  $B^1$ . Can this data set be rationalized by a preference that extends the

standard total order  $\geq$  on  $\mathbb{R}$ ? Clearly, such a rationalization exists; in fact,  $\geq$  is itself the unique rationalization. However, this is not a nice rationalization since  $(0, 1)$  does not have an optimum according to  $\geq$ . On the other hand, Theorem 3.3.10 guarantees that there *is* a nice rationalization of the strict preference of  $A^1$  over  $B^1$  if we do not require the rationalizing preference to extend  $\geq$ . And indeed it does: simply let  $1 \succ r$  for all  $r < 1$ , and for all  $r, r' \in (0, 1)$  let  $r \sim r'$ .

### 3.3.5 Strict Menu Preferences and Strict Rationalization

We now turn to the case in which  $\succeq$  is trivial and  $T = S$ , so that the rationalizability of a data set  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  reduces to the following: there exists a preference  $\succsim$  on  $X$  such that for each  $t$ , there exists  $x \in A^t$  with  $x \succ B^t$ . This case is of particular interest, as the procedure of iterated exclusion of dominated observations ends in one round and the never-covered property has a much simpler form. To wit, since  $T = S$ , for any nonempty  $T' \subseteq T$ ,  $B(T' \cap S) = B(T')$ . Therefore, the procedure of iterated exclusion of dominated observations reduces to the following: for any nonempty  $T' \subseteq T$ ,

$$\begin{aligned}\Phi^1(T') &= \{t \in T' : A^t \subseteq B(T' \cap S)\} = \{t \in T' : A^t \subseteq B(T')\} \text{ and} \\ \Phi^2(T') &= \{t \in T' : A^t \subseteq B((T' \cap S) \cup \Phi^1(T'))\} = \{t \in T' : A^t \subseteq B(T')\} = \Phi^1(T').\end{aligned}$$

Therefore, the set of dominated observations is  $\Phi(T') = \{t \in T' : A^t \subseteq B(T')\}$ , and the never-covered property  $\Phi(T') \neq T'$  holds if and only if  $A(T') \not\subseteq B(T')$ .

We note that this special case of our result is (in its essentials) covered by Fishburn (1976), who establishes that there is  $\succsim$  such that  $x^t \in A^t$  with  $x^t \succ B^t$  for each  $t$  if and only if

$$A(T') \not\subseteq B(T') \text{ for all nonempty } T' \subseteq T. \quad (3.2)$$

Following Fishburn, we shall refer to the property (3.2) as the *partial congruence axiom*. In fact, Fishburn's result is somewhat more general because it partially covers the case in which  $T$  is infinite.<sup>13</sup> We have confined our attention to the case in which  $T$  is finite because it is the case most relevant to empirical applications and it allows us to formulate an efficient algorithm for checking the never-covered property.<sup>14</sup>

The following result summarizes our findings when menu preferences are strict.

**Corollary 3.3.12** *For a data set  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  where  $T = S$ , the following statements are equivalent:*

- (i)  $\mathcal{M}$  is nicely rationalizable.
- (ii)  $\mathcal{M}$  is nicely rationalizable by a strict preference.
- (iii)  $\mathcal{M}$  satisfies the partial congruence axiom.

The equivalence of the first and second statements in this corollary is due to Fishburn (1976, Lemma 1). As we have explained, the partial congruence axiom and the never-covered property are equivalent when  $T = S$  and thus the equivalence of the first and third statements follows from Theorem 3.3.10.

It almost goes without saying that when  $S$  is a strict subset of  $T$  (so that there are some observations where rationalization only requires  $x^t \in A^t$  such that  $x^t \succsim B^t$  rather than  $x^t \succ B^t$ ), the partial congruence axiom no longer characterizes rationalizable data sets and one needs to appeal to the never-covered property. Indeed, consider the data set in Example 3.3.8; is it possible for  $A^t \succ B^t$  for all  $t$ ? The answer is 'No' because  $A(T) = B(T)$  and the partial congruence axiom is violated. However,

---

<sup>13</sup>Fishburn (1976) considers two separate cases. For the case in which  $T$  could be infinite but  $A^t$  is required to be *finite* for each  $t$ , he shows that rationalizability is characterized by the partial congruence axiom. When  $T$  is countable and  $A^t$  is allowed to be infinite, his characterization result (Theorem 3) takes a different form, but it is equivalent to the partial congruence axiom when  $T$  is finite. The case in which  $T$  is more than countable and  $A^t$  is infinite is not covered by his results (or ours).

<sup>14</sup>If  $T$  is infinite, then there is obviously no hope of any algorithm for checking rationalizability. Fishburn's paper (perhaps partly because of its emphasis on the case of infinite  $T$ ) does not discuss algorithms for checking the partial congruence axiom.

if we only require the relation in the first observation to be strict (as we did in that example), then the data set *is* rationalizable because it satisfies the (weaker) never-covered property.

We now discuss the relationship between our work and the paper of de Clippel and Rozen (forthcoming). We first describe the problem they solve using our terminology. They assume that the preorder  $\succeq$  is trivial. They consider finite data sets where each observation  $t$  has the form  $(\{A_j^t\}_{j \in J(t)}, x^t)$ , where  $\{A_j^t\}_{j \in J(t)}$  is a collection of subsets of  $X$  and develop an algorithm that enables them to determine if  $\{(\{A_j^t\}_{j \in J(t)}, x^t)\}_{t \in T}$  admits an *upper contour rationalization* in the following sense: there is a strict preference  $\succ$  such that, at each  $t$ , there is a set in the collection  $\{A_j^t\}_{j \in J(t)}$  that is contained in the upper contour of  $x^t$ , i.e., there is  $A_{j(t)}^t$  in  $\{A_j^t\}_{j \in J(t)}$  such that  $A_{j(t)}^t \succ x$ .<sup>15</sup>

This problem and our menu rationalization problem have different economic motivations; however, when  $X$  is finite (so that all the relevant subsets in both problems are also finite), the two problems could be thought of as equivalent in the sense that it is always possible to convert one problem into the other, which also means that any algorithm developed for one could, in principle, be used to solve the other. That said, it should be clear from the conversion procedure (outlined in the Appendix) that there is no general computational reason for solving either problem in this roundabout fashion, since the converted data set would typically have more observations than the original data set. Thus, the two algorithms are best understood as distinct and serving different purposes.

### 3.3.6 Menu Preferences in Kreps (1979) and Tyson (2018)

Let  $\hat{\mathcal{X}} \subseteq \mathcal{X}$  be a nonempty collection of menus and let  $\succsim^M$  be a preference over  $\hat{\mathcal{X}}$  (which means that  $\succsim^M$  is a reflexive, transitive, and complete binary relation on  $\hat{\mathcal{X}}$ ). Abusing our terminology somewhat, we say that a preference  $\succsim$  on  $X$  nicely

---

<sup>15</sup>A version of their algorithm is contained in the first working paper version of their paper; see de Clippel and Rozen (2012).

rationalizes  $\succsim_M$  if, for all  $D \in \hat{\mathcal{X}}$ , the set  $\arg \max(D; \succsim)$  is nonempty and, for any  $x' \in \arg \max(D'; \succsim)$  and  $x'' \in \arg \max(D''; \succsim)$ , we have  $x' \succsim x''$  if  $D' \succsim^M D''$  and  $x' \succ x''$  if  $D' \succ^M D''$ .

Tyson (2018) shows that a menu preference  $\succsim^M$  on  $\hat{\mathcal{X}}$  is nicely rationalizable if and only if it satisfies the cover dominance condition, as defined below.

**Definition 3.3.13 (Tyson (2018))** *A menu preference  $\succsim^M$  over  $\hat{\mathcal{X}}$  satisfies the cover dominance condition if for any  $A, D \in \hat{\mathcal{X}}$  and  $\{B_i\}_{i \in I} \subseteq \hat{\mathcal{X}}$ ,*

$$A \succ^M B_i \text{ for each } i \in I \text{ and } D \subseteq \cup_{i \in I} B_i \Rightarrow A \succ^M D.$$

Furthermore, when  $\hat{\mathcal{X}}$  is finite and closed under union, the cover dominance condition can be equivalently stated in the following way.

**Definition 3.3.14** *Let  $\hat{\mathcal{X}}$  be closed under union. A menu preference  $\succsim^M$  over  $\hat{\mathcal{X}}$  is Kreps-consistent if, for any  $A, B \in \mathcal{X}$ ,  $A \succsim^M B$  implies that  $A \sim^M A \cup B$ .<sup>16</sup>*

Hence, Tyson (2018) generalizes the result in Kreps (1979) which states that, when  $X$  is finite, a menu preference on  $\mathcal{X}$  (the set of all nonempty subsets of  $X$ ) is nicely rationalizable if and only if it is Kreps-consistent.

In our setup, it is assumed that  $\mathcal{M}$ , a finite list of preference pairs between menus  $A^t$  and  $B^t$ , is observed. Its key difference with Kreps (1979) and Tyson (2018) is that these observations need not constitute a preference over all the menus in  $\hat{\mathcal{X}} = \{A^t\}_{t \in T} \cup \{B^t\}_{t \in T}$ ; in other words, while we posit that it is known that  $A^t$  is weakly preferred to  $B^t$ , we do not require the observer to know how  $A^t$  compares with  $A^{t'}$  or with  $B^{t'}$ .

Provided that the collection of menus  $\hat{\mathcal{X}}$  is finite, we could always construct a finite set of menu preferences  $\mathcal{M}^*$  from a preference  $\succsim^M$  on  $\hat{\mathcal{X}}$  in the following way:

---

<sup>16</sup>We denote by  $\sim^M$  the equivalence relation induced from  $\succsim^M$ .

$(A^t, B^t) \in \mathcal{M}^*$  if and only if  $A^t \succsim^M B^t$  and  $t \in S$  if and only if  $A^t \succ^M B^t$ .

Clearly,  $\mathcal{M}^*$  is nicely rationalizable (in the sense of Definition 3.3.9) by a preference  $\succsim$  if and only if  $\succsim$  nicely rationalizes the menu preference  $\succsim^M$ . By Theorem 3.3.10, the nice rationalizability of  $\succsim^M$  is characterized by the never-covered property on  $\mathcal{M}^*$ . In particular, this means that one can efficiently check if  $\succsim^M$  is nicely rationalizable by implementing our algorithm for checking the never-covered property. It also follows immediately from Theorem 3.3.10 and the results in Tyson (2018) that, when  $\hat{\mathcal{X}}$  is finite, the cover dominance condition on  $\succsim^M$  and the never-covered property on  $\mathcal{M}$  are equivalent; furthermore, if  $\hat{\mathcal{X}}$  is also closed under union, then each of these conditions is equivalent to Kreps-consistency.

### 3.3.7 Continuous Rationalizability

When the space of alternatives  $X$  is infinite, it is helpful to endow it with a topology and study continuous preferences. This guarantees (among other things) that the preference generates an optimum choice on compact menus and that the optimum varies continuously with the menu. For example, a continuous preference on the consumption space  $\mathbb{R}_+^n$  would guarantee that the demand correspondence is nonempty when prices are strictly positive (so that the budget set is compact) and varies continuously with prices.<sup>17</sup>

We say that a set of menu preference pairs  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  is  $\succeq$ -rationalized by a continuous utility function  $u : X \rightarrow \mathbb{R}$  if  $u$  represents a preference  $\succsim$  that  $\succeq$ -rationalizes  $\mathcal{M}$  in the sense of Definition 3.3.1. Theorem 3.3.15 below provides conditions under which  $\mathcal{M}$  can be rationalized by a continuous utility function.

---

<sup>17</sup>More generally, if we endow the collection of compact menus with the Hausdorff metric, then the correspondence mapping a compact menu to its optima is well-defined and upper hemicontinuous.

**Theorem 3.3.15** *Suppose that  $X$  is a locally compact and separable metric space and  $\succeq$  is a continuous preorder on  $X$ .<sup>18</sup> For the set of menu preference pairs  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  where  $B^t$  is compact for each  $t \in T$ , the following statements are equivalent:*

- (i)  $\mathcal{M}$  is  $\succeq$ -rationalizable.
- (ii)  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$ .
- (iii)  $\mathcal{M}$  is  $\succeq$ -rationalized by a continuous utility function  $u$ .

Note that this theorem does not assume that  $A^t$  is a compact set. It does assume that  $B^t$  is a compact set, which guarantees that for any continuous utility function  $u$ , the set  $\arg \max_{x \in B^t} u(x)$  is nonempty. If, in addition,  $A^t$  is a compact set for all  $t$ , then  $\arg \max_{x \in A^t} u(x)$  is also nonempty for all  $t$  and thus  $\mathcal{M}$  is *nicely*  $\succeq$ -rationalized by a continuous utility function  $u$  if  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$ .

The following example illustrates the application of Theorem 3.3.15.

**Example 3.3.16** Let  $X = \mathbb{R}_+^n$  be the consumption space with  $n$  goods. To capture the notion that a consumer strictly prefers more goods to less, we let the product order  $\geq$  to be underlying preorder. Then utility function  $u$  extends  $\geq$  if and only if it is *strictly increasing*, in the sense that  $u(x') > u(x)$  whenever  $x' > x$ . The order  $\geq$  is continuous in the Euclidean topology on  $\mathbb{R}_+^n$ . Suppose that  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$ , where  $A^t$  and  $B^t$  are compact sets; Theorem 3.3.15 guarantees that  $\mathcal{M}$  can be nicely rationalized by a strictly increasing and continuous utility function if and only if it obeys the never-covered property under  $\geq$ .

There are other preorders besides the product order that could be natural in this setting. For example,  $X = \mathbb{R}_+^n$  could be the space of contingent consumption, where the probabilities of each state are known (or part of the hypothesis). Then, based on those probabilities, different bundles in  $X$  could be ranked according to the first

---

<sup>18</sup>Terminology: a preorder  $\succsim$  is *continuous* (or *closed*) if  $\{(x, y) \in X \times X : x \succsim y\}$  is a closed set in  $X \times X$ .

order stochastic dominance, i.e.,  $x \geq_{FSD} y$  if  $x$  first order stochastically dominates  $y$ . For example, suppose that the states are equiprobable; then  $x \geq_{FSD} y$  and  $y \geq_{FSD} x$  if the entries in  $y$  is a permutation of those in  $x$ . In this case, a utility function that extends  $\geq_{FSD}$  is simply a utility function that is strictly increasing in  $\geq$  and symmetric.

Obviously,  $\geq_{FSD}$  is a finer order than  $\geq$  in the sense that  $\geq \subset \geq_{FSD}$ . It is also straightforward to check that  $\geq_{FSD}$  is a continuous preorder. Suppose that  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  and  $A^t$  and  $B^t$  are compact for all  $t$ ; by Theorem 3.3.15,  $\mathcal{M}$  satisfies the never-covered property under  $\geq_{FSD}$  if and only if it admits a nice rationalization by a continuous utility function that extends  $\geq_{FSD}$ .

Consider the example depicted in Figure 3.1(b), where  $\hat{K}^p$  is weakly preferred to  $\hat{K}^q$  in observation 1 and  $\hat{K}^q$  is weakly preferred to  $\hat{K}^r$  in observation 2. Since the never-covered property under  $\geq$  is satisfied, these observations can be rationalized by a continuous and strictly increasing utility function. However, they are not rationalizable by a preference that extends  $\geq_{FSD}$  when states 1 and 2 are equiprobable. Notice that for every  $(a, b) \in \hat{K}^p$ , there is  $(a', b') \in \hat{K}^q$  such that either  $(a', b') > (a, b)$  or  $(b', a') > (a, b)$ ; thus  $\hat{K}^p \subseteq \hat{K}^{q\downarrow}$ .<sup>19</sup> In formal terms,  $\Phi(\{1\}) = \{1\}$ , so the never-covered property is violated.

### 3.4 Application: Revealed Price Preference

One of the major themes in classical consumer theory is the recovery of the utility function from indirect utility. Formally, the question can be posed in the following way. Let  $v : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$  be a function. What necessary and sufficient conditions on  $v$  guarantee that

$$v(p) = \max\{u(x) : p \cdot x \leq 1\}$$

---

<sup>19</sup>From Figure 3.1(b), it is clear that  $\hat{K}^p$  is contained in the interior of  $\hat{K}^q \cup (\hat{K}^q)'$ , where  $(\hat{K}^q)'$  is the reflection of  $\hat{K}^q$  on the 45 degree line.

for some function  $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$  (interpreted as the consumer's utility function)? This question has been thoroughly studied (see, for example, Krishna and Sonnenschein (1990) and Jackson (1986)) and it is well-known that the distinctive property that  $v$  necessarily satisfies is quasi-convexity.

Our objective in this section is to develop a finite analog of this question, in the sense that instead of recovering  $u$  from the function  $v$  we ask what conditions would allow us to recover a preference on the underlying bundles that are consistent with a *finite list* of preferences over prices. Of course, the quick and short answer to the issue before us is the never-covered property, but the additional structure of the consumer problem, with linear budget sets in Euclidean space, allows us to say more.

We work with a data set with  $T$  observations, where at each observation  $t$ , the consumer reports either a weak or strong preference between two price vectors. Following our convention, if  $t \in W$ , then the consumer weakly prefers the price vector  $p^t$  to  $q^t$ . If  $t \in S$ , then the consumer strictly prefers the price vector  $p^t$  to  $q^t$ . Without loss of generality, we normalize the income of the consumer to be 1, so that the consumer's budget set at price  $p \in \mathbb{R}_+^n$  is

$$L(p) := \{x \in \mathbb{R}_+^n : p \cdot x \leq 1\}.$$

A preference for  $p^t$  over  $q^t$  means a preference for the budget  $L(p^t)$  over  $L(q^t)$ . Thus the set of preferences over budget sets may be denoted by  $\mathcal{M} = \{(L(p^t), L(q^t))\}_{t \in T}$ . The next result is an application of Theorem 3.3.15 to this environment.

**Theorem 3.4.1** *The following statements on  $\mathcal{M} = \{(L(p^t), L(q^t))\}_{t \in T}$  are equivalent:*

(i)  $\mathcal{M}$  can be rationalized by a locally nonsatiated preference on  $\mathbb{R}_+^n$ .<sup>20</sup>

(ii)  $\mathcal{M}$  satisfies the never-covered property under the product order  $\geq$ .

<sup>20</sup>Terminology: A preference  $\succsim$  is *locally nonsatiated* if for every  $x \in \mathbb{R}_+^n$  and every open neighborhood  $N$  around  $x$ , there  $x' \in N$  such that  $x' \succ x$ .

(iii)  $\mathcal{M}$  can be nicely rationalized by a strictly increasing, continuous, and concave utility function  $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$ .

A straightforward application of Theorem 3.3.15 tells us that  $\mathcal{M}$  satisfies the never-covered property under  $\geq$  if and only if it can be rationalized by a strictly increasing and continuous utility function. The latter statement is replaced in Theorem 3.4.1 by both a weaker statement (rationalization by a locally nonsatiated preference) and a stronger statement (rationalization by a strictly increasing, continuous, and *concave* utility function). A proof of this result is in the Appendix, but it is worth noting the following here. In establishing that Statement (i) implies (ii), we cannot simply appeal to the argument in Section 3.3.1 because in that case we make the assumption that the rationalizing preference  $\succsim$  extends a preorder; however, the local nonsatiation assumption on  $\succsim$  in this context allows us to retrace that argument, essentially because for any linear budget set  $L$  (assuming strictly positive prices),  $L^\downarrow = L$  and  $L^{\downarrow\downarrow}$  is the interior of  $L$ . As for the implication from (ii) to (iii), the linearity of the budget sets is crucial in guaranteeing that the rationalizing utility function can be chosen to be concave; our proof of that implication combines Theorem 3.3.15 with the Afriat's Theorem (see Afriat (1967)) which guarantees rationalization with a concave utility function.

We know from standard consumer theory that quasiconvexity plays a crucial role in the characterization of indirect utility functions. The never-covered property could be thought of as the finite analog to quasiconvexity. In the case in which all the price preferences are strict, i.e.,  $T = S$ , this connection is especially clear and is presented in Corollary 3.4.2 below.

To motivate the characterizing condition in Corollary 3.4.2, suppose that  $\mathcal{M}$  is rationalized by a strictly increasing, continuous, and concave utility function  $u$ . This implies that the indirect utility function  $v$  is strictly decreasing in prices and quasiconvex. Given  $T' \subseteq T$ , suppose that  $v(p^{t*}) \geq v(p^t)$  for all  $t \in T'$  and  $v(q^{s*}) \geq v(q^t)$

for all  $t \in T'$ . Since  $T = S$ ,  $v(p^{s*}) > v(q^{s*})$ . Since  $v$  is quasiconvex,  $v(q^{s*}) \geq v(q)$  for all  $q \in \text{conv}(\{q^s\}_{s \in T'})$ . Thus

$$v(p^{t*}) \geq v(p^{s*}) > v(q^{s*}) \geq v(q)$$

for all  $q \in \text{conv}(\{q^s\}_{s \in T'})$ . Since  $v$  is decreasing, we conclude that  $p^{t*} \not\geq q$  for any  $q \in \text{conv}(\{q^s\}_{s \in T'})$ . It turns out that this quasiconvex-like property is precisely equivalent to the never-covered property.

**Corollary 3.4.2** *The set of preferences over budget sets  $\mathcal{M} = \{(L(p^t), L(q^t))\}_{t \in T}$ , where  $T = S$ , is rationalizable by a strictly increasing, continuous, and concave utility function  $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$  if and only if the following property holds: for any nonempty  $T' \subseteq T$ ,*

*there exists some  $t \in T'$  such that  $p^t \not\geq q$  for any  $q \in \text{conv}(\{q^s\}_{s \in T'})$ .*

The proof of this result, as well as its extension to the case in which  $W$  is nonempty, can be found Appendix C.

## 3.5 Application: Coarse Rationalizability

So far in this chapter, we have considered the rationalization of a set of menu preference pairs. In this section, we discuss a formally related but economically distinct issue, namely, *the rationalization of choices from menus*. Our contribution is to provide a method for testing rationalizability in situations where observations are *coarse*, in a sense we shall make specific. Among other things, we provide an extension of the Afriat's Theorem to this environment.

### 3.5.1 Four Concepts of Rationalization

Suppose that at observation  $t$ , there is a menu  $C^t$  and a set  $A^t \subseteq C^t$ . Two notions of rationalization are commonly used in analyses of this type. The first concept requires a preference  $\succsim$  such that  $A^t = \max(C^t; \succsim)$  for all  $t \in T$ ; the Richter's Theorem (see Richter (1966)) characterizes data sets which are rationalizable in this sense. The second concept requires a preference  $\succsim$  such that  $A^t \subseteq \max(C^t; \succsim)$  for all  $t \in T$ ; the Afriat's Theorem (and its generalizations to nonlinear domains) characterize data sets that satisfy this concept of rationalization. Loosely speaking, the first notion of rationalization is the one most commonly used in the theoretical revealed preference literature; on the other hand, empirical work using revealed preference has mostly relied on the second, weaker notion, which is unsurprising since it does not posit that the observer has observed all the optimal choices, but only one (or some) of them.

A third concept of rationalization has been characterized in Fishburn (1976), where the set of optimal points is required to be contained in  $A^t$ ; in other words,  $\max(C^t; \succsim) \subseteq A^t$ . Obviously, Fishburn's concept generalizes the one in Richter's Theorem by allowing some elements of  $A^t$  to be nonoptimal, but it retains the requirement that nothing outside of  $A^t$  is optimal. This suggests that a fourth concept of rationalization may be useful in empirical applications: one that allows for the possibility that some elements in  $A^t$  are nonoptimal (following Fishburn) and also that some elements outside of  $A^t$  are optimal (following Afriat). In formal terms, it requires that  $\max(C^t; \succsim) \cap A^t \neq \emptyset$ .

The revealed preference literature since the 1970s have by and large neglected Fishburn's rationalization concept and its corresponding result. We think that Fishburn's concept, as well as the relaxation of that concept which we just proposed, deserves notice because they are relevant to empirical applications of revealed preference.

These concepts are applicable whenever there are *coarse observations*, where the observer knows (or hypothesizes) that there is an optimal choice found in  $A^t$ , but is

agnostic about precisely which alternatives within  $A^t$  are optimal. There are at least three broad scenarios where it is useful to think of coarse observations.

(i) The most obvious cases are those where the observations are simply known to be imprecise. For example, a researcher may have information on how much is spent on broad categories of goods, without knowing the allocation within each category. Alternatively, a researcher may have records on a consumer's credit card purchases, which puts a *lower* bound on how money is spent each month on different goods, but does not provide the precise breakdown of monthly expenditure since there could be goods bought with cash.

(ii) There could be situations where some alternative  $y^t$  is recorded as the choice from  $C^t$  but, in testing for rationality or estimating the preference, the researcher may wish to accommodate the possibility that choices were observed with error; this could be accomplished by defining a neighborhood  $A^t$  around  $y^t$  (in some sense appropriate to the specific context) and then checking if there is a preference  $\succsim$  with  $\max(C^t; \succsim) \cap A^t \neq \emptyset$  for all  $t \in T$ .

(iii) In experimental settings, it is common to find subjects whose choice behavior are not exactly consistent with rationality. Since the choices  $y^t$  are typically observed perfectly, it is implausible to attribute the rationality violations to observational errors. However, one could still use the size of the neighborhood  $A^t$  (suitably measured) as a way of comparing the rationality of different experimental subjects; those who require large  $A^t$ 's to rationalize their behavior can be deemed less rational than those where  $A^t$  is just a small neighborhood of  $y^t$ .

### 3.5.2 Coarse Data Sets and Menu Preferences

We consider an observer who has a finite set of coarse observations of a decision maker's choices. We denote a *coarse data set* by  $\mathcal{O} = \{(A^t, C^t)\}_{t \in T}$  where  $\{W, S\}$  is a partition of  $T$  and for each  $t \in T$  we have  $\emptyset \neq A^t \subseteq C^t$ . The interpretation is as

follows. When  $t \in W$ ,  $A^t$  contains at least one choice of the decision maker in  $C^t$ ; when  $t \in S$ ,  $A^t$  contains all the choices of the decision maker in  $C^t$ . The observer would like to recover a preference  $\succsim$  that rationalizes the data in the following sense.

**Definition 3.5.1** *A preference  $\succsim$  on  $X$  rationalizes the coarse data set  $\mathcal{O} = \{(A^t, C^t)\}_{t \in T}$  if*

- (i)  $\max(C^t; \succsim) \cap A^t \neq \emptyset$  for each  $t \in W$ , and
- (ii)  $\max(C^t; \succsim) \subseteq A^t$  for each  $t \in S$ .

*If  $\succsim$  exists, we say that  $\mathcal{O}$  is rationalizable. If  $\succsim$  can be chosen to rationalize  $\mathcal{O}$  and extend a given preorder  $\succeq$ , then  $\mathcal{O}$  is  $\succeq$ -rationalizable.*

Obviously, conditions (i) and (ii) in this definition correspond precisely to the fourth and third concepts of rationalization discussed in Section 3.5.1. Note that if  $T = W$ , then every coarse data set is trivially rationalized by a preference that is indifferent across all alternatives. So in the case, the rationalizability problem is interesting only if the preference is required to be locally nonsatiated or to extend some preorder  $\succeq$ .

Checking if a coarse data set is rationalizable is straightforward, given the results on menu preferences we have developed in Section 3.3. Indeed, for any coarse data set  $\mathcal{O} = \{(A^t, C^t)\}_{t \in T}$ , we could construct the following set  $\mathcal{M}^*$  of menu preference pairs: the menu  $A^t$  is weakly preferred to the menu  $C^t$  for all  $t \in T$  and  $A^t$  is strictly preferred to the menu  $C^t \setminus A^t$  for  $t \in S$ . Clearly,  $\mathcal{M}^*$  is rationalized by a preference  $\succsim$  if and only if the coarse data set  $\mathcal{O}$  is rationalizable by the same preference  $\succsim$ . Thus every result we have on the rationalizability (or  $\succeq$ -rationalizability) of the set of menu preference pairs  $\mathcal{M}^*$  has an analog for  $\mathcal{O}$ .

In the following subsection, we extend the Afriat's Theorem to coarse data sets.

### 3.5.3 A Generalization of Afriat's Theorem

We consider a data set  $\mathcal{O} = \{(A^t, L(p^t, y^t))\}_{t \in T}$  where for each  $t \in T$ ,  $p^t \in \mathbb{R}_{++}^n$  is the price vector,  $y^t$  is the total expenditure, and  $L(p^t, y^t) := \{x \in \mathbb{R}_+^n : p^t \cdot x \leq y^t\}$  is the budget set at observation  $t$ .<sup>21</sup> Departing from the standard setting of Afriat's Theorem, the observer does not know the exact choice of the consumer and only knows that the choice lies in  $A^t \subseteq L(p^t)$ . The following result provides us with a test of coarse rationalizability in this setting.

**Theorem 3.5.2** *Let  $\mathcal{O} = \{(A^t, L(p^t, y^t))\}_{t \in T}$  be a coarse data set where  $T = W$  and  $A^t \subseteq L(p^t, y^t)$  for all  $t \in T$ . The following statements are equivalent:*

- (i)  $\mathcal{O}$  can be rationalized by a locally nonsatiated preference.
- (ii)  $\mathcal{O}$  satisfies the never-covered property under the product order  $\geq$ .<sup>22</sup>
- (iii)  $\mathcal{O}$  can be rationalized by a strictly increasing, continuous, and concave utility function.

**Example 3.5.3** In studies of consumer demand, a researcher would often not have information on the demand for every relevant good. A common way to address this issue is to perform some aggregation procedure across goods, even though this approach is strictly valid only under stringent conditions on the utility function and/or the pattern of prices changes.

To be more specific, suppose that at observation  $t$ , the information available consists of the prices of all goods  $p^t \in \mathbb{R}_{++}^n$ , the demand for the first  $m - 1$  goods, and the total expenditure on the remaining goods (which we denote by  $c_{m,n}^t$ ). In other words, the actual demand for goods  $m, m + 1, \dots, n$  is not observed. To get round this problem, the researcher could construct a price index for those goods,  $\bar{p}_m^t$ , which

---

<sup>21</sup>Note that we departing from the convention and notation of the previous section by *not* normalizing expenditure at 1. This presentation is more appropriate in this section to highlight the fact that total expenditure  $y^t$  is part of the observer's data.

<sup>22</sup>In this statement, we are interpreting  $\mathcal{O}$  as a set of weak menu preferences.

would be a function of their prices  $(p_m^t, p_{m+1}^t, \dots, p_n^t)$ , with the corresponding demand for the composite good being  $\bar{x}_m^t = c_{m,n}^t / \bar{p}_m^t$ . In this way, the researcher creates a data set of the standard form, with prices  $(p_1^t, p_2^t, \dots, \bar{p}_m^t)$  and demand  $(x_1^t, x_2^t, \dots, \bar{x}_m^t)$  for  $m$  goods at each observation.

Coarse data sets offer a potentially useful alternative approach to tackle this problem. At observation  $t$ , the researcher observes  $x_i^t$  for  $i = 1, \dots, m-1$  and  $c_{m,n}^t$ . Thus the demand of the consumer must lie in the set

$$A^t = \{x \in \mathbb{R}_+^n : x_i = x_i^t \text{ for } i = 1, \dots, m-1 \text{ and } \sum_{i=m}^n p_i x_i = c_{m,n}^t\}.$$

The corresponding coarse data set is  $\mathcal{O} = \{(A^t, L(p^t, y^t))\}_{t \in T}$ , where  $y^t = \sum_{i=1}^{m-1} p_i^t x_i^t + c_{m,n}^t$ . This can be analyzed using Theorem 3.5.2.

As an illustration, suppose that  $\mathcal{O}$  consists of two observations where

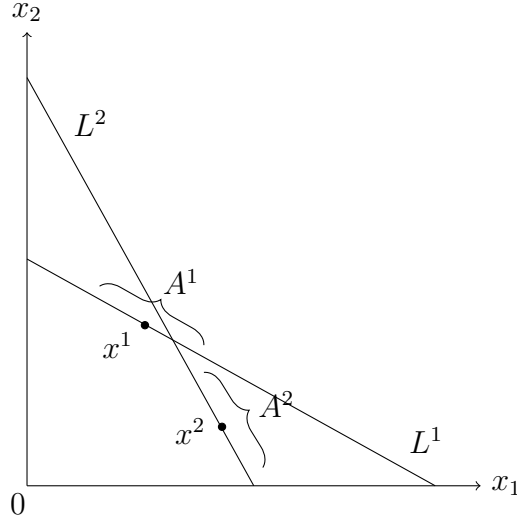
$$\begin{aligned} p^1 &= (2, 2.5, 3.5), & x_1^1 &= 1.5, & c_{2,3}^1 &= 9, & y^1 &= 12 \\ p^2 &= (4, 3, 3), & x_1^1 &= 3, & c_{2,3}^1 &= 4.5, & y^2 &= 16.5. \end{aligned}$$

This data set is coarse rationalizable. Indeed, the bundle  $\tilde{x} = (1.5, 9/2.5, 0)$  is in  $A^1$  but  $p^2 \cdot \tilde{x} = 16.8 > 16.5$ , so it is not in  $L(p^2, y^2)$ . Given that there are just two observations, this is enough to guarantee that  $\mathcal{O}$  satisfies the never-covered property under  $\geq$ .

On the other hand, suppose we were to aggregate goods 2 and 3 into a composite commodity, with the price of the composite good fixed at 3 at both observations 1 and 2. Then the demand for the composite good at these observations are  $\bar{x}_2^1 = 9/3 = 3$  and  $\bar{x}_2^2 = 4.5/3 = 1.5$ . The corresponding two-good data set has

$$\begin{aligned} p^1 &= (2, 3), & x^1 &= (1.5, 3), & y^1 &= 12; \\ p^2 &= (4, 3), & x^2 &= (3, 1.5), & y^2 &= 16.5. \end{aligned}$$

Figure 3.2: An Example of Observation Errors



Notes: The data set  $\{(x^1, L^1), (x^2, L^2)\}$  is not rationalizable, but  $\{(A^1, L^1), (A^2, L^2)\}$  (as depicted) is rationalizable.

It is straightforward to check that this data set violates GARP and is not rationalizable.

**Example 3.5.4** Consider a researcher who observes that the consumer chooses the bundle  $x^t$  from the budget set  $L(p^t, y^t)$ . To allow for the possibility that  $x^t$  was observed with error, the researcher could allow for the true consumption bundle to be in the set

$$A^t = \{x \in L(p^t, y^t) : p^t \cdot x = y^t \text{ and } |p_i^t x_i - p_i^t x_i^t| \leq k y^t \text{ for all } i\},$$

where  $k \in [0, 1]$ . In other words, the expenditure on good  $i$  is allowed to deviate from  $p_i^t x_i^t$  but not by more than  $k y^t$ . This is illustrated in Figure 3.2, where the ‘original’ data set  $\{(x^1, L^1), (x^2, L^2)\}$  is not rationalizable, but  $\{(A^1, L^1), (A^2, L^2)\}$  (as depicted) is rationalizable. More generally, our extension of the Afriat’s Theorem provides a way to check if  $\mathcal{O} = \{(A^t, L(p^t, y^t))\}_{t \in T}$  is rationalizable.

### 3.6 Application: Multiple Preferences

In this section, we investigate the observable restrictions of the multiple preferences model; see, for example, Aizerman and Malishevski (1981) and Salant and Rubinstein (2008). In contrast with the single preference model, the choice behavior of the DM may be a result of multiple rationales. Formally, the DM has a set  $\Pi$  of strict preferences, and she chooses

$$f_{\Pi}(A) := \left\{ x : x = \max(A; \succ) \text{ for some } \succ \in \Pi \right\}$$

in each menu  $A$ .

We represent the observed choice behavior of the DM by a data set  $(\Sigma, f)$ , where  $\Sigma \subseteq \mathcal{X}$  and  $f(A)$  is the collection of alternatives that the DM chooses in  $A \in \Sigma$ . We say that  $(\Sigma, f)$  is rationalizable by multiple preferences if there exists a set  $\Pi$  of strict preferences such that

$$f_{\Pi}(A) = f(A)$$

for all  $A \in \Sigma$ . For notational simplicity, we denote by  $g(A) := A \setminus f(A)$  the collection of alternatives that are not chosen in  $A \in \Sigma$ . In what follows, we identify necessary and sufficient conditions under which  $(\Sigma, f)$  is rationalizable by multiple preferences by applying our results in Section 3.3.

For a given data set  $(\Sigma, f)$ , let us first consider the following data set of menu preferences

$$\mathcal{O} := \{(f(A), g(A))\}_{A \in \Sigma},$$

where all menu preferences are strict. It is easy to see that if  $(\Sigma, f)$  is rationalizable by multiple preferences, then there necessarily exists a strict preference  $\succ$  such that  $\max(A; \succ) \in f(A)$  for all  $A \in \Sigma$ . In other words,  $\mathcal{O}$  must be rationalizable by a strict

preference. However, this is not sufficient for rationalizability by multiple preferences, as illustrated by the following example.

**Example 3.6.1** Let  $X = \{x, y, z\}$ . Consider the data set  $(\Sigma, f)$  consisting of the following three observations:<sup>23</sup>

- (i)  $f(\{x, y\}) = x$ ;
- (ii)  $f(\{y, z\}) = y$ ;
- (iii)  $f(\{x, y, z\}) = \{x, z\}$ .

Consider the data set of menu preferences  $\mathcal{O} = \{(x, y), (y, z), (\{x, z\}, y)\}$ . It follows from Corollary 3.3.12 that  $\mathcal{O}$  is rationalizable by a strict preference (for example, by the strict preference  $x \succ y \succ z$ ). However,  $(\Sigma, f)$  is not rationalizable by multiple preferences. To see this, suppose to the contrary that  $(\Sigma, f)$  is rationalizable by multiple preferences. By definition, there exists a set  $\Pi$  of strict preferences such that  $f_\Pi(A) = f(A)$  for all  $A \in \Sigma$ . The first observation  $f(\{x, y\}) = x$  reveals that  $x \succ y$  for all  $\succ \in \Pi$ , and the second observation  $f(\{y, z\}) = y$  reveals that  $y \succ z$  for all  $\succ \in \Pi$ . By transitivity, it must be that  $x \succ z$  for all  $\succ \in \Pi$ , which contradicts with the third observation that  $f(\{x, y, z\}) = \{x, z\}$ .

We now present a characterization of the multiple preferences model as follows. By Corollary 3.3.12, the data set of menu preferences  $\mathcal{O} = \{(f(A), g(A))\}_{A \in \Sigma}$  is rationalizable by a strict preference if and only if for all  $\emptyset \neq \Sigma' \subseteq \Sigma$ ,

$$\bigcup_{A \in \Sigma'} f(A) \not\subseteq \bigcup_{A \in \Sigma'} g(A).$$

---

<sup>23</sup>We abuse the notation by suppressing the set delimiters, e.g., writing  $x$  rather than  $\{x\}$ .

As illustrated by Example 3.6.1, this is not sufficient. We strengthen this condition as follows: for any nonempty  $\Sigma' \subseteq \Sigma$  and  $B \in \Sigma$ ,

$$\left( \bigcup_{A \in \Sigma'} f(A) \setminus \bigcup_{A \in \Sigma'} g(A) \right) \subseteq B \Rightarrow f(B) \cap \left( \bigcup_{A \in \Sigma'} g(A) \right) = \emptyset.^{24}$$

To understand the condition, note that it is plainly a necessary condition. Indeed, suppose that the data set is rationalizable by multiple preferences. If an alternative is not chosen in  $A$ , then under each preference, it is dominated by some alternative chosen in  $A$ . Hence, any alternative in  $g(A)$  is not chosen in a menu that contains  $f(A)$ . This implies that the set of choices made in  $\bigcup_{A \in \Sigma'} A$  does not contain any alternative in  $\bigcup_{A \in \Sigma'} g(A)$ . Furthermore, since  $f(A)$  contains all maximal alternatives under each preference in  $A$ , choices made in  $\bigcup_{A \in \Sigma'} A$  should be a subset of  $\bigcup_{A \in \Sigma'} f(A)$ , and thus a subset of  $\bigcup_{A \in \Sigma'} f(A) \setminus \bigcup_{A \in \Sigma'} g(A)$ . By a similar argument, any alternative in  $\bigcup_{A \in \Sigma'} g(A)$  is not chosen in any menu containing  $\bigcup_{A \in \Sigma'} f(A) \setminus \bigcup_{A \in \Sigma'} g(A)$ , as stated by the condition. Theorem 3.6.2 shows that this new condition is both necessary and sufficient for the rationalizability of a data set by multiple preferences.

From a practical perspective, we provide a divide-and-conquer approach to test whether a give data set  $(\Sigma, f)$  is rationalizable by multiple preferences. Rationalizability by multiple preferences requires that for each  $A \in \Sigma$  and  $x \in f(A)$ , there exist a strict preference such that  $x$  is the maximal element in  $A$  according to this strict preference. Furthermore, the maximal element in any other set  $A' \in \Sigma$  according to this strict preference must lie in  $f(A')$ . Thus, for each  $A \in \Sigma$  and  $x \in f(A)$ , the following data set of menu preferences

$$\mathcal{O}_{A,x} := (x, A \setminus x) \cup \left\{ (f(A'), A' \setminus f(A')) \right\}_{A' \in \Sigma, A' \neq A},$$

---

<sup>24</sup>To see that this is a strengthening of the previous condition, suppose that there exists some nonempty  $\Sigma' \subseteq \Sigma$  such that  $\bigcup_{A \in \Sigma'} f(A) \subseteq \bigcup_{A \in \Sigma'} g(A)$ . Fix an arbitrary nonempty set  $B \in \Sigma'$ . Since  $\bigcup_{A \in \Sigma'} f(A) \setminus \bigcup_{A \in \Sigma'} g(A) = \emptyset \subseteq B$ , it follows from the new condition that  $f(B) \cap \left( \bigcup_{A \in \Sigma'} g(A) \right) = \emptyset$ . We have arrived at a contradiction, since  $f(B) \subseteq \bigcup_{A \in \Sigma'} f(A) \subseteq \bigcup_{A \in \Sigma'} g(A)$ .

where all menu preferences are strict, must be rationalizable by a strict preference. Let

$$\mathfrak{D} := \{\mathcal{O}_{A,x}\}_{A \in \Sigma, x \in f(A)}.$$

A necessary condition for the data set  $(\Sigma, f)$  to be rationalizable by multiple preferences is that each  $\mathcal{O}_{A,x}$  in  $\mathfrak{D}$  is rationalizable by a strict preference. Theorem 3.6.2 below shows that the converse is also true. Recall that it is straightforward to check whether a data set of menu preferences is rationalizable by a strict preference (see Section 3.3.3). Thus, it is straightforward to check whether a data set  $(\Sigma, f)$  is rationalizable by multiple preferences using the divide-and-conquer approach.

**Theorem 3.6.2** *The following statements are equivalent:*

- (i)  $(\Sigma, f)$  is rationalizable by multiple preferences.
- (ii) For any nonempty  $\Sigma' \subseteq \Sigma$  and  $B \in \Sigma$ ,

$$\left( \bigcup_{A \in \Sigma'} f(A) \setminus \bigcup_{A \in \Sigma'} g(A) \right) \subseteq B \Rightarrow f(B) \cap \left( \bigcup_{A \in \Sigma'} g(A) \right) = \emptyset.$$

- (iii) Each  $\mathcal{O}_{A,x}$  in  $\mathfrak{D}$  is rationalizable by a strict preference.

Theorem 3.6.2 holds for any data set  $(\Sigma, f)$ . In the special case of complete data, that is, when  $\Sigma = \mathcal{X}$ , Aizerman and Malishevski (1981) show that the multiple preferences model could be characterized by the following two axioms:

**Chernoff:**  $A \subseteq B \Rightarrow f(B) \cap A \subseteq f(A)$  for all  $A, B \in \mathcal{X}$ .

**Aizerman:**  $f(B) \subseteq A \subseteq B \Rightarrow f(A) \subseteq f(B)$  for all  $A, B \in \mathcal{X}$ .

In words, the Chernoff axiom says that a best choice in some set is still best if the set shrinks. The Aizerman axiom says that deleting from a given set some choices outside the choice set cannot make new choices chosen. We establish this result as a corollary of Theorem 3.6.2 by showing that our characterization for any data set

(complete or incomplete) follows from the Chernoff axiom and the Aizerman axiom in the case of complete data.

**Corollary 3.6.3 (Aizerman and Malishevski (1981))** *The data set  $(\mathcal{X}, f)$  is rationalizable by multiple preferences if and only if it satisfies the Chernoff axiom and the Aizerman axiom.*

### 3.7 Application: Minimax Regret

The decision criterion of minimax regret is first suggested in Savage (1951)’s reading of Wald (1950) to model a DM who anticipates regret and thus incorporates in her choice the desire to minimize the worst-case regret. In this section, we investigate the observable restrictions of the minimax regret model by applying our results in Section 3.3.

We adopt the standard framework of the state space that has been discussed by, for example, Kreps (1979) and Dekel, Lipman, and Rustichini (2001). We consider any nonempty finite set  $X$ . Let  $U$  denote a state space with a typical element  $u$ . To economize on notation, we also write  $u$  to denote the utility function of the DM associated with the state  $u$ . The regret of choosing  $x$  relative to  $y$  is  $u(y) - u(x)$  if the state is  $u$ , the worst-case regret of choosing  $x$  relative to  $y$  is

$$\phi_U(x, y) := \max_{u \in U} \{u(y) - u(x)\},$$

and the worst-case regret of choosing  $x$  in a menu  $A$  is

$$\max_{y \in A} \phi_U(x, y).$$

The DM who uses the minimax regret decision criterion chooses all alternatives that generate the lowest worst-case regret. That is, the DM chooses

$$f_U(A) = \arg \min_{x \in A} \left\{ \max_{y \in A} \phi_U(x, y) \right\}$$

in each menu  $A$ .<sup>25</sup>

We represent the observed choice behavior of the DM by  $(\Sigma, f)$ , where  $\Sigma \subseteq \mathcal{X}$  and  $f(A)$  is the choice of the DM in  $A \in \Sigma$ . For the sake of simplicity, we assume that  $f$  is a choice function.<sup>26</sup> We say that  $(\Sigma, f)$  is rationalizable under the minimax regret model if there exists a finite set  $U$  of utility functions such that

$$f(A) = \arg \min_{x \in A} \left\{ \max_{y \in A} \phi_U(x, y) \right\} \quad (3.3)$$

for all  $A \in \Sigma$ .

Given the flexibility to construct the set of utility functions, the readers might wonder, whether there are any observable restrictions of the minimax regret model. To get this out of the way, let us first present a data set  $(\Sigma, f)$  that is not rationalizable under the minimax regret model in Example 3.7.1. We shall also use this data set in Example 3.7.4 (Continued) to illustrate the approach that we establish in this section.

**Example 3.7.1** Let  $X = \{x, y, z, w\}$ . Consider a data set  $(\Sigma, f)$  consisting of the following three observations:  $f(X) = x$ ,  $f(X \setminus z) = y$ , and  $f(X \setminus w) = y$ .

Suppose that  $(\Sigma, f)$  is rationalizable under the minimax regret model, say, by the set  $U$  of utility functions. In what follows, we show that (1)  $f(X) = x$  and  $f(X \setminus z) = y$  imply that  $\phi_U(y, z) > \phi_U(y, w)$ ; and (2)  $f(X) = x$  and  $f(X \setminus w) = y$

---

<sup>25</sup>Our minimax regret model differs from that in Milnor (1954) and Stoye (2011), who axiomatize the preference ordering represented by (the negative of)

$$\max_{s \in \mathcal{S}} \left\{ \max_{g \in M} u \circ g(s) - u \circ f(s) \right\},$$

where  $f$  and  $g$  are acts in a menu  $M$ ,  $\mathcal{S}$  is a state space, and  $u$  is an expected utility functional.

<sup>26</sup>The case of a choice correspondence can be accommodated using similar logic.

imply that  $\phi_U(y, w) > \phi_U(y, z)$ . Thus, we have a contradiction. Since  $f(X) = x$ ,  $x$  generates a lower worst-case regret in  $X$  than  $y$  does. It follows that the worst-case regret of  $x$  in  $X \setminus z$  is lower than the worst-case regret of  $y$  in  $X$ . Since  $f(X \setminus z) = y$ , the worst-case regret of  $y$  in  $X \setminus z$  is lower than the worst-case regret of  $x$  in  $X \setminus z$ . Therefore, the worst-case regret of  $y$  in  $X$  and  $X \setminus z$  are not the same. This is possible only if the worst-case regret of  $y$  relative to  $z$  is higher than the worst-case regret of  $y$  relative to  $w$ . Similarly, since  $f(X) = x$  and  $f(X \setminus w) = y$ , we can conclude that the worst-case regret of  $y$  relative to  $w$  is higher than the worst-case regret of  $y$  relative to  $z$ .

We present a lemma that simplifies our analysis below. This lemma shows that, without loss of generality, we can work with the following choice function:

$$f(A) = \arg \min_{x \in A} \left\{ \max_{y \in A \setminus x} \phi_U(x, y) \right\} \quad (3.4)$$

for all  $A \in \Sigma$ . The proof is trivial and purely algebraic, and hence omitted.

**Lemma 3.7.2** *There exists a finite set  $U$  of utility function such that (3.3) holds for all  $A \in \Sigma$  if and only if (3.4) holds for all  $A \in \Sigma$  under the same  $U$ .*

By Lemma 3.7.2, the rationalizability of a data set  $(\Sigma, f)$  under the minimax regret model is equivalent to the existence of a finite set  $U$  of utility functions such that

$$\max_{y \in A \setminus f(A)} \phi_U(f(A), y) < \max_{y \in A \setminus x} \phi_U(x, y)$$

for all  $A \in \Sigma$  and  $x \in A \setminus f(A)$ . For notational simplicity, throughout the rest of this section, we shall write  $(x, y)$  rather than  $\phi_U(x, y)$  to denote the worst-case regret of  $x$  relative to  $y$ . This should not cause any confusion.

Interestingly, the rationalizability of a data set under the minimax regret model is related to menu preferences. For any  $(\Sigma, f)$ , we construct its corresponding data

set  $\bar{\mathcal{O}}$  as follows: for each  $A \in \Sigma$  and  $x \in A$  with  $x \neq f(A)$ , let  $(A_i, B_i) \in \bar{\mathcal{O}}$  where

$$\bar{A}_i = x \times (A \setminus x) \text{ and } \bar{B}_i = f(A) \times (A \setminus f(A)).$$

The interpretation of the data point  $(\bar{A}_i, \bar{B}_i)$  is that the menu  $\bar{A}_i$  is strictly better than  $\bar{B}_i$  in the numeric order, which captures that alternative  $x$  which is not chosen in  $A$  generates a higher worst-case regret in  $A$  than  $f(A)$  does. Theorem 3.7.3 below makes the formal link between the rationalizability of a data set under the minimax regret model and the rationalizability of menu preferences.

**Theorem 3.7.3** *The data set  $(\Sigma, f)$  is rationalizable under the minimax regret model if and only if the corresponding data set  $\bar{\mathcal{O}}$  is rationalizable by a strict preference.*

Recall that it is straightforward to check whether a data set of menu preferences is rationalizable by a strict preference (see Section 3.3.3). Thus, it is straightforward to check whether a data set  $(\Sigma, f)$  is rationalizable under the minimax regret model. We now revisit Example 3.7.1 to illustrate how to use Theorem 3.7.3 to show that the data set is not rationalizable under the minimax regret model.

**Example 3.7.4 (Continued)** We construct the corresponding data set  $\bar{\mathcal{O}}$  as follows:

$$\begin{aligned} \bar{A}_1 &= \{(y, x), (y, z), (y, w)\}, & \bar{B}_1 &= \{(x, y), (x, z), (x, w)\}; \\ \bar{A}_2 &= \{(z, x), (z, y), (z, w)\}, & \bar{B}_2 &= \{(x, y), (x, z), (x, w)\}; \\ \bar{A}_3 &= \{(w, x), (w, y), (w, z)\}, & \bar{B}_3 &= \{(x, y), (x, z), (x, w)\}; \\ \bar{A}_4 &= \{(x, y), (x, w)\}, & \bar{B}_4 &= \{(y, x), (y, w)\}; \\ \bar{A}_5 &= \{(w, x), (w, y)\}, & \bar{B}_5 &= \{(y, x), (y, w)\}; \\ \bar{A}_6 &= \{(x, y), (x, z)\}, & \bar{B}_6 &= \{(y, x), (y, z)\}; \\ \bar{A}_7 &= \{(z, x), (z, y)\}, & \bar{B}_7 &= \{(y, x), (y, z)\}. \end{aligned}$$

By Corollary 3.3.12,  $\bar{\mathcal{O}}$  is not rationalizable by a strict preference. In particular,  $(\bar{A}_1 \cup \bar{A}_4 \cup \bar{A}_6) \subseteq (\bar{B}_1 \cup \bar{B}_4 \cup \bar{B}_6)$ . By Theorem 3.7.3,  $(\Sigma, f)$  is not rationalizable under the minimax regret model.

# Appendix A

## Appendices for Chapter 1

### A.1 Proof of Proposition 1.2.5

Let  $V \in \mathcal{V}$  be represented by  $\mathcal{P}$ . Consider  $f^*$  and  $g^*$  satisfying the conditions stated in the proposition. We have

$$\begin{aligned} V(g^*) - V(f^*) &= \mathbb{E}_{\mathcal{P}}(g^*) - \mathbb{E}_{\mathcal{P}}(f^*) \\ &= \mathbb{E}_{\mathcal{P}}(g^*) - \mathbb{E}_{p^*}(f^*) \\ &\leq \mathbb{E}_{p^*}(g^*) - \mathbb{E}_{p^*}(f^*) \\ &= p^*(s, \theta)(g^*(s, \theta) - f^*(s, \theta)) \\ &\leq \max_{p \in \mathcal{P}} p(s, \theta)(g^*(s, \theta) - f^*(s, \theta)), \end{aligned}$$

where  $p^* \in \mathcal{P}$  minimizes the expectation of  $f^*$ . Other inequalities are obvious. Let  $\mu_\theta$  be the CML posterior when signal  $\theta$  is observed. We have

$$\begin{aligned} V_\theta(g^*|\theta) - V_\theta(f^*|\theta) &= \mu_\theta(s)(g^*(s, \theta) - f^*(s, \theta)) \\ &= \frac{\max_{p \in \mathcal{P}} p(s, \theta)}{\sum_{s' \in S} (\max_{p \in \mathcal{P}} p(s', \theta))} (g^*(s, \theta) - f^*(s, \theta)) \\ &\geq \max_{p \in \mathcal{P}} p(s, \theta)(g^*(s, \theta) - f^*(s, \theta)). \end{aligned}$$

The last inequality holds since  $\sum_{s' \in S} (\max_{p \in \mathcal{P}} p(s', \theta)) \leq \sum_{s' \in S} p'(s' \times \Theta) \leq 1$  for each  $p' \in \mathcal{P}$ . The proposition is thus shown.

## A.2 Proof of Theorem 1.3.1

For any extended evaluation function  $V \in \mathcal{V}$  that is represented by  $\mathcal{P}$ , it is obvious that  $\theta \in \Theta_V$  if and only if there exists  $p \in \mathcal{P}$  and  $s \in S$  such that  $p(s, \theta) > 0$ .

Consider a nonempty and finite subset  $E$  of  $S \times \Theta$ . Let  $\Sigma(E)$  be the algebra generated by all subsets of  $E$  in  $S \times \Theta$ . For any set of probability distributions  $\mathcal{P}^*$  over  $S \times \Theta$ , we let  $\mathcal{P}^*|_{\Sigma(E)}$  denote the set of probability distributions over  $\Sigma(E)$  induced by  $\mathcal{P}^*$ . For any two extended evaluation functions  $V$  and  $V'$  that are represented by  $\mathcal{P}$  and  $\mathcal{P}'$  respectively, it is clear that  $V$  and  $V'$  agree on  $E$  if and only if  $\mathcal{P}|_{\Sigma(E)} = \mathcal{P}'|_{\Sigma(E)}$ . We proceed to the proof of the theorem.

**(Necessity.)** For axiom IIS, consider  $V, V' \in \mathcal{V}$ . Let  $V$  be represented by  $\mathcal{P}$  and  $V'$  by  $\mathcal{P}'$ . If  $V$  and  $V'$  agree on  $S \times \theta$ , then we have  $\mathcal{P}|_{\Sigma(S \times \theta)} = \mathcal{P}'|_{\Sigma(S \times \theta)}$ . Therefore,  $\theta \in \Theta_V$  implies that  $\theta \in \Theta_{V'}$ . Moreover, for each  $s \in S$ ,  $\max_{p \in \mathcal{P}} p(s, \theta) = \max_{p' \in \mathcal{P}'} p'(s, \theta)$ . By the formula of the CML posterior (1.1), we have  $V_\theta = V'_\theta$ . Axiom MI is shown similarly. Axiom ISU is shown by Proposition 1.2.5.

**(Sufficiency.)** Assume that axioms IIS, MI and ISU all hold. Through out the proof of the sufficiency part, assume the payoff space to be  $\mathbb{R}$ . This is without loss of generality since for any payoff  $y$  not in  $\mathbb{K}$ , we can pick  $x$  in the interior of  $\mathbb{K}$  and take the convex combination  $\alpha x + (1 - \alpha)y$  of  $x$  and  $y$  such that  $\alpha x + (1 - \alpha)y$  is in  $\mathbb{K}$ .

**Lemma A.2.1** *Let  $V \in \mathcal{V}$  be represented by  $\mathcal{P} = (\mu, \{c^t(\cdot|s)_{s \in S}\}_{t \in T})$  and  $V_\theta$  represented by  $\mathcal{Q}_\theta$ , where  $\theta \in \Theta_V$ . If*

$$\sum_{s \in S} \mu(s) \max_{t \in T} c^t(\theta|s) = 1, \quad (\text{A.1})$$

then  $\mathcal{Q}_\theta = \{\mu\}$ .

**Proof.** Fix some  $s \in S$  with  $\mu(s) > 0$ . By condition (A.1),  $\max_{t \in T} c^t(\theta|s) = 1$ .

Consider an extended act  $f^*$  satisfying that

$$\begin{aligned} f^*(s, \theta) &= 0, \\ f^*(s, \theta') &= 1, \forall \theta' \in \Theta \setminus \theta, \\ f^*(s', \theta') &= 0, \forall (s', \theta') \in (S \times \Theta) \setminus (s \times \Theta). \end{aligned}$$

For such an extended act, we have

$$V(f^*) = \mu(s)(1 - \max_{t \in T} c^t(\theta|s)) = 0 = V_\theta(f^*|\theta).$$

Consider an extended act  $g^*$  satisfying that

$$\begin{aligned} g^*(s, \theta) &= \frac{1}{2}, \\ g^*(s, \theta') &= 1, \forall \theta' \in \Theta \setminus \theta, \\ g^*(s', \theta') &= 0, \forall (s', \theta') \in (S \times \Theta) \setminus (s \times \Theta). \end{aligned}$$

We have

$$V(g^*) = \frac{1}{2}\mu(s), \text{ and } V_\theta(g^*|\theta) = \frac{1}{2} \min_{\mu^* \in \mathcal{Q}_\theta} \mu^*(s).$$

For extended acts  $f^*$  and  $g^*$ , axiom ISU implies that

$$\min_{\mu^* \in \mathcal{Q}_\theta} \mu^*(s) \geq \mu(s).$$

Since this holds for each  $s \in S$  with  $\mu(s) > 0$ , we have for each  $\mu^* \in \mathcal{Q}_\theta$  and each  $s \in S$ ,  $\mu^*(s) = \mu(s)$ . That is,  $\mathcal{Q}_\theta = \{\mu\}$ . ■

Now, consider an arbitrary extended evaluation function  $V$  that is represented by  $\mathcal{P} = (\mu, \{c^t(\cdot|s)_{s \in S}\}_{t \in T})$ . Let  $\theta$  be a  $V$  non-null signal and  $V_\theta$  be represented by  $\mathcal{Q}_\theta$ . Let  $\mu_\theta$  be the CML posterior of  $\mathcal{P}$ . We want to show that  $\mathcal{Q}_\theta = \mu_\theta$ .

First, consider an extended evaluation function  $V'$ . Let  $V'$  be represented by  $\mathcal{P}' = (\mu', \{d^t(\cdot|s)_{s \in S}\}_{t \in T})$  such that

- (i)  $\mu' = \mu_\theta$ ,
- (ii)  $\mu'(s)d^t(s, \theta) = \mu(s)c^t(s, \theta)$  for all  $t \in T$  and  $s \in S$ .

Since  $\mu_\theta(s) \geq \mu(s)c^t(s, \theta)$  for each  $t \in T$  and  $s \in S$ ,  $d^t(s, \theta)$  can be well-defined. Clearly,  $\mathcal{P}'|_{\Sigma(S \times \theta)} = \mathcal{P}|_{\Sigma(S \times \theta)}$ . Therefore,  $V$  and  $V'$  agree on  $S \times \theta$ . By axiom IIS,  $V_\theta = V'_\theta$ . We can ensure that  $\max_{t \in T} d^t(\theta|s) = \max_{t \in T} d^t(\theta|s')$  for each  $s, s' \in S$ . Let this constant number be  $\alpha$ , which is in  $(0, 1]$ .

Next, consider two extended evaluation functions  $V''$  and  $V^n$ .  $V''$  is represented by  $\mathcal{P}'' = (\mu'', \{e^t(\cdot|s)_{s \in S}\}_{t \in T})$  such that

- (i)  $\mu'' = \mu_\theta$ ,
- (ii)  $e^t(\theta|s) = \frac{d^t(\theta|s)}{\alpha}$  for all  $s \in S$  and  $t \in T$ .

Let  $V^n$  satisfy that  $\theta \notin V_\theta^n$ . Clearly,  $\alpha V'' + (1 - \alpha)V^n$  and  $V'$  agree on  $S \times \theta$ . Since  $\theta$  is  $V^n$  null, by axiom IIS and axiom MI,  $V_\theta'' = V'_\theta$ . That is,  $V_\theta'' = V_\theta$ . By the definition of  $\alpha$ ,  $\max_{t \in T} e^t(\theta|s) = 1$  for each  $s \in S$ . By Lemma A.2.1,  $V_\theta''$  is represented by  $\mu_\theta$ , i.e.,  $\mathcal{Q}_\theta = \mu_\theta$ . This completes the proof of the theorem.

### A.3 Proof of Theorem 1.3.4

**(Necessity.)** Suppose that the DM believes the possible set of conditional distributions is  $\{(c^t(\cdot|s))_{s \in S}\}_{t \in T}$  which satisfies conditions (1.3) and (1.4). For each  $s \in S$  and each  $\theta \in \Theta^*$ ,

$$\frac{\mu_\theta(s)}{\mu(s)} \left( \max_{s' \in S'} \frac{\mu_\theta(s')}{\mu(s')} \right)^{-1} = \frac{\max_{t \in T} c^t(\theta|s)}{\max_{s' \in S} (\max_{t \in T} c^t(\theta|s'))} \geq \max_{t \in T} c^t(\theta|s).$$

Therefore,

$$\sum_{\theta \in \Theta^*} \frac{\mu_\theta(s)}{\mu(s)} \left( \max_{s' \in S'} \frac{\mu_\theta(s')}{\mu(s')} \right)^{-1} \geq \sum_{\theta \in \Theta^*} \left( \max_{t \in T} c^t(\theta|s) \right) \geq 1.$$

This shows the necessity part.

**(Sufficiency.)** Suppose that the belief profile  $(\mu, \{\mu_\theta\}_{\theta \in \Theta^*})$  satisfies that for each  $s \in S$ ,

$$\sum_{\theta \in \Theta^*} \frac{\mu_\theta(s)}{\mu(s)} \left( \max_{s' \in S'} \frac{\mu_\theta(s')}{\mu(s')} \right)^{-1} \geq 1. \quad (\text{A.2})$$

We construct the finite set of conditional distributions. Define  $\{c^\theta(\cdot|s)_{s \in S}\}_{\theta \in \Theta^*}$ , where each  $c^\theta$  is labeled by a signal  $\theta \in \Theta^*$ . Let  $S_\theta := \arg \max_{s \in S} (\mu_\theta(s)/\mu(s))$ . For each  $s \in S$  and  $\theta \in \Theta^*$ , let

$$c^\theta(\theta|s) = \frac{\mu_\theta(s)}{\mu(s)} \left( \max_{s' \in S'} \frac{\mu_\theta(s')}{\mu(s')} \right)^{-1}.$$

Obviously,  $c^\theta(\theta|s) \leq 1$  for each  $s \in S$  and  $c^\theta(\theta|s) = 1$  for each  $s \in S_\theta$ . By condition (A.2), we have

$$\sum_{\theta \in \Theta^*} c^\theta(\theta|s) \geq 1$$

for each  $s \in S$ . Hence, we can find non-negative numbers  $\{c^\theta(\theta'|s)\}_{\theta' \in \Theta^* \setminus \theta}$  for each  $\theta \in \Theta^*$  and each  $s \in S$  such that

$$\sum_{\theta' \in \Theta^*} c^\theta(\theta'|s) = 1,$$

and

$$c^\theta(\theta'|s) \leq c^{\theta'}(\theta'|s).$$

Since  $\{c^{\theta'}(\theta|s)\}_{\theta' \in \Theta^*}$  is maximized at  $\theta' = \theta$  for each  $s$  and each  $\theta$ , we have the desired posterior  $\mu_\theta$  for each  $\theta \in \Theta^*$ .

## A.4 Proof of Theorem 1.4.2

If there exists  $a^* \in \bigcap_{s \in S} \arg \max_{a \in A^*} v(s, a)$ , let  $a^s = a^*$  for each  $s \in S$ . Otherwise, for each  $s \in S$ , pick  $a^s \in \arg \max_{a \in A^*} v(s, a)$ . Pick  $\mu^s \in \Delta^o(S)$  such that  $a^s \in \mathcal{BR}(\mu^s)$ .  $\{\mu^s\}_{s \in S}$  are the target posteriors. For each  $\mu^s$ , pick a vector  $(\lambda_{s'}^s)_{s' \in S} \in (0, 1)^S$  such that

$$\frac{\mu(s')\lambda_{s'}^s}{\mu(s'')\lambda_{s''}^s} = \frac{\mu^s(s')}{\mu^s(s'')}$$

for each  $s', s'' \in S$ . Fix the vector  $(\lambda_{s'}^s)_{s' \in S}$  for each  $s \in S$ . For any  $\{r^s\}_{s \in S} \subseteq (0, 1)$ , let  $n(\{r^s\}_{s \in S})$  be an integer satisfying that

$$\max_{s' \in S, s'' \in S} \frac{1 - r^{s'}\lambda_{s''}^{s'}}{n(\{r^s\}_{s \in S}) - 1} < \min_{s' \in S, s'' \in S} r^{s'}\lambda_{s''}^{s'}. \quad (\text{A.3})$$

For the given  $\epsilon > 0$ , pick  $\{r^s\}_{s \in S} \subseteq (0, 1)$  such that

$$\forall s \in S, \sum_{s' \in S} \mu(s') (r^s \lambda_{s'}^s v(s', a^s) + (1 - r^s \lambda_{s'}^s) v(s', a^{s'})) = l^*, \quad (\text{A.4})$$

where  $l^*$  satisfies that

$$l^* > \sum_{s \in S} \left( \mu(s) \max_{a \in A^*} v(s, a) \right) - \epsilon. \quad (\text{A.5})$$

In the case  $a^s = a^*$ , we have

$$l^* = \sum_{s \in S} \left( \mu(s) \max_{a \in A^*} v(s, a) \right),$$

which implies that  $\{r^s\}_{s \in S}$  exist. For the case where  $\bigcap_{s \in S} \arg \max_{a \in A^*} v(s, a) = \emptyset$ , since as  $r^s$  converges to zero,

$$\sum_{s' \in S} \mu(s') (r^s \lambda_{s'}^s v(s', a^s) + (1 - r^s \lambda_{s'}^s) v(s', a^{s'}))$$

is strictly lower than but converges to

$$\sum_{s' \in S} \mu(s') v(s', a^{s'}) = \sum_{s \in S} \left( \mu(s) \max_{a \in A^*} v(s, a) \right),$$

the numbers  $\{r^s\}_{s \in S}$  that satisfy the desired conditions exist. Fix the numbers  $\{r^s\}_{s \in S}$ .

We proceed to construct the information structure. Let  $N = n(\{r^s\}_{s \in S})$ . Let  $S = \{s_1, \dots, s_m\}$ . Let  $\Theta = \{\theta_{j,l}\}_{j \in \{1, \dots, m\}, l \in \{1, \dots, N\}}$ . Let there be  $m \cdot N$  signal generating systems. Each signal generating system is denoted by  $c^{j,l}$  for some  $j \in \{1, \dots, m\}$  and  $l \in \{1, \dots, N\}$ . For each  $c^{j,l}$ , let

$$\begin{aligned} c^{j,l}(\theta_{j,l}|s) &= r^{s_j} \lambda_s^{s_j}, \forall s \in S, \\ c^{j,l}(\theta_{j',l'}|s) &= \frac{1 - r^{s_j} \lambda_s^{s_j}}{N - 1}, \text{ if } s = s_{j'} \text{ and } l \neq l', \text{ and} \\ c^{j,l}(\theta_{j',l'}|s) &= 0 \text{ otherwise.} \end{aligned}$$

By condition (A.3), for each  $\theta_{j,l}$ , its conditional maximum probability on each  $s \in S$  is achieved uniquely by  $c^{j,l}$  and equal to  $r^{s_j} \lambda_s^{s_j}$ . Hence, the CML posterior at signal  $\theta_{j,l}$  is  $\mu^{s_j}$ .

Consider  $G = (\Theta, \{c^{j,l}(\cdot|s)_{s \in S}\}_{j \in \{1, \dots, m\}, l \in \{1, \dots, N\}}, \{a^{\theta_{j,l}}\}_{\theta_{j,l} \in \Theta})$ , where  $a^{\theta_{j,l}} = a^{s_j}$ . We argue  $G$  satisfies the desired conditions of the theorem. Note that  $a^{\theta_{j,l}}$  is indeed the optimal action of the agent given the CML posterior  $\mu^{s_j}$ . For each signal generating system  $c^{j,l}$ , conditional on state  $s_j$ , action  $a^{s_j}$  is taken by the agent with probability 1. Conditional on state  $s \neq s_j$ , action  $a^s$  is taken by the agent with probability  $1 - \frac{1 - r^{s_j} \lambda_s^{s_j}}{N - 1}$ , and action  $a^{s_j}$  is taken with probability  $\frac{1 - r^{s_j} \lambda_s^{s_j}}{N - 1}$ . By conditions (A.4) and (A.5), we know that  $G$  satisfies the conditions in the statement of the theorem.

# Appendix B

## Appendices for Chapter 2

### B.1 Proof of Theorem 2.3.1

Throughout the proof, for any  $A \in \mathcal{M}$ , any  $u \in \mathcal{U}$  and any nonempty and closed subset  $\mathcal{V} \subseteq \mathcal{U}$ , we define

$$t(u, \mathcal{V}, A) := \arg \max_{p \in c(\mathcal{V}, A)} u(p).$$

Define function  $\Gamma : \mathcal{U} \times \Delta(X) \rightarrow \mathbb{R}$  as  $\Gamma(u, p) := u(p)$ . One can verify that  $\Gamma$  is continuous on the product space  $\mathcal{U} \times \Delta(X)$ . We first show that  $t(u, \mathcal{V}, A)$  is nonempty.

**Lemma B.1.1** *For any  $A \in \mathcal{M}$ , any  $u \in \mathcal{U}$  and any nonempty and closed  $\mathcal{V} \subseteq \mathcal{U}$ ,  $t(u, \mathcal{V}, A)$  is nonempty.*

**Proof.** It suffices to show that  $c(\mathcal{V}, A)$  is closed. Consider a sequence  $\{p_n\}_{n=1}^{\infty} \subseteq c(\mathcal{V}, A)$  converging to  $p$ . By the definition of  $c(\mathcal{V}, A)$ , we can find a sequence  $\{v_n\}_{n=1}^{\infty} \subseteq \mathcal{V}$  such that  $\Gamma(v_n, p_n) \geq \Gamma(v_n, q)$  for any  $n$  and any  $q \in A$ . By compactness of  $\mathcal{V}$ , we can find a subsequence  $\{v_{n_k}\}_{k=1}^{\infty}$  of  $\{v_n\}_{n=1}^{\infty}$  such that  $v_{n_k}$  converges to some  $v \in \mathcal{V}$ . By continuity of  $\Gamma$ , we know  $\Gamma(v, p) \geq \Gamma(v, q)$  for any  $q \in A$ . As a result,  $p \in c(\mathcal{V}, A)$ . This implies that  $c(\mathcal{V}, A)$  is closed. ■

**(Necessity.)** Suppose that  $\succsim$  is represented by  $(u, \mathcal{V})$ , where  $u \in \mathcal{U}$  and  $\mathcal{V} \subseteq \mathcal{U} \setminus \{o\}$  is nonempty and compact. Define  $V : \mathcal{M} \rightarrow \mathbb{R}$  such that

$$V(A) = \max_{p \in c(\mathcal{V}, A)} u(p).$$

By Lemma B.1.1,  $V(A)$  is well-defined. Hence, axiom WO holds. It is also easy to verify axioms IDD and WC.

For axiom IR, note that for any  $A \in \mathcal{M}$ ,  $c(\mathcal{V}, A) \subseteq c(\mathcal{V}, \text{conv}(A)) \subseteq \text{conv}(c(\mathcal{V}, A))$ . Since  $u$  is linear,  $V(\text{conv}(A)) = V(A)$ . That is,  $\text{conv}(A) \sim A$ . For axiom IIC, note that for any  $B \in \mathcal{M}$ , if  $A^\downarrow \subseteq B \subseteq A$  for some  $A \in \mathcal{M}$ , then  $c(\mathcal{V}, B) = c(\mathcal{V}, A)$ . Therefore,  $A \sim B$ . For axiom PSB, consider  $A, B \in \mathcal{M}$  such that  $A \succsim B$ . We have

$$\max_{p \in c(\mathcal{V}, A)} u(p) \geq \max_{q \in c(\mathcal{V}, B)} u(q).$$

Note that  $c(\mathcal{V}, A \cup B) \subseteq c(\mathcal{V}, A) \cup c(\mathcal{V}, B)$ . Therefore,

$$\max_{r \in c(\mathcal{V}, A \cup B)} u(r) \leq \max \left\{ \max_{p \in c(\mathcal{V}, A)} u(p), \max_{q \in c(\mathcal{V}, B)} u(q) \right\} = \max_{p \in c(\mathcal{V}, A)} u(p).$$

This indicates that  $A \succsim A \cup B$ .

Finally, for axiom PERU, consider  $\alpha \in (0, 1)$  and  $A, B \in \mathcal{M}$  such that  $A \succsim B$ . We show that  $A \succsim \alpha A + (1 - \alpha)B$ . It is easy to see that

$$c(\mathcal{V}, \alpha A + (1 - \alpha)B) \subseteq \alpha c(\mathcal{V}, A) + (1 - \alpha)c(\mathcal{V}, B). \quad (\text{B.1})$$

By condition (B.1), for any  $p \in t(u, \mathcal{V}, \alpha A + (1 - \alpha)B)$ , there exist  $q_1 \in c(\mathcal{V}, A)$ ,  $q_2 \in c(\mathcal{V}, B)$  such that  $p = \alpha q_1 + (1 - \alpha)q_2$ . Hence,

$$\begin{aligned} V(\alpha A + (1 - \alpha)B) &= u(p) = \alpha u(q_1) + (1 - \alpha)u(q_2) \\ &\leq \alpha V(A) + (1 - \alpha)V(B) \\ &\leq V(A). \end{aligned}$$

**(Sufficiency.)** We first show that both the current preference and the maximal set of future preferences can be identified by restricting  $\succsim$  to finite menus. We then derive the representation for finite menus. Finally, we extend the representation to compact menus. Denote the collection of nonempty finite menus as  $\mathcal{M}^F$ . Throughout the proof, assume that all axioms hold.

**Lemma B.1.2** *There exists a unique  $u \in \mathcal{U}$  such that for any  $p, q \in \Delta(X)$ ,  $p \succsim q \Leftrightarrow u(p) \geq u(q)$ .*

**Proof.**  $\succsim$  is a weak order over singleton menus. By axiom IDD,  $p \succ q$  implies  $\alpha p + (1 - \alpha)r \succ \alpha q + (1 - \alpha)r$  for any  $\alpha \in (0, 1)$  and  $r \in \Delta(X)$ . Axiom WC implies that for any lotteries  $p, q$  and  $r$ , if  $p \succ r \succ q$ , then there exists  $\alpha, \beta \in (0, 1)$  such that  $\alpha p + (1 - \alpha)q \succ r \succ \beta p + (1 - \beta)q$ . These three conditions are standard for an expected utility representation. The expected utility is unique up to a positive affine transformation. Thus, we can uniquely identify such an expected utility in  $\mathcal{U}$ . ■

**Lemma B.1.3** *For any  $A \in \mathcal{M}$  and  $\{B_i\}_{i=1}^n \subseteq \mathcal{M}$ , if  $A = \cup_{i=1}^n B_i$ , then for some  $i \in \{1, \dots, n\}$ ,  $B_i \succsim A$ .*

**Proof.** It is a simple implication of axiom PSB. ■

**Lemma B.1.4** *For any  $A \in \mathcal{M}$ , there exists  $p^*, q^* \in A$  such that  $p^* \succsim A \succsim q^*$ .*

**Proof.** Suppose that  $A \succ p$  for all  $p \in A$ . By Lemma B.1.3, for any collection  $\{B_i\}_{i=1}^n \subseteq \mathcal{M}$  satisfying that  $\cup_{i=1}^n B_i = A$ , we have  $B_i \succsim A$  for some  $i \in \{1, \dots, n\}$ . We

can choose the collection  $\{B_i\}_{i=1}^n$  such that the diameter of each menu  $B_i$  is smaller than  $\frac{1}{k}$  for some positive integer  $k$ , i.e.,

$$\max_{p \in B_i, p' \in B_i} d(p, p') \leq \frac{1}{k},$$

for each  $i \in \{1, \dots, n\}$ . By this, we can find a convergent sequence of menus  $\{B^k\}_{k=1}^{+\infty}$  such that  $B^k \subseteq A$  and  $B^k \succsim A$  for each  $k$  and the diameters of  $\{B^k\}_{k=1}^{+\infty}$  converge to 0. Let  $p^*$  be the limiting singleton menu of the sequence of menus. Obviously,  $p^* \in A$ . By axiom WC, we know  $p^* \succsim A$ . Hence, it is impossible that  $A \succ p$  for all  $p \in A$ .

Next, suppose that  $q \succ A$  for all  $q \in A$ . Therefore,  $A^\downarrow = \emptyset$ . By axiom IIC, we have  $q \sim A$  for each  $q \in A$  since  $A^\downarrow \subseteq q$ . Hence, it is impossible that  $q \succ A$  for all  $q \in A$ . ■

**Lemma B.1.5** *For any  $A \in \mathcal{M}$ , there exists  $p \in A^\downarrow$  such that*

- (i) *if  $B \in \mathcal{M}$  and  $p \in B \subseteq A$ , then  $B \succsim A$ ;*
- (ii) *if  $B \in \mathcal{M}$  and  $p \in B \subseteq A^\downarrow$ , then  $B \sim A$ ;*
- (iii)  *$A \sim A^\downarrow \sim p$ .*

**Proof.** Consider any menu  $A$ . By Lemma B.1.4,  $A^\downarrow$  is not empty. We first prove statement (i) by contradiction. If the statement is not true, for any  $p \in A^\downarrow$ , we can find  $B \in \mathcal{M}$  such that  $p \in B \subseteq A$  and  $A \succ B$ . Hence, we can find for each  $p \in A^\downarrow$  some menu  $B_p$  such that  $p \in B_p \subseteq A$  and  $A \succ B_p$ . For each  $p$ , let  $B_p^o$  be a superset of  $B_p$  such that  $B_p^o$  is a subset of  $A$  and open in  $A$ . Let  $B_p^c$  be the closure of  $B_p^o$  in  $A$ . Obviously,  $B_p^c$  is also a well-defined menu. For each  $p \in A^\downarrow$ , we make  $d_h(B_p, B_p^c)$  small enough such that by axiom WC,  $A \succ B_p^c$ . Note that  $\{B_p^o\}_{p \in A^\downarrow}$  is an open cover of  $A^\downarrow$ . Therefore, we can find a finite set  $\{p_i\}_{i=1}^n \subseteq A^\downarrow$  such that  $\{B_{p_i}^o\}_{i=1}^n$  covers  $A^\downarrow$ . Obviously,  $\{B_{p_i}^c\}_{i=1}^n$  also covers  $A^\downarrow$ , i.e.,  $A^\downarrow \subseteq \cup_{i=1}^n B_{p_i}^c$ . By axiom IIC,  $A \sim \cup_{i=1}^n B_{p_i}^c$ . By Lemma B.1.3, there exists  $i \in \{1, \dots, n\}$  such that  $B_{p_i}^c \succsim A$ , which is a contradiction. Therefore, statement (i) holds.

For statement (ii), consider  $B$  such that  $p \in B \subseteq A^\downarrow$ . By statement (i), we know  $B \succsim A$ . By Lemma B.1.4, there exists  $p' \in B$  such that  $p' \succsim B$ , which implies  $A \succsim B$ . Hence,  $B \sim A$ . Statement (iii) is directly implied by statement (ii). ■

For any menu  $A$ , let  $\text{ext}(A)$  denote the set of extreme points of  $A$ , i.e.,  $p \in \text{ext}(A)$  if and only if  $p$  is not a convex combination of any two different points in  $A$ . We proceed to the next lemma.

**Lemma B.1.6** *For any  $A \in \mathcal{M}^F$ , there exists  $p \in \text{ext}(A)$  such that  $A \sim p$ .*

**Proof.** By axiom IR, for any  $A \in \mathcal{M}^F$ ,  $A \sim \text{conv}(A)$ . Since  $A$  is finite,  $\text{ext}(A)$  is also a well-defined menu and  $\text{conv}(\text{ext}(A)) = \text{conv}(A)$ . Therefore,  $A \sim \text{ext}(A)$ . By Lemma B.1.5, there exists  $p \in \text{ext}(A)$  such that  $p \sim \text{ext}(A) \sim A$ . The lemma is proved. ■

For the remaining part of the proof, let  $u \in \mathcal{U}$  be the current utility identified in Lemma B.1.2. When  $u = o$ , Lemma B.1.5 implies that for any  $A \in \mathcal{M}$ ,  $A \sim p$  for some  $p \in A$ . It implies that  $A \sim B$  for any two menus  $A$  and  $B$  since each singleton menu is indifferent. Therefore,  $(o, \mathcal{V})$  represents the menu preference  $\succsim$  for any nonempty and compact  $\mathcal{V} \subseteq \mathcal{U} \setminus \{o\}$ . In particular, the maximal set of future preferences is  $\mathcal{U} \setminus \{o\}$ . From now on, we focus on the nontrivial case where  $u \neq o$ . To proceed, we introduce some notations. For any  $A \in \mathcal{M}$  and any  $p \in A$ , define  $N(p, A) := \{v \in \mathcal{U} : v(p) \geq v(q), \forall q \in A\}$ .  $N(p, A)$  is the set of preferences over lotteries rationalizing the choice of  $p$  in  $A$ . We say that  $A$  dominates  $B$ , denoted by  $A \succ^* B$ , if  $\forall p \in A$  and  $\forall q \in B$ , we have  $p \succ q$ . The next lemma is a technical result.

**Lemma B.1.7** *Fix any collection of menus  $\{A_i\}_{i=0}^n \subseteq \mathcal{M}^F$  such that  $p_i \in A_i$  for each  $i \in \{0, \dots, n\}$  and  $N(p_0, A_0) \subseteq \cup_{i=1}^n N(p_i, A_i)$ . Let  $\{\alpha_i\}_{i=0}^n$  and  $\{q_i\}_{i=1}^n$  satisfy that  $\alpha_i > 0$  for each  $i \in \{0, \dots, n\}$ ,  $q_i \in A_i$  for each  $i \in \{1, \dots, n\}$  and  $\sum_{i=0}^n \alpha_i = 1$ . If  $\alpha_0 p_0 + \sum_{i=1}^n \alpha_i q_i$  is an extreme point of  $\sum_{i=0}^n \alpha_i A_i$ , then  $q_k = p_k$  for some  $k \in \{1, \dots, n\}$ .*

**Proof.** Suppose  $\alpha_0 p_0 + \sum_{i=1}^n \alpha_i q_i$  is an extreme point of the menu  $\sum_{i=0}^n \alpha_i A_i$ . Since the menu is finite, there is an expected utility  $v' \in \mathbb{R}^X$  such that  $v'(\alpha_0 p_0 + \sum_{i=1}^n \alpha_i q_i) > v'(r)$  for any distinct  $r \in \sum_{i=0}^n \alpha_i A_i$ .<sup>1</sup> There exists  $v \in \mathcal{U}$  such that  $v(\alpha_0 p_0 + \sum_{i=1}^n \alpha_i q_i) > v(r)$  for any distinct  $r \in \sum_{i=0}^n \alpha_i A_i$ . Therefore, we know  $v \in N(p_0, A_0)$ . By the fact that  $N(p_0, A_0) \subseteq \cup_{i=1}^n N(p_i, A_i)$ , we know that there is some  $k$  such that  $v \in N(p_k, A_k)$ . This implies that  $v(\alpha_0 p_0 + \sum_{i=1}^n \alpha_i q_i) = v(\alpha_0 p_0 + \sum_{i \neq k} \alpha_i q_i + \alpha_k p_k)$ . It happens only when  $\alpha_0 p_0 + \sum_{i=1}^n \alpha_i q_i = \alpha_0 p_0 + \sum_{i \neq k} \alpha_i q_i + \alpha_k p_k$ . Therefore, we conclude that for some  $k \in \{1, \dots, n\}$ ,  $q_k = p_k$ . ■

**Lemma B.1.8** *For any  $A \in \mathcal{M}$  and lottery  $p$ , if  $p \succ^* A$ , then for any  $\alpha \in (0, 1)$  and any lottery  $q$ , we have*

$$p \succ A \cup p \Leftrightarrow \alpha p + (1 - \alpha)q \succ \alpha A \cup p + (1 - \alpha)q.$$

**Proof.** By axiom IDD, the left-hand side implies the right-hand side. To see the other direction, first note that by Lemma B.1.5,  $p \succ^* A$  implies  $p \succsim A \cup p$ . Assume for some  $\alpha \in (0, 1)$ , we have  $p \sim A \cup p$  and  $\alpha p + (1 - \alpha)q \succ \alpha A \cup p + (1 - \alpha)q$ . By compactness of  $A$  and Lemma B.1.2, we can find  $r \in A$  such that  $u(r) \geq u(p')$  for any  $p' \in A$ . Since  $p \sim A \cup p$ , we know  $A \cup p \succ r$ , and thus  $\alpha A \cup p + (1 - \alpha)q \succ \alpha r + (1 - \alpha)q$ . Lemma B.1.5 implies that there exists some lottery  $l \in \alpha A \cup p + (1 - \alpha)q$  such that  $l \sim \alpha A \cup p + (1 - \alpha)q$ . Since  $\alpha A \cup p + (1 - \alpha)q \succ \alpha r + (1 - \alpha)q$ ,  $l$  can only be  $\alpha p + (1 - \alpha)q$ , which contradicts to the fact that  $\alpha p + (1 - \alpha)q \succ \alpha A \cup p + (1 - \alpha)q$ . The lemma is thus proved. ■

**Lemma B.1.9** *For any  $A, B \in \mathcal{M}^F$  and any two lotteries  $p, q$  such that  $p \succ^* A$  and  $q \succ^* B$ , if  $N(q, B \cup q) \subseteq N(p, A \cup p)$ , then  $p \succ A \cup p$  implies  $q \succ B \cup q$ .*

---

<sup>1</sup>See, for example, Theorem 2.3 in Bertsimas and Tsitsiklis (1997).

**Proof.** Consider  $p, q, A$  and  $B$  that satisfy the primitive conditions stated by the lemma. Suppose that  $p \succ A \cup p$  and  $q \not\succ B \cup q$ . We want to derive a contradiction. Since  $q \succ^* B$ , by Lemma B.1.5,  $q \not\succ B \cup q$  implies  $q \sim B \cup q$ .

First, we argue that we can find  $B^*$  and  $q^*$  such that  $q^* \succ^* B^*$ ,  $q^* \sim B^* \cup q^*$ ,  $N(q^*, B \cup q^*) \subseteq N(p, A \cup p)$ , and  $p \succ q^* \succ^* A$ . Choose  $r \in A$  such that  $r \succsim p'$  for each  $p' \in A$ . Take any  $\lambda \in (0, 1)$  and define  $w = \lambda r + (1 - \lambda)p$ . By Lemma B.1.2, we know  $p \succ w \succ^* A$ . Let  $\beta \in (0, 1)$  and define  $B^* := \beta w + (1 - \beta)B$  and  $q^* = \beta w + (1 - \beta)q$ . We require  $\beta$  to be close enough to 1 such that  $p \succ q^* \succ^* A$ . Thanks to Lemma B.1.8, we have  $q^* \sim B^* \cup q^*$ . Since  $B$  and  $q$  are transformed linearly to  $B^*$  and  $q^*$ , we know that  $N(q^*, B^* \cup q^*) = N(q, B \cup q)$  and  $q^* \succ^* B^*$ . Therefore, the desired conditions for  $B^*$  and  $q^*$  are satisfied.

Since  $p \succ A \cup p$ , by Lemma B.1.5, we can find  $p' \in A$  such that  $A \cup p \sim p'$ . Thus,  $B^* \cup q^* \sim q^* \succ A \cup p$ . By axiom PERU, we know for any  $\alpha \in (0, 1)$ ,

$$q^* \sim B^* \cup q^* \succsim \alpha B^* \cup q^* + (1 - \alpha)A \cup p. \quad (\text{B.2})$$

Consider  $l \in \Delta(X)$  such that  $q^* \succ l \succ^* B^*$ . We can find  $\alpha^* \in (0, 1)$  such that for any  $\alpha \in (\alpha^*, 1)$ ,

$$\alpha q^* + (1 - \alpha)A \cup p \succ^* l \succ^* \alpha B^* + (1 - \alpha)A \cup p. \quad (\text{B.3})$$

Since  $B^* \cup q^* \sim q^* \succ l$ , by axiom WC, there is some  $\alpha^{**} \in (0, 1)$  such that for any  $\alpha \in (\alpha^{**}, 1)$ ,

$$\alpha B^* \cup q^* + (1 - \alpha)A \cup p \succ l. \quad (\text{B.4})$$

For any  $\alpha \in (\max\{\alpha^*, \alpha^{**}\}, 1)$ , by Lemma B.1.6, there is an extreme point  $\hat{q}$  of  $\alpha B^* \cup q^* + (1 - \alpha)A \cup p$  such that  $\hat{q} \sim \alpha B^* \cup q^* + (1 - \alpha)A \cup p$ . Since  $N(q^*, B^* \cup q^*) \subseteq N(p, A \cup p)$ , Lemma B.1.7 implies that  $\hat{q} \in \alpha B^* + (1 - \alpha)A \cup p$  or  $\hat{q} = \alpha q^* + (1 - \alpha)p$ . Conditions (B.3) and (B.4) imply that  $\hat{q} = \alpha q^* + (1 - \alpha)p$ . Since  $p \succ q^*$ , condition (B.2) leads to a contradiction. The lemma is proved. ■

A immediate corollary of Lemma B.1.9 is that when  $p, q, A, B$  satisfy  $p \succ^* A, q \succ^* B$  and  $N(p, A \cup p) = N(q, B \cup q)$ , we have  $p \succ A \cup p \Leftrightarrow q \succ B \cup q$ . The next lemma generalizes this result.

**Lemma B.1.10** *Let  $\{A_i\}_{i=0}^n \subseteq \mathcal{M}^F$  and  $\{p_i\}_{i=0}^n \subseteq \Delta(X)$  satisfy that (i)  $p_i \succ^* A_i$  and  $p_i \succ A_i \cup p_i$  for each  $i \in \{1, \dots, n\}$ , and (ii)  $p_0 \succ^* A_0$ . If  $N(p_0, A_0 \cup p_0) \subseteq \cup_{i=1}^n N(p_i, A_i \cup p_i)$ , then  $p_0 \succ A_0 \cup p_0$ .*

**Proof.** For each  $i \in \{1, \dots, n\}$ , we can find  $A'_i$  such that

- (i)  $p_i \succ^* A'_i$ ,
- (ii)  $N(p_i, A'_i \cup p_i) = N(p_i, A_i \cup p_i)$ , and
- (iii)  $q_i, q'_i \in A'_i$  implies  $q_i \sim q'_i$ .

To construct  $A'_i$ , note that  $p_i \succ q$  for any  $q \in A_i$ . For any  $q \in A_i$ , consider the linear combination  $\alpha p_i + (1 - \alpha)q$  such that  $\alpha u(p_i) + (1 - \alpha)u(q) = \frac{1}{2}u(p_i) + \frac{1}{2} \max_{r \in A_i} u(r)$ .

One can easily verify that

$$\alpha = \frac{\frac{1}{2}u(p_i) + \frac{1}{2} \max_{r \in A_i} u(r) - u(q)}{u(p_i) - u(q)} \in (0, 1).$$

Therefore, let

$$\alpha_{p_i, q} := \frac{\frac{1}{2}u(p_i) + \frac{1}{2} \max_{r \in A_i} u(r) - u(q)}{u(p_i) - u(q)},$$

we can define

$$A'_i := \{q' : \exists q \in A_i, q' = \alpha_{p_i, q} p_i + (1 - \alpha_{p_i, q})q\}.$$

To show that  $A'_i$  satisfies the desired conditions, we just need to show  $N(p_i, A'_i \cup p_i) = N(p_i, A_i \cup p_i)$ . This is obvious since for any  $\alpha \in (0, 1)$ ,  $v(p_i) \geq v(q)$  if and only if  $v(p_i) \geq \alpha v(p_i) + (1 - \alpha)v(q)$ . By the construction of  $A'_i$ , we know the desired conditions are satisfied. By Lemma B.1.9, we know  $p_i \succ A'_i \cup p_i$  for  $i \in \{1, \dots, n\}$ . By a similar construction, it is without loss of generality to assume that for any  $i, j \in \{1, \dots, n\}$ ,  $p_i \sim p_j$  and  $q_i \sim q_j$  for any  $q_i \in A'_i$  and  $q_j \in A'_j$ .

Since  $p_i \succ p_i \cup A'_i$  for  $i \in \{1, \dots, n\}$ , by Lemma B.1.5, we know  $A'_i \cup p_i \sim A'_i \sim q_i$  for any  $q_i \in A'_i$ . Therefore,  $A'_i \cup p_i \sim A'_j \cup p_j$  for  $i, j \in \{1, \dots, n\}$ . By axiom PERU,  $A'_j \cup p_j \gtrsim \sum_{i=1}^n \frac{1}{n} A'_i \cup p_i$  for any  $j \in \{1, \dots, n\}$ . By Lemma B.1.5, we have  $A'_j \cup p_j \sim \sum_{i=1}^n \frac{1}{n} A'_i \cup p_i$  for any  $j \in \{1, \dots, n\}$ .

By a similar construction as in Lemma B.1.9, we can further assume that  $p_i \succ p_0 \sim q_i \succ^* A_0$  for any  $i \in \{1, \dots, n\}$  and  $q_i \in A'_i$ . Now suppose  $A_0 \cup p_0 \sim p_0$ . We want to derive a contradiction. Since  $A_0 \cup p_0 \sim p_0 \sim \sum_{i=1}^n \frac{1}{n} A'_i \cup p_i$ , by axiom PERU, for any  $\alpha \in (0, 1)$ , we have

$$A_0 \cup p_0 \gtrsim \alpha A_0 \cup p_0 + (1 - \alpha) \left( \sum_{i=1}^n \frac{1}{n} A'_i \cup p_i \right). \quad (\text{B.5})$$

By Lemma B.1.6, we can find an extreme point  $\hat{p}$  of  $\alpha A_0 \cup p_0 + (1 - \alpha) \left( \sum_{i=1}^n \frac{1}{n} A'_i \cup p_i \right)$  such that  $\hat{p} \sim \alpha A_0 \cup p_0 + (1 - \alpha) \left( \sum_{i=1}^n \frac{1}{n} A'_i \cup p_i \right)$ . By axiom WC, there exists  $\alpha^* \in (0, 1)$  such that for any  $\alpha \in (\alpha^*, 1)$ , we know  $\hat{p} \in \alpha p_0 + (1 - \alpha) \left( \sum_{i=1}^n \frac{1}{n} A'_i \cup p_i \right)$ . By condition (B.5) and that  $p_i \succ p_0$ , we have  $\hat{p} \in \alpha p_0 + (1 - \alpha) \sum_{i=1}^n \frac{1}{n} A'_i$ . This is impossible since by Lemma B.1.7,  $\alpha p_0 + (1 - \alpha) \sum_{i=1}^n \frac{1}{n} A'_i$  contains no extreme points. We are done. ■

Define  $\mathcal{T} := \{(A, p) : A \in \mathcal{M}^F, p \succ^* A, p \succ A \cup p\}$ . If  $\mathcal{T}$  is empty, let  $\mathcal{V} = \mathcal{U} \setminus \{o\}$ . Otherwise, define

$$\mathcal{V} = (\mathcal{U} \setminus \{o\}) \setminus \left( \bigcup_{(A, p) \in \mathcal{T}} N(p, A \cup p) \right).$$

$\mathcal{V}$  contains all the preferences that are potentially in the DM's set of future preferences. Define  $\hat{N}(p, A \cup p) := N(p, A \cup p) \setminus \{o\}$ . It is easy to see that  $\hat{N}(p, A \cup p)$  is closed (and thus compact) for each pair of  $(A, p) \in \mathcal{T}$ . Finally, define  $\hat{N}^o(p, A \cup p)$  to be the relative interior of  $\hat{N}(p, A \cup p)$  with respect to  $\mathcal{U}$ .

**Lemma B.1.11**  $\mathcal{V}$  is not empty.

**Proof.** We argue that  $-u \in \mathcal{V}$ . By the construction of  $\mathcal{V}$ , if  $\mathcal{T} = \emptyset$ , we must have  $-u \in \mathcal{V}$ . If not, consider any  $(A, p) \in \mathcal{T}$ .  $p \succ^* A$  implies that  $-u \notin N(p, A \cup p)$ . Therefore,  $\mathcal{V}$  is never empty. ■

**Lemma B.1.12** *For any  $p \in \Delta(X)$ , the set  $\{A \in \mathcal{M} : p \succ^* A, p \succ A \cup p\}$  is open.*

**Proof of Lemma B.1.12.** We show both  $\{A \in \mathcal{M} : p \succ^* A\}$  and  $\{A \in \mathcal{M} : p \succ A \cup p\}$  are open. Continuity of  $u$  ensures that  $\{A \in \mathcal{M} : p \succ^* A\}$  is open. The second part follows from axiom WC and the fact that  $d_h(A \cup p, B \cup p) \leq d_h(A, B)$ . ■

**Lemma B.1.13** *Suppose  $\mathcal{T} \neq \emptyset$ . For each  $(A, p) \in \mathcal{T}$ , there is a collection  $\{(A_i, p_i)\}_{i \in I} \subseteq \mathcal{T}$  such that  $\hat{N}(p, A \cup p) \subseteq \cup_{i \in I} \hat{N}^o(p_i, A_i \cup p_i)$ .*

**Proof.** It is without loss of generality to consider  $(A, p) \in \mathcal{T}$  such that  $A \cup p$  is in the relative interior of  $\Delta(X)$  by Lemma B.1.9. For any  $v \in \hat{N}(p, A \cup p)$ , we have  $v(p) \geq v(q)$  for any  $q \in A$ . By the definition of  $\hat{N}(p, A \cup p)$ ,  $v$  is not constant over  $X$ . Thus, we can find  $\epsilon \in \mathbb{R}^X$  such that  $\sum_x \epsilon_x = 0$ ,  $\sum_x |\epsilon_x|$  being small enough and  $v \cdot \epsilon < 0$ . Consider  $A'$  such that  $A' := \{q' : \exists q \in A, q' = q + \epsilon\}$ . Let  $\sum_x |\epsilon_x|$  be smaller enough to ensure that  $A' \in \mathcal{M}^F$ . By Lemma B.1.12, we have  $p \succ^* A'$  and  $p \succ A' \cup p$ , i.e.  $(A', p) \in \mathcal{T}$ . Moreover, we know that  $v \in \hat{N}^o(p, A' \cup p)$  since  $v(p) - v(q') > 0$  for any  $q' \in A'$  and is bounded away from 0 uniformly. This finishes the proof of the lemma. ■

An immediate implication of Lemma B.1.13 is that  $\cup_{(A,p) \in \mathcal{T}} \hat{N}(p, A \cup p)$  is relatively open in  $\mathcal{U}$ . Thus, when  $\mathcal{T} \neq \emptyset$ ,  $\mathcal{V} = \mathcal{U} - \left[ \cup_{(A,p) \in \mathcal{T}} \hat{N}(p, A \cup p) \right] - \{o\}$  is compact. Lemma B.1.13 also implies the following result of finite covering.

**Lemma B.1.14** *Suppose  $\mathcal{T} \neq \emptyset$ . Consider  $A \in \mathcal{M}^F$  and a lottery  $p$  such that  $N(p, A \cup p) \subseteq \mathcal{U} \setminus \mathcal{V}$ , there is a finite collection  $\{(A_i, p_i)\}_{i=1}^n \subseteq \mathcal{T}$  such that  $N(p, A \cup p) \subseteq \cup_{i=1}^n N(p_i, A_i \cup p_i)$ .*

**Proof.** We just need to show that there is a finite collection  $\{(A_i, p_i)\}_{i=1}^n \subseteq \mathcal{T}$  such that  $\hat{N}(p, A \cup p) \subseteq \cup_{i=1}^n \hat{N}(p_i, A_i \cup p_i)$  since  $o \in N(p', A' \cup p')$  for any  $A', p'$ . By  $N(p, A \cup p) \subseteq \mathcal{U} \setminus \mathcal{V}$ , we know  $\hat{N}(p, A \cup p) \subseteq \cup_{(B,q) \in \mathcal{T}} \hat{N}(B, B \cup q) \subseteq \cup_{(B,q) \in \mathcal{T}} \hat{N}^o(B, B \cup q)$  by Lemma B.1.13. By the open covering theorem, we can find a finite collection  $\{(A_i, p_i)\}_{i=1}^n \subseteq \mathcal{T}$  such that  $\hat{N}(p, A \cup p) \subseteq \cup_{i=1}^n \hat{N}^o(p_i, A_i \cup p_i) \subseteq \cup_{i=1}^n \hat{N}(p_i, A_i \cup p_i)$ . This finishes the proof of the lemma. ■

**Lemma B.1.15** *Consider  $A \in \mathcal{M}^F$  and a lottery  $p$  such that  $p \succ^* A$ .  $p \succ A \cup p$  if and only if  $N(p, A \cup p) \cap \mathcal{V} = \emptyset$ .*

**Proof.** If  $N(p, A \cup p) \cap \mathcal{V} = \emptyset$ , we know  $\mathcal{T} \neq \emptyset$ . We have  $N(p, A \cup p) \subseteq \mathcal{U} \setminus \mathcal{V}$ . By Lemma B.1.14, there is a finite collection of  $\{(A_i, p_i)\}_{i=1}^n \subseteq \mathcal{T}$ , such that  $N(p, A \cup p) \subseteq \cup_{i=1}^n N(p_i, A_i \cup p_i)$ . By Lemma B.1.10, we know  $p \succ A \cup p$ . If  $N(p, A \cup p) \cap \mathcal{V} \neq \emptyset$ , either  $\mathcal{T} = \emptyset$  which directly implies that  $A \cup p \sim p$ , or by the construction of  $\mathcal{V}$ ,  $A \cup p \sim p$ . This finishes the proof of the lemma. ■

**Lemma B.1.16** *Consider a menu  $A \in \mathcal{M}^F$ . The following statements are true.*

- (i) *For any  $p \in A \setminus A^\downarrow$ ,  $N(p, A) \cap \mathcal{V} = \emptyset$ ;*
- (ii)  *$\exists p \in A$  such that  $N(p, A) \cap \mathcal{V} \neq \emptyset$  and  $A \sim p$ .*

**Proof.** We first show statement (i). Consider  $p \in A \setminus A^\downarrow$ . By Lemma B.1.5, we know  $p \succ A \sim A^\downarrow \cup p$ . Thus  $(A^\downarrow, p) \in \mathcal{T}$ . By Lemma B.1.15, we know that  $N(p, A^\downarrow \cup p) \cap \mathcal{V} = \emptyset$ . Therefore,  $N(p, A) \cap \mathcal{V} = \emptyset$  since  $N(p, A) \subseteq N(p, A^\downarrow \cup p)$ .

For statement (ii), consider first when  $\mathcal{T} = \emptyset$ . In this case,  $\mathcal{V} = \mathcal{U} \setminus \{o\}$ , we are done. When  $\mathcal{T} \neq \emptyset$ , choose  $p \in A$  such that  $p$  satisfies the condition stated in Lemma B.1.5. Consider the sets  $B, C$  such that  $B := \{q \in A : p \succ q\}$  and  $C := \{r \in A : p \sim r\}$ . First note that  $B \cup p \sim p$ . By Lemma B.1.15,  $N(p, B \cup p) \cap \mathcal{V} \neq \emptyset$ . This implies that there exists  $v \in \mathcal{V}$  such that  $v(p) \geq v(q)$  for any  $q \in B$ . By statement (i),  $v \notin N(q, A)$  for any  $q \in A \setminus A^\downarrow$ . It implies that  $\exists r \in C$  such that  $v(r) \geq v(q)$  for any  $q \in A$ . ■

In the remaining part of the proof, we extend the representation from finite menus to compact menus. We have shown that  $(u, \mathcal{V})$  represents the menu preference  $\succsim$  restricted on  $\mathcal{M}^F$ . To finish the proof of the theorem, we just need to show that we can extend the representation to  $\mathcal{M}$ .

Consider first when  $\mathcal{T} = \emptyset$ . We argue that for any  $A \in \mathcal{M}$ , it must be true that  $A \sim q$  for some  $q$  such that  $q \succsim r$  for any  $r \in A$ . Otherwise, by Lemma B.1.5, we know  $\exists A$  and for any  $p \in A \setminus A^\downarrow$ , we have  $p \succ A^\downarrow \cup p$ . By the compactness of  $A^\downarrow$ , for any  $\epsilon > 0$ , we can find a finite set  $A'$  such that  $d_h(A', A^\downarrow) < \epsilon$ . Let  $\epsilon$  be small enough. By Lemma B.1.12, we have  $p \succ A' \cup p$  and  $p \succ^* A'$ . Therefore,  $(A', p) \in \mathcal{T}$ , which is a contradiction. Thus, when  $\mathcal{T} = \emptyset$ , the construction that  $\mathcal{V} = \mathcal{U} \setminus \{o\}$  indeed ensures that  $(u, \mathcal{V})$  represents the menu preference  $\succsim$ .

From now on, assume  $\mathcal{T} \neq \emptyset$ . Consider  $A \in \mathcal{M}$ , we first show that for any  $p \in A$  such that  $p \succ A$ , we have  $N(p, A) \cap \mathcal{V} = \emptyset$ . By Lemma B.1.5, we have  $p \succ^* A^\downarrow$  and  $p \succ A^\downarrow \cup p$ . We want to show that  $N(p, A^\downarrow \cup p) \cap \mathcal{V} = \emptyset$ , and this immediately leads to  $N(p, A) \cap \mathcal{V} = \emptyset$ . We prove by contradiction. Assume  $v \in N(p, A^\downarrow \cup p) \cap \mathcal{V}$ . By Lemma B.1.12, there exists a positive number  $\delta$  such that for any  $B \in \mathcal{M}^F$  with  $d_h(B, A^\downarrow) < \delta$ , we have  $p \succ^* B$  and  $p \succ B \cup p$ . Pick such a finite menu  $B \subseteq A^\downarrow$ . However, since  $v \in N(p, A)$ , we know  $v \in N(p, A^\downarrow \cup p)$ , and thus  $v \in N(p, B \cup p)$ , which is a contradiction. Therefore,  $N(p, A) \cap \mathcal{V} = \emptyset$ .

Next, we show that  $\exists p \in A$  such that  $p \sim A$  and  $N(p, A) \cap \mathcal{V} \neq \emptyset$ . By Lemma B.1.5, we can find  $p \in A$  such that  $p \sim A$  and  $p \sim B$  for any  $B$  containing  $p$  with  $B \subseteq A^\downarrow$ . Define  $A^\sim := \{p' \in A : A \sim p'\}$ . We consider a sequence of finite menus  $\{B_n\}_{n=1}^\infty$  such that for each  $n$ , (i)  $B_n \subseteq A^\downarrow$ , (ii)  $B_n \subseteq B_{n+1}$ , (iii)  $p \succ^* B_n$ , and (iv)  $d_h(A^\downarrow, B_n \cup A^\sim)$  converges to 0. Define  $C_n := B_n \cup A^\sim$ . We first show that for each  $n$ , there is some  $v_n \in \mathcal{V}$  such that  $c(\{v_n\}, C_n) \cap A^\sim \neq \emptyset$ . To see this, note that for each  $n$ , Lemma B.1.15 implies that  $\exists v_n \in \mathcal{V}$ , such that  $\Gamma(v_n, p) \geq \Gamma(v_n, q)$  for each  $q \in B_n$ . Thus,  $c(\{v_n\}, C_n) \cap A^\sim \neq \emptyset$ . For each  $n$ , select  $p_n \in c(\{v_n\}, C_n) \cap A^\sim$ .

Consider a subsequence  $\{n_k\}_{k=1}^\infty$  such that  $v_{n_k}$  converges to  $v^* \in \mathcal{V}$  and  $p_{n_k}$  converges to  $p^* \in A^\sim$ . For each lottery  $q$  in  $A^\downarrow$ , we can find a selection  $q_{n_k} \in C_{n_k}$  converging to  $q$ . The fact that  $\Gamma(v_{n_k}, p_{n_k}) \geq \Gamma(v_{n_k}, q_{n_k})$  implies that  $\Gamma(v^*, p^*) \geq \Gamma(v^*, q)$  for each  $q \in A^\downarrow$ . This indicates that  $c(\{v^*\}, A^\downarrow) \cap A^\sim \neq \emptyset$ . Since for any  $q \in A$  with  $q \succ A$ ,  $N(q, A) \cap \mathcal{V} = \emptyset$ , we conclude that  $c(\{v^*\}, A) \cap A^\sim \neq \emptyset$ . Therefore, there exists some  $p^* \in A^\sim$  such that  $N(p^*, A) \cap \mathcal{V} \neq \emptyset$ . The theorem is proved.

## B.2 Proof of Theorem 2.4.3

We prove the theorem through a sequence of lemmas.

**Lemma B.2.1** *For any  $u, v_1, v_2 \in \mathcal{U} \setminus \{o\}$ , if  $(\eta_1; \theta_1, w) \in \mathcal{D}_u(v_1)$ ,  $(\eta_2; \theta_2, w) \in \mathcal{D}_u(v_2)$  and  $\eta_1 \geq \eta_2$ , then for any  $A \in \mathcal{M}$ ,*

$$\max_{p \in c(\{v_1\}, A)} u(p) \geq \max_{q \in c(\{v_2\}, A)} u(q).$$

**Proof.** By the primitive conditions, we have  $v_1 = \eta_1 u + \theta_1 w$  and  $v_2 = \eta_2 u + \theta_2 w$ . Since  $u \cdot w = 0$ ,  $\eta_1^2 + \theta_1^2 = \eta_2^2 + \theta_2^2 = 1$ . Since  $\theta_1, \theta_2 \in [0, 1]$ , we know  $\theta_1 = \sqrt{1 - \eta_1^2}$  and  $\theta_2 = \sqrt{1 - \eta_2^2}$ . Since  $\eta_1 \geq \eta_2$ , we can find a positive number  $a > 0$  and a nonnegative number  $b \geq 0$  such that  $v_1 = av_2 + bu$ . Find  $p_1 \in c(\{v_1\}, A), p_2 \in c(\{v_2\}, A)$  such that  $u(p_1) = \max_{p \in c(\{v_1\}, A)} u(p)$  and  $u(p_2) = \max_{q \in c(\{v_2\}, A)} u(q)$ . Suppose to the contrary that  $u(p_2) > u(p_1)$ . We must have  $b = 0$ . Otherwise,  $b > 0$  and  $v_2(p_2) \geq v_2(p_1)$  imply  $v_1(p_2) > v_1(p_1)$ , which is a contradiction. Since  $b = 0$ , we know  $v_1 = v_2$ , and thus  $u(p_1) = u(p_2)$ , which is again a contradiction. ■

To proceed, for any OMP  $\succsim$ , recall that  $\mathcal{V}(\succsim) = \bigcup_{(u, v) \in \mathcal{R}(\succsim)} \mathcal{V}$  is the maximal set of future preferences. We note that for any  $u, v \in \mathcal{U} \setminus \{o\}$ ,  $v$  has a unique  $u$ -decomposition  $(\eta; \theta, w)$  when  $v \notin \{u, -u\}$ . To see this, note that  $\eta = u \cdot v$  and  $\theta = \sqrt{1 - \eta^2}$ . Since  $v \notin \{-u, u\}$ , we know  $|\eta| < 1$ . Thus,  $\theta > 0$ .  $w$  is pinned down by  $w = \frac{v - \eta u}{\theta}$ , which is unique. If  $v \in \{-u, u\}$ , for any  $w \in \mathcal{U} \setminus \{o\}$ , we can

find a  $u$ -decomposition  $(\eta; \theta, w)$  where  $\eta = 1$  or  $-1$  and  $\theta = 0$ . Moreover, for each  $\mathcal{V} \subseteq \mathcal{U} \setminus \{o\}$ , we define

$$\gamma(\mathcal{V}, u, w) = \{\eta : (\eta; \theta, w) \in \mathcal{D}_u(v) \text{ for some } \theta \in [0, 1], v \in \mathcal{V}\}.$$

For the following Lemma B.2.2, we maintain all the notations introduced here.

**Lemma B.2.2** *Consider an OMP  $\succsim$ . If  $(u, \mathcal{V}) \in \mathcal{R}(\succsim)$  and  $u \neq o$ , then for any  $w \in \mathcal{U} \setminus \{o\}$  such that  $u \cdot w = 0$ , we have*

$$\max_{\eta \in \gamma(\mathcal{V}, u, w)} \eta = \max_{\hat{\eta} \in \gamma(\mathcal{V}(\succsim), u, w)} \hat{\eta}.$$

**Proof.** We first show that for any nonempty compact  $\mathcal{V} \subseteq \mathcal{U} \setminus \{o\}$ , any  $u \neq o$  and any  $w \in \mathcal{U} \setminus \{o\}$  with  $u \cdot w = 0$ ,  $\max_{\eta \in \gamma(\mathcal{V}, u, w)} \eta$  is well-defined. For any sequence  $\{v_n\}_{n=1}^\infty$  such that  $v_n \in \mathcal{V}$  and  $(\eta_n; \theta_n, w) \in \mathcal{D}_u(v_n)$  for some  $\eta_n$  and  $\theta_n$ , we can find a convergent subsequence  $\{(\eta_{n_k}, \theta_{n_k})\}_{k=1}^\infty$  converging to  $(\eta, \theta)$ . By this, we know that  $(\eta; \theta, w) \in \mathcal{D}_u(\eta u + \theta w)$  and  $\eta u + \theta w \in \mathcal{V}$ . Thus, the maximum is achieved.

To proceed, consider any  $(u, \mathcal{V}) \in \mathcal{R}(\succsim)$ . We have  $\mathcal{V} \subseteq \mathcal{V}(\succsim)$ , and thus

$$\max_{\eta \in \gamma(\mathcal{V}, u, w)} \eta \leq \max_{\hat{\eta} \in \gamma(\mathcal{V}(\succsim), u, w)} \hat{\eta}.$$

First note that  $\gamma(\mathcal{V}(\succsim), u, w) = \{-1\}$  if and only if  $\mathcal{V}(\succsim) = \{-u\}$ . In this case,  $\mathcal{V} = \mathcal{V}(\succsim) = \{-u\}$ , and the lemma trivially holds.

Next, consider the case where  $1 \in \gamma(\mathcal{V}(\succsim), u, w)$  for some  $w \in \mathcal{W}_u$ . This happens if and only if  $u \in \mathcal{V}(\succsim)$ . We argue that  $u \in \mathcal{V}$ , and thus the lemma holds for this case. Suppose that  $u \notin \mathcal{V}$ , we know there is some positive number  $\epsilon \in (0, 1)$  such that

$$\max_{v \in \mathcal{V}} u \cdot v < 1 - \epsilon$$

by the compactness of  $\mathcal{V}$ . We construct a menu  $A$  such that the menu preference represented by  $(u, \mathcal{V})$  is different from the menu preference represented by  $(u, \mathcal{V}(\succsim))$  over menu  $A$  and the singleton menu  $l = (\frac{1}{|X|}, \dots, \frac{1}{|X|})$ . This leads to a contradiction. To see this, let

$$A := \{l\} \cup \{p : p = l - \beta u + \delta w, w \in \mathcal{W}_u\}$$

where  $\beta, \delta > 0$  are constants and close to 0 to ensure that each lottery in  $A$  is well-defined. It is not hard to verify that  $A$  is compact. We further require that

$$\frac{\delta}{\beta} > \frac{1 - \epsilon}{\sqrt{1 - (1 - \epsilon)^2}}.$$

Given menu  $A$ , since  $u \in \mathcal{V}(\succsim)$ , we know that  $A \sim l$  if the menu preference  $\succsim$  is represented by  $(u, \mathcal{V}(\succsim))$ . If the menu preference  $\succsim$  is represented by  $(u, \mathcal{V})$ , we argue that  $l \notin c(\mathcal{V}, A)$ , and thus  $l \succ A$ . To see this, consider any  $v \in \mathcal{V}$ . We have  $v(l) = 0$ . If (i)  $v = -u$ , then  $\max_{p \in A} v(p) = \beta u \cdot u = \beta > 0$ , and if (ii)  $v \neq -u$ ,  $v = \eta u + \theta w$  for some  $w \in \mathcal{W}_u$ , then  $\eta \in (-1, 1 - \epsilon)$  and  $\theta \neq 0$ . For case (ii), we know  $v(p) = -\eta\beta + \theta\delta$  for  $p = l - \beta u + \delta w$ . Since  $\theta > 0$  and  $\delta > 0$  and  $\beta > 0$ , if  $\eta < 0$ , then  $v(p) > 0$  trivially. If instead  $\eta \geq 0$ , then  $v(p) > -(1 - \epsilon)\beta + \sqrt{1 - (1 - \epsilon)^2}\delta > 0$ . The last inequality is by the assumption that  $\frac{\delta}{\beta} > \frac{1 - \epsilon}{\sqrt{1 - (1 - \epsilon)^2}}$ . Thus, in both cases,  $v(l) < v(p)$  for some  $p \in A$ . This shows that  $u \in \mathcal{V}$ .

Finally, we consider the case where  $\gamma(\mathcal{V}(\succsim), u, w) \neq \{-1\}$  and  $1 \notin \gamma(\mathcal{V}(\succsim), u, w)$  for each  $w \in \mathcal{W}_u$ . Suppose there is some  $w^* \in \mathcal{W}_u$  such that

$$\max_{\eta \in \gamma(\mathcal{V}, u, w^*)} \eta < \max_{\hat{\eta} \in \gamma(\mathcal{V}(\succsim), u, w^*)} \hat{\eta}.$$

Thus, there is some  $\eta \in (-1, 1)$  such that

$$\eta u + \sqrt{1 - \eta^2} w^* \in \mathcal{V}(\succsim)$$

and some  $\epsilon > 0$  close to 0 such that  $\forall (w, \hat{\eta}) \in \mathcal{W}_u \times (\eta - \epsilon, 1]$  with  $\|w - w^*\| < \epsilon$ , we have

$$\hat{\eta}u + \sqrt{1 - \hat{\eta}^2}w \notin \mathcal{V}, \quad (\text{B.6})$$

where  $\|\cdot\|$  denotes the sup-norm. For notation convenience, define  $f(x) := \sqrt{1 - x^2}$  for  $x \in [-1, 1]$ . Note that  $f(\eta) > 0$ .

Define menu

$$B = \{p : p = l + \delta w, w \in \mathcal{W}_u\}$$

where  $\delta > 0$  and is close enough to 0 to ensure each lottery of  $B$  is well-defined.

Define

$$q = l + \delta w^* + \alpha u + \beta w^*$$

where  $\alpha > 0$  and both  $\alpha$  and  $\beta$  are close enough to 0 so that  $q$  is well-defined. Define menu  $C = B \cup \{q\}$ . We will require certain properties over  $\delta, \alpha, \beta$  conditional on different cases. Recall that we assume that  $1 \notin \gamma(\mathcal{V}(\succsim), u, w)$  for each  $w \in \mathcal{W}_u$  and thus both  $\{u \cdot v'\}_{v' \in \mathcal{V}(\succsim)}$  and  $\{u \cdot v'\}_{v' \in \mathcal{V}}$  are bounded above by some constant  $c \in (0, 1)$ . Consider two cases.

**The first case.** If  $\eta > 0$ , we let  $\delta, \alpha, \beta$  satisfy that

$$\alpha\eta + \beta f(\eta) = 0 \quad (\text{B.7})$$

$$\delta - (\delta + \beta)(1 - \frac{\epsilon^2}{2}) > 0 \quad (\text{B.8})$$

$$f(c) \left( \delta - (\delta + \beta)(1 - \frac{\epsilon^2}{2}) \right) > \alpha. \quad (\text{B.9})$$

In this case, we know  $\beta < 0$  and as long as  $|\alpha|$  and  $|\beta|$  are small enough compared to  $\delta$ , the conditions are satisfied. We want to show that  $q \notin c(\mathcal{V}, C)$  and  $q \in c(\mathcal{V}(\succsim), C)$ , which leads to a contradiction.  $q \in c(\mathcal{V}(\succsim), C)$  is from the observation that under preference  $\eta u + f(\eta)w^*$ ,  $q$  is indifferent with lottery  $l + \delta w^*$  by condition (B.7) and

weakly better than  $l + \delta w$  for any  $w \in \mathcal{W}_u$ . To see  $q \notin c(\mathcal{V}, B)$ , we consider two situations.

**Situation (i)**  $v' \in \mathcal{V}$  such that  $v' = \eta' u + f(\eta') w'$  for  $w' \in \mathcal{W}_u$  and  $\|w' - w^*\| \geq \epsilon$ .

It is easy to see that  $w^* \cdot w' < 1 - \frac{\epsilon^2}{2}$ . Define  $q' = l + \delta w'$ . We have

$$v'(q') - v'(q) = f(\eta')(\delta - (\delta + \beta)w^* \cdot w') - \alpha\eta'.$$

When  $\eta' \geq 0$ , we have  $f(\eta') \in (f(c), 1]$  and thus

$$\begin{aligned} v'(q') - v'(q) &\geq f(\eta') \left( \delta - (\delta + \beta)(1 - \frac{\epsilon^2}{2}) \right) - \alpha\eta' \\ &> f(c) \left( \delta - (\delta + \beta)(1 - \frac{\epsilon^2}{2}) \right) - \alpha \\ &> 0. \end{aligned}$$

where the second inequality comes from condition (B.8) and the last inequality comes from condition (B.9). When  $\eta' < 0$ , condition (B.8) ensures that  $v'(q') - v'(q) > 0$  since  $-\alpha\eta'$  is strictly positive in this case.

**Situation (ii)**  $v' \in \mathcal{V}$  such that  $v' = \eta' u + f(\eta') w'$  for  $w' \in \mathcal{W}_u$  and  $\|w' - w^*\| < \epsilon$ .

By condition (B.6), we know  $\eta' \leq \eta - \epsilon$ . Define  $q'$  similarly. We have

$$\begin{aligned} v'(q') - v'(q) &= f(\eta')(\delta - (\delta + \beta)w^* \cdot w') - \alpha\eta' \\ &\geq -\beta f(\eta') - \alpha\eta' > 0. \end{aligned}$$

The last inequality is by condition (B.7) and that  $\eta' < \eta$ .

**The second case.** If  $\eta \leq 0$ , we let  $\delta, \alpha, \beta$  satisfy conditions (B.8), (B.9) and

$$\exists \epsilon' \in (0, \epsilon), \alpha\eta + \beta f(\eta) = \alpha\epsilon'. \quad (\text{B.10})$$

The only difference is now  $\beta > 0$ . Consider  $v' \in \mathcal{V}$  such that  $v' = \eta' u + f(\eta') w'$  for some  $w' \in \mathcal{W}_u$ . Let  $q' = l + \delta w'$ . We know

$$v'(q') - v'(q) = f(\eta')(\delta - (\delta + \beta)w^* \cdot w') - \alpha\eta'$$

When  $\eta' > 0$ , a similar argument as Case 1.1 implies that  $v'(q') > v'(q)$ . When  $\eta' \leq 0$ , again by condition (B.6), we know  $\eta' \leq \eta - \epsilon < \eta$ . Hence,

$$\begin{aligned} & f(\eta')(\delta - (\delta + \beta)w^* \cdot w') - \alpha\eta' \\ & \geq -\beta f(\eta') - \alpha\eta' \\ & = -\beta(f(\eta') - f(\eta)) - \alpha(\eta' - \eta) - \alpha\epsilon' \\ & \geq -\beta(f(\eta') - f(\eta)) + \alpha\epsilon - \alpha\epsilon' \\ & > \alpha\epsilon - \alpha\epsilon' \\ & > 0. \end{aligned}$$

The first equality is by condition (B.10). The last inequality is by  $\epsilon' < \epsilon$ . The lemma is thus proved. ■

By Lemma B.2.2, we know the sets of  $\succeq_u$ -undominated future preferences in  $\mathcal{V}$  and  $\mathcal{V}(\succsim)$  are the same. This completes the proof of the theorem.

### B.3 Proof of Theorem 2.4.6

By the proof of Theorem 2.3.1,  $\mathcal{V}(\succsim)$  is compact and  $v \notin \mathcal{V}(\succsim)$  if and only if there exists  $p$  and a finite menu  $A$  such that  $p \succ^* A$ ,  $p \succ A \cup p$  and  $v \in N(p, A \cup p)$ . We first show that  $\succsim_1$  is more optimistic than  $\succsim_2$  if and only if  $u_1 = u_2$  and  $\mathcal{V}(\succsim_2) \subseteq \mathcal{V}(\succsim_1)$ . Assume that  $\succsim_1$  is more optimistic than  $\succsim_2$ . We argue that  $p \sim_1 q$  implies  $p \sim_2 q$ . Suppose not, then we can find  $p, q$  such that  $p \sim_1 q$  and  $p \succ_2 q$ . However, since there

exist some  $p', q'$  such that  $p' \succ_1 q'$ , by taking  $\alpha \in (0, 1)$  close to 1, we have

$$\alpha q + (1 - \alpha)p' \succ_1 \alpha p + (1 - \alpha)q'$$

$$\alpha p + (1 - \alpha)q' \succ_2 \alpha q + (1 - \alpha)p'.$$

This violates the fact that  $\succsim_1$  is more optimistic than  $\succsim_2$ . Thus,  $p \sim_1 q$  implies  $p \sim_2 q$ , which further implies that  $p \succsim_1 q$  if and only if  $p \succsim_2 q$ . Consequently, the current preferences for  $\succsim_1$  and  $\succsim_2$  are the same. Denote it as  $u$ . To see that  $\mathcal{V}(\succsim_2) \subseteq \mathcal{V}(\succsim_1)$ , note that  $v \notin \mathcal{V}(\succsim_1)$  implies there exist  $p$  and  $A$  where  $p \succ_1^* A$ ,  $p \succ_1 A \cup p$  and  $v \in N(p, A \cup p)$ . Since  $\succsim_1$  is more optimistic, we know  $p \succ_2^* A$ ,  $p \succ_2 A \cup p$  and  $v \in N(p, A \cup p)$ . Therefore,  $v \notin \mathcal{V}(\succsim_2)$ , and thus  $\mathcal{V}(\succsim_2) \subseteq \mathcal{V}(\succsim_1)$ .

The inverse direction is obvious since  $\mathcal{V}(\succsim_2) \subseteq \mathcal{V}(\succsim_1)$  implies  $c(\mathcal{V}(\succsim_2), A) \subseteq c(\mathcal{V}(\succsim_1), A)$  for any menu  $A$ .

## B.4 Proof of Theorem 2.6.1

Let  $\succsim$  be represented by some  $(u, \mathcal{V})$ . The two cases correspond to  $u \in \mathcal{V}$  and  $\mathcal{V} = \{-u\}$  respectively. Sufficiency is trivial. We prove necessity by contradiction. Assume that  $\succsim$  is continuous and  $\mathcal{V}$  does not contain  $u$ , and contains some  $v \neq -u$ . Define  $\bar{\eta} = \max_{v \in \mathcal{V}} u \cdot v$ . We know  $\bar{\eta} \in (-1, 1)$ . Take  $v \in \mathcal{V}$  such that  $v = \bar{\eta}u + \theta w$  for some  $w \in \mathcal{U} \setminus \{o\}$  with  $u \cdot w = 0$ .

First suppose that  $\bar{\eta} > 0$ . Define menu  $A_{\alpha, \delta}$  as

$$A_{\alpha, \delta} = \left\{ \left( \frac{1}{|X|}, \dots, \frac{1}{|X|} \right) + \alpha v' : v' \in \mathcal{U} \setminus \{o\}, u \cdot v' \leq \bar{\eta} \right\} \cup \left\{ \left( \frac{1}{|X|}, \dots, \frac{1}{|X|} \right) + \delta u \right\}.$$

Compactness of  $A_{\alpha, \delta}$  is easy to verify.  $\alpha$  and  $\delta$  are taken to be positive and close enough to 0 to ensure that the menu is well-defined. When  $\alpha = \delta \bar{\eta} > 0$ , it is easy to see that  $\left( \frac{1}{|X|}, \dots, \frac{1}{|X|} \right) + \delta u$  is rationalized by  $v$  in  $A_{\alpha, \delta}$ . By taking any  $\epsilon \in (0, \delta)$ , within the menu

$A_{\alpha, \delta - \epsilon}, \left(\frac{1}{|X|}, \dots, \frac{1}{|X|}\right) + (\delta - \epsilon)u$  is never rationalized. Thus,  $A_{\alpha, \delta} \sim \left(\frac{1}{|X|}, \dots, \frac{1}{|X|}\right) + \delta u$  while  $A_{\alpha, \delta - \epsilon} \succsim \left(\frac{1}{|X|}, \dots, \frac{1}{|X|}\right) + \alpha \bar{\eta} u$ . Note that  $\delta = \frac{\alpha}{\bar{\eta}} > \alpha > \alpha \bar{\eta}$ . As a result, continuity is violated.

Next suppose that  $\bar{\eta} = 0$ . Consider the menu  $A_{\alpha, \delta}$  where  $\alpha = 0$  and  $\delta > 0$ . It is easy to verify that  $A_{0, \delta} \sim \left(\frac{1}{|X|}, \dots, \frac{1}{|X|}\right) + \delta u$  and  $A_{\alpha', \delta} \sim \left(\frac{1}{|X|}, \dots, \frac{1}{|X|}\right)$  for any positive  $\alpha'$ . Therefore, continuity is violated.

Finally, suppose that  $\bar{\eta} < 0$ . Consider the menu  $A_{\alpha, \delta}$  where  $\alpha < 0$  and  $\alpha = \delta \bar{\eta}$ . Similar to the proof of case  $\bar{\eta} > 0$ , increasing  $\delta$  a bit discontinuously decreases the utility of the menu.

## B.5 Proof of Theorem 2.6.2

If  $\succsim$  can be represented by  $(u, \{v\})$ , independence is easy to verify. If  $\succsim$  satisfies independence, then  $A \sim A'$  and  $B \sim B'$  implies  $\alpha A + (1 - \alpha)B \sim \alpha A' + (1 - \alpha)B'$  for any  $\alpha \in (0, 1)$ . We show that if  $\succsim$  cannot be represented by  $(u, \{v\})$  for some  $v$ , then independence is violated.  $\succsim$  cannot be represented by some  $(u, \{v\})$  if and only if  $\succsim$  is represented by some  $(u, \mathcal{V})$  such that (i)  $u \notin \mathcal{V}$  and (ii) there exists  $v_1, v_2 \in \mathcal{V}$  such that neither  $v_1$  nor  $v_2$  is more  $u$ -aligned than the other. Consider such  $v_1, v_2 \in \mathcal{V}$  where  $v_1 = \eta_1 u + \theta_1 w_1$  and  $v_2 = \eta_2 u + \theta_2 w_2$  such that  $w_1, w_2 \in \mathcal{W}_u$ . Without loss of generality, assume there exists no  $v \in \mathcal{V}$  which is more  $u$ -aligned than  $v_1$  or  $v_2$ . By our proof of Theorem 2.4.3, for any  $\epsilon > 0$ , we can find menus  $A_1$  and  $A_2$  with  $p_1 \in A_1$  and  $p_2 \in A_2$  such that (i)  $p_1 \succ p$  for any  $p \in A_1 \setminus \{p_1\}$ , (ii)  $p_2 \succ q$  for any  $q \in A_2 \setminus \{p_2\}$ , (iii)  $v \in \mathcal{V}$  rationalizes  $p_1$  in  $A_1$  only if  $d(v, v_1) \leq \epsilon$ , and (iv)  $v \in \mathcal{V}$  rationalizes  $p_2$  in  $A_2$  only if  $d(v, v_2) \leq \epsilon$ . Therefore, by taking  $\epsilon$  close enough to 0 such that there exists no  $v$  with  $d(v, v_1) \leq \epsilon$  and  $d(v, v_2) \leq \epsilon$ , we have for any  $\alpha \in (0, 1)$ , any lottery  $p$  rationalized by some  $v \in \mathcal{V}$  in  $\alpha A_1 + (1 - \alpha)A_2$  cannot be  $\alpha p_1 + (1 - \alpha)p_2$ . It

implies that  $\alpha p_1 + (1 - \alpha)p_2 \succ \alpha A_1 + (1 - \alpha)A_2$ , which contradicts independence. The theorem is thus proved.

## B.6 Equivalence to C18 with Finite Alternatives

We show the equivalence between our model and the model by C18 in this section when the alternative space is finite. Throughout this section, let  $Z$  be a nonempty and finite alternative space. A menu is a nonempty subset of  $Z$ . Let  $\mathcal{M}_Z$  be the set of menus. A utility function is a function mapping  $Z$  to  $\mathbb{R}$ . A menu preference is complete and transitive binary relation  $\succsim$  over  $\mathcal{M}_Z$ . For any menu  $A$  and any set of utility functions  $\mathcal{V}$ , we continue to use the notation  $c(\mathcal{V}, A)$  to denote the set of choices in  $A$  that are rationalized by  $\mathcal{V}$ .

Our model has a natural counterpart in this discrete setting. A menu preference  $\succsim$  is an OMP if there exists a tuple  $(u, \mathcal{V})$  where  $u$  is the DM's current utility function and  $\mathcal{V}$  is a finite set of future utility functions that are anticipated by the DM such that for any two menus  $A$  and  $B$ ,  $A \succsim B$  if and only if

$$\max_{z \in c(\mathcal{V}, A)} u(z) \geq \max_{z' \in c(\mathcal{V}, B)} u(z').$$

We use  $\succsim^{OMP}$  to denote an OMP.

The model proposed by C18 is called the planner-doer model with subjective commitment (PDSC), where the planner's preference over menus is denoted by  $\succsim^{PDSC}$ . The menu preference  $\succsim^{PDSC}$  can be characterized by a tuple  $(u, v, \mathcal{C})$  where  $u$  is the planner's utility function,  $v$  is the doer's utility function, and  $\mathcal{C}$  is a finite collection of nonempty subsets of  $Z$  that covers  $Z$ , in which each  $C \in \mathcal{C}$  is interpreted as one possible subjective commitment of the planner. For each given menu  $A$ , the planner can pick a commitment  $C$  such that the doer can only make choices from  $A \cap C$ .

Hence, for any two menus  $A$  and  $B$ ,  $A \succsim^{PDSC} B$  if and only if

$$\max_{C \in \mathcal{C}} \left( \max_{z \in c(\{v\}, A \cap C)} u(z) \right) \geq \max_{C' \in \mathcal{C}} \left( \max_{z' \in c(\{v\}, B \cap C')} u(z') \right).$$

We argue that a menu preference is an OMP if and only if it is a PDSC. To see this, consider  $\succsim^{OMP}$  and let it be characterized by  $(u, \mathcal{V})$ . Consider  $(u, v, \mathcal{C})$  defined as follows:

- (i)  $v = -u$ ;
- (ii)  $C \in \mathcal{C}$  if and only if there exists  $v' \in \mathcal{V}$  and  $k \in \mathbb{R}$  such that  $\{z : z \in C, v'(z) > k\} \subseteq C$  and  $|\{z : z \in C, v'(z) = k\}| = 1$ .

We argue that the menu preference  $\succsim^{PDSC}$  constructed above is equivalent to  $\succsim^{OMP}$ . For any menu  $A$ , since  $v = -u$ , the planner wants to make the commitment set  $C \cap A$  as small as possible. By condition (ii), the planner only considers  $C \in \mathcal{C}$  such that  $|C \cap A| = 1$ . For such a commitment set  $C$ , we can find some  $v' \in \mathcal{V}$  such that  $C \cap A \subseteq \arg \max_{z \in A} v'(z)$ . Obviously, we have

$$\max_{C \in \mathcal{C}} \left( \max_{z \in c(\{v\}, A \cap C)} u(z) \right) = \max_{z \in c(\mathcal{V}, A)} u(z).$$

Inversely, consider an arbitrary menu preference  $\succsim^{PDSC}$  that is characterized by  $(u, v, \mathcal{C})$ . We construct the menu preference  $\succsim^{OMP}$  that is characterized by  $(u, \mathcal{V})$  such that each  $v_C \in \mathcal{V}$  corresponds to one  $C \in \mathcal{C}$  and satisfies:

- (i)  $v_C(z) = v(z') > v(z'')$  for all  $z, z' \in C$  and  $z'' \in Z \setminus C$ ;
- (ii)  $v_C(z) \geq v_C(z')$  if and only if  $u(z) \leq u(z')$  for all  $z, z' \in Z \setminus C$ .

By the construction, if  $C \cap A \neq \emptyset$ , then  $c(\{v_C\}, A) = A \cap C$ . If  $C \cap A = \emptyset$ , then  $c(\{v_C\}, A) = \arg \min_{z \in A} u(z)$ . Obviously,  $\succsim^{OMP}$  is the same as  $\succsim^{PDSC}$ .

# Appendix C

## Appendices for Chapter 3

### C.1 Proof of Theorem 3.3.5

We have argued in the main text that Statement (i) implies Statement (ii). In what follows, we show that Statement (ii) implies Statement (iii) and that Statement (iii) implies Statement (i).

**(Statement (ii)  $\Rightarrow$  Statement (iii))** Suppose that  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$ . We shall explicitly provide a way of selecting  $x^t$  in  $A^t$  for each  $t \in T$  such that  $\{x^t\}_{t \in T}$  is a no-cycling selection under  $\succeq$ .

For ease of notation, let us denote by  $\mathcal{E}(T')$  the set of alternatives that are revealed to be dominated through the procedure of iterated exclusion of dominated observations, i.e.,

$$\mathcal{E}(T') := B(T')^{\downarrow\downarrow} \bigcup B((T' \cap S) \cup \Phi(T'))^{\downarrow}.$$

Since  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$ , for any nonempty  $T' \subseteq T$ ,  $\Phi(T')$  is strict subset of  $T'$ , which implies that  $A(T') \setminus \mathcal{E}(T') \neq \emptyset$ .

Let  $T_1 := T$  and  $S_1 := A(T_1) \setminus \mathcal{E}(T_1)$ . We proceed by induction. Suppose that we have constructed  $T_k$  and  $S_k$  for some  $k \geq 1$ . If  $T_k \neq \emptyset$ , we construct  $T_{k+1}$  and  $S_{k+1}$

as follows:

$$T_{k+1} := \Phi(T_k) = \{t \in T_k : A^t \subseteq \mathcal{E}(T_k)\}, \text{ and}$$

$$S_{k+1} := A(T_{k+1}) \setminus \mathcal{E}(T_{k+1}).$$

Since  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$ , if  $T_k \neq \emptyset$ , then  $T_{k+1} = \Phi(T_k)$  is a strict subset of  $T_k$  and  $S_k = A(T_k) \setminus \mathcal{E}(T_k) \neq \emptyset$ . The construction stops when  $T_{k^*} \neq \emptyset$  and  $T_{k^*+1} = \emptyset$  for some  $k^*$ .

We are now ready to select  $x^t$  in  $A^t$  for each  $t \in T$  such that  $\{x^t\}_{t \in T}$  is a no-cycling selection under  $\succeq$ . For each  $1 \leq k \leq k^*$ , let  $V_k := T_k \setminus T_{k+1}$  denote the collection of observations that are eliminated when constructing  $T_{k+1}$  from  $T_k$ . Clearly,  $\{V_k\}_{k=1}^{k^*}$  is a partition of  $T$ . By definition, for each  $k$  and each  $t \in V_k = T_k \setminus T_{k+1}$ , we have  $A^t \setminus \mathcal{E}(T_k) \neq \emptyset$  and hence  $A^t \cap S_k = A^t \cap (A(T_k) \setminus \mathcal{E}(T_k)) \neq \emptyset$ .

For each  $k$  and each  $t \in V_k = T_k \setminus T_{k+1}$ , we pick an arbitrary  $x^t \in A^t \cap S_k$ . We proceed to verify that the revealed preference relations defined on  $\{x^t\}_{t \in T}$  obey the no-cycling property. Let  $k(t)$  be the corresponding index  $k$  such that  $t \in V_k$ . It suffices to show that (1)  $x^t R x^{t'}$  implies that  $k(t) \leq k(t')$ ; and (2)  $x^t P x^{t'}$  implies that  $k(t) < k(t')$ . Suppose that  $x^t R x^{t'}$  but  $k(t) > k(t')$ . Then  $t \in \Phi(T_{k(t')})$  due to the construction of  $\{V_k\}_{k=1}^{k^*}$ . It follows that  $A^t \subseteq \mathcal{E}(T_{k(t')})$  and  $B^{t\downarrow} \subseteq \mathcal{E}(T_{k(t')})$ . Since  $x^t R x^{t'}$ , we have  $x^{t'} \in B^{t\downarrow} \subseteq \mathcal{E}(T_{k(t')})$ , which contradicts with  $x^{t'} \in S_{k(t')} = A(T_{k(t')}) \setminus \mathcal{E}(T_{k(t')})$ . Hence,  $x^t R x^{t'}$  implies  $k(t) \leq k(t')$ . Suppose that  $x^t P x^{t'}$  but  $k(t) \geq k(t')$ . If  $k(t) > k(t')$ , then we have the same contradiction as argue above. If  $k(t) = k(t') = k$  for some  $k$ , then both  $x^t$  and  $x^{t'}$  belong to  $S_k$ . Since  $S_k = A(T_k) \setminus \mathcal{E}(T_k)$  and  $B(T_k)^{\downarrow\downarrow} \cup B(T_k \cap S)^{\downarrow} \subseteq \mathcal{E}(T_k)$ , we have

$$x^t, x^{t'} \in S_k \subseteq A(T_k) \setminus (B(T_k)^{\downarrow\downarrow} \cup B(T_k \cap S)^{\downarrow}).$$

But this is impossible since  $x^t P x^{t'}$  implies that either  $x^{t'} \in B^{t\downarrow\downarrow}$  or  $t \in S$  and  $x^{t'} \in B^{t\downarrow}$ . Hence,  $x^t P x^{t'}$  implies  $k(t) < k(t')$ .

**(Statement (iii)  $\Rightarrow$  Statement (i))** Suppose that  $\mathcal{M}$  admits a no-cycling selection under  $\succeq$ ,  $\{x^t\}_{t \in T}$ . Let  $R^*$  be the binary relation on  $X$  where  $\hat{x} R^* \hat{y}$  if there is  $t \in T$  such that  $\hat{x} = x^t$  and  $\hat{y} \in B^t$ . Let  $\succsim^*$  be the transitive closure of the binary relation  $R^* \cup \succeq$ . By the Szpilrajn's extension theorem (see Szpilrajn (1930)),  $\succsim^*$  admits an extension  $\succsim$ .<sup>1</sup>

We claim that the preference  $\succsim$  has two properties: (a) it rationalizes the data set and (b) it extends  $\succeq$ . It follows from the construction that  $x^t \succsim B^t$  for all  $t \in T$ , so it remains to show that  $x^t \succ B^t$  for  $t \in S$ . To show (a), suppose to the contrary that for some  $t' \in S$ ,  $x^{t'} \sim y$  for some  $y \in B^{t'}$ . Given that  $\succsim$  extends  $\succsim^*$ , this can only occur if  $y \succsim^* x^{t'}$ . This means there is  $t'' \in T$  such that  $y \succeq x^{t''} \succsim^* x^{t'}$ . Therefore we obtain  $x^{t'} P x^{t''} \succsim^* x^{t'}$ , which is excluded by the no-cycling property. This completes the proof of (a). To show (b), note that  $x \succsim y$  if  $x \succeq y$  by construction, so it remains to show that  $x \succ y$  if  $x \triangleright y$ . Suppose instead we have  $x \triangleright y$  but  $x \sim y$ . This can only occur if  $y \succsim^* x$ . Since  $\succeq$  is a preorder, if  $x \triangleright y$  and  $y \succsim^* x$ , there must be  $t', t'' \in T$  such that  $y \succeq x^{t''} \succsim^* x^{t'} \succeq x$ . So we obtain  $x^{t'} \succeq x \triangleright y \succeq x^{t''}$ , which means that  $x^{t'} P x^{t''}$ , which is incompatible with  $x^{t''} \succsim x^{t'}$ , given the no-cycling property.

## C.2 Proof of Proposition 3.3.6

We prove the first statement below. The second statement is an immediate implication of the first statement. Fix a nonempty  $T' \subseteq T$  and  $\Phi \subsetneq T'$  such that for any  $t \in T' \setminus \Phi$ ,

$$A^t \not\subseteq B(T')^{\downarrow\downarrow} \bigcup B((T' \cap S) \cup \Phi)^{\downarrow}.$$

---

<sup>1</sup> This means that there is a complete preorder  $\succsim$  such that  $x \succsim y$  if  $x \succsim^* y$  and  $x \succ y$  if  $x \succ^* y$  where  $\succ$  is the asymmetric part of  $\succsim$  and  $\succ^*$  the asymmetric part of  $\succsim^*$ .

To show that  $\Phi(T') \subseteq \Phi$ , we proceed by induction. Obviously,

$$\Phi^0(T') = \emptyset \subseteq \Phi,$$

$$\Phi^1(T') = \{t \in T' : A^t \subseteq B(T')^{\downarrow\downarrow} \bigcup B(T' \cap S)^{\downarrow}\} \subseteq \Phi.$$

Suppose that we have  $\Phi^m(T') \subseteq \Phi$  for some  $m$ . It follows that

$$B(T')^{\downarrow\downarrow} \bigcup B((T' \cap S) \cup \Phi^m(T'))^{\downarrow} \subseteq B(T')^{\downarrow\downarrow} \bigcup B((T' \cap S) \cup \Phi)^{\downarrow}.$$

Thus,  $t \in T' \setminus \Phi$  implies  $t \in T' \setminus \Phi^{m+1}(T')$ , which further implies that  $\Phi^{m+1}(T') \subseteq \Phi$ .

By induction, we have  $\Phi(T') \subseteq \Phi$ .

### C.3 Proof of Proposition 3.3.7

By Theorem 3.3.5,  $\mathcal{M}$  is  $\succeq$ -rationalizable if and only if it satisfies the never-covered property under  $\succeq$ . Thus, it suffices to show that  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$  if and only if the ALGORITHM outputs  $\succeq$ -Rationalizable.

**(Only if-part)** If  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$ , then  $\Phi(T') \neq T'$  for any nonempty  $T' \subseteq T$ . Thus, Case (iii) never occurs when we run the ALGORITHM on this data set. Furthermore,  $T^k$  is strictly decreasing in  $k$  in the set inclusion sense, and  $T^{k^*} = \emptyset$  for some  $k^*$ . Therefore, the ALGORITHM outputs  $\succeq$ -Rationalizable.

**(If-part)** Conversely, suppose that the ALGORITHM outputs  $\succeq$ -Rationalizable. We then have a sequence of subsets of  $T$ ,  $\{T^0, T^1, \dots, T^{k^*}\}$ , where  $T^k = \Phi(T^{k-1}) \subsetneq T^{k-1}$  for all  $k = 1, 2, \dots, k^*$  and  $T^{k^*} = \emptyset$ . For any nonempty  $T' \subseteq T$ , there exists some  $k$  such that  $T' \subseteq T^k$  and  $T' \not\subseteq T^{k+1}$ . It is straightforward to verify that the operator  $\Phi(\cdot)$  is monotonically increasing in the set inclusion sense. As such,

$\Phi(T') \subseteq \Phi(T^k) = T^{k+1}$ . Since  $T' \not\subseteq T^{k+1}$ , we have  $\Phi(T') \neq T'$ . Thus, the data set satisfies the never-covered property under  $\supseteq$ .

## C.4 Proof of Theorem 3.3.10

By Theorem 3.3.5, Statements (i) and (ii) are equivalent, and obviously (iii) implies (i). So it remains to show that (ii) implies (iii).

To the original data set  $\mathcal{M}$  we add the observations  $\{(A^{at}, A^{at})\}_{t \in T}$  and  $\{(B^{bt}, B^{bt})\}_{t \in T}$ , where  $A^{at} = A^t$  and  $B^{bt} = B^t$  for each  $t \in T$ . Consider the augmented data set

$$\mathcal{M}^* = \{(A^t, B^t)\}_{t \in T} \cup \{(A^{at}, A^{at})\}_{t \in T} \cup \{(B^{bt}, B^{bt})\}_{t \in T}.$$

where all the added observations are considered weak preferences. Notice that  $\mathcal{M}$  is nicely rationalizable if and only if  $\mathcal{M}^*$  is rationalizable.

We denote a typical observation of  $\mathcal{M}^*$  by  $z$ ; the observation  $z$  may be in  $T$ , in  $aT := \{a1, a2, \dots, at, \dots\}$ , or in  $bT := \{b1, b2, \dots, bt, \dots\}$ . Suppose that  $\mathcal{M} = \{(A^t, B^t)\}_{t \in T}$  satisfies the never-covered property, i.e., for any  $T' \subseteq T$ , the set of dominated observations at  $T'$ ,  $\Phi(T')$ , satisfies  $\Phi(T') \neq T'$ . We claim that this implies that  $\mathcal{M}^*$  also satisfies the never-covered property, which will guarantee (by Theorem 3.3.5) that  $\mathcal{M}^*$  is rationalizable.

Given  $Z'$ , a nonempty subset of  $Z$ , we denote its dominated observations by  $\Phi^*(Z')$ . We need to show that  $\Phi^*(Z') \neq Z'$ . Suppose that  $Z' \cap S = \emptyset$ , where  $S \subseteq T$  is the set of observations with strict preferences. Then  $\Phi^*(Z') = \emptyset$  and so obviously  $\Phi^*(Z') \neq Z'$ . Now suppose that  $Z' \cap S$  is nonempty, which means that  $Z' \cap T$  is also nonempty. Since  $\mathcal{M}$  satisfies the never-covered property,  $(Z' \cap T) \setminus \Phi(Z' \cap T)$  is nonempty. It is straightforward to check in this case that if  $\hat{t} \in (Z' \cap T) \setminus \Phi(Z' \cap T)$

then  $\hat{t} \in Z' \setminus \Phi^*(Z')$ . Therefore,  $Z' \setminus \Phi^*(Z')$  is nonempty since  $(Z' \cap T) \setminus \Phi(Z' \cap T)$  is nonempty. We conclude that  $\mathcal{M}^*$  satisfies the never-covered property.

## C.5 Menu Preference Rationalization and Upper Contour Rationalization

It is straightforward to check that menu  $A$  is strictly preferred to menu  $B$  if and only if there is an upper contour rationalization of the following set of observations:

$$(\{\{y\}\}_{y \in A}, x) \text{ for each } x \in B.$$

For example,  $\{x, y\}$  is strictly preferred to  $\{z, w\}$  if and only if there is an upper contour rationalization of the observations  $(\{\{x\}, \{y\}\}, z)$  and  $(\{\{x\}, \{y\}\}, w)$ . Conversely, suppose we wish to guarantee that there is a set in the collection  $\{A_j\}_{j \in J}$  that is contained in the upper contour set of  $x$ ; this is equivalent to a rationalization of the following set of menu preference pairs:

$$(\cup_{j \in J} \{y_j\}, x) \text{ for each } \cup_{j \in J} \{y_j\} \text{ where } y_i \in A_j \text{ for all } j \in J.$$

For example, either  $\{x, y\}$  or  $\{z\}$  is in the upper contour set of  $w$  if and only if there is a rationalization of the following menu preferences:  $\{x, z\}$  is preferred to  $\{w\}$  and  $\{y, z\}$  is preferred to  $\{w\}$ .

Thus, it is always possible to convert an upper contour rationalization problem into a menu preference rationalization problem and vice versa when  $X$  is finite (but not when it is infinite) and any algorithm developed for one problem could, in principle, be used to solve the other. However, it should also be clear from the conversion procedure we outlined above that there is no general reason for solving either problem in this roundabout fashion, since the converted data set would have more (and in some case

many more) observations than the original data set. The two algorithms are best understood as distinct and serving different purposes.

## C.6 Proof of Theorem 3.3.15

The equivalence of Statements (i) and (ii) follows from Theorem 3.3.5 and obviously (iii) implies (i). It suffices to show that Statement (ii) implies (iii). By Theorem 3.3.5, since  $\mathcal{M}$  satisfies the never-covered property under  $\succeq$ , it admits a no-cycling selection  $\{x^t\}_{t \in T}$  under  $\succeq$ . From the proof of Theorem 3.3.5, we know that any preference that extends  $\text{tran}(R^* \cup \succeq)$  (the transitive closure of  $R^* \cup \succeq$ ) will rationalize the data and extend  $\succeq$ . It remains to show that there is a preference representable by a continuous utility function that extends  $\text{tran}(R^* \cup \succeq)$ . By Levin's Theorem, such an extension exists so long as  $\text{tran}(R^* \cup \succeq)$  is a closed preorder. That is indeed the case (see the proof of this claim contained in the proof of Theorem 2 in Nishimura, OK, and Quah (2017)) and it follows from the compactness of  $B^t$  for all  $t \in T$  and the finite number of observations.

## C.7 Proof of Theorem 3.4.1

Statement (iii) obviously implies (i). To show that (i) implies (ii), suppose that  $\mathcal{M}$  is rationalized by a locally nonsatiated preference  $\succsim$  and  $x^t \in L(p^t)$  satisfies  $x^t \succsim L(q^t)$  for all  $t \in T$  and  $x^t \succ L(q^t)$  for all  $t \in S$ . Let  $\hat{x} = \max(\{x^t\}_{t \in T'}; \succsim)$ . Denoting  $L_p(T') = \bigcup_{t \in T'} L(p^t)$  and  $L_q(T')$  in an analogous fashion, we note that  $L_p(T')$  cannot be covered by  $L_q(T')^{\downarrow\downarrow} \cup L_q(T' \cap S)^{\downarrow}$  since the former contains  $\hat{x}$  and the latter does not. Indeed,  $\hat{x} \succ L_q(T' \cap S)^{\downarrow} = L_q(T' \cap S)$ , and so  $\hat{x}$  is not in  $L_q(T' \cap S)$ . Nor can  $\hat{x}$  be in  $L_q(T')^{\downarrow\downarrow} = L_q(T')^o$  (the interior of  $L_q(T')$ ) because  $\hat{x} \succsim L_q(T')$  and the preference is locally nonsatiated. Taking it one step further, we know that  $\hat{x}$  is

not in  $L_q(T')^o \cup L_q((T' \cap S) \cup \Phi^1(T'))$ , where

$$\begin{aligned}\Phi^1(T') &:= \{t \in T' : L(p^t) \subseteq L_q(T')^{\downarrow\downarrow} \cup L_q(T' \cap S)^{\downarrow}\} \\ &= \{t \in T' : L(p^t) \subseteq L_q(T')^o \cup L_q(T' \cap S)\}.\end{aligned}$$

This is because  $\hat{x} \succ L_q((T' \cap S) \cup \Phi^1(T'))$ . Repeating this argument, we eventually conclude that  $\hat{x}$  is in  $A(T')$  but not in  $L_q(T')^o \cup L_q((T' \cap S) \cup \Phi(T'))$  and hence  $\Phi(T') \neq T'$ .

To show that (ii) implies (iii), note that Theorem 3.3.15 guarantees that there is a strictly increasing and continuous utility function  $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$  that rationalizes  $\mathcal{M}$ . So it suffices to show that there is a utility function  $\hat{u}$  that also rationalizes  $\mathcal{M}$ , which has the additional property of concavity. Each budget,  $L(p^t)$  is compact and so it has an optimum under  $u$  which we denote by  $x^t$  (if there are multiple optimal alternatives we may pick any one of them); similarly we denote the optimum bundle in  $L(q^t)$  by  $y^t$ . Since  $u$  is strictly increasing,  $p^t \cdot \bar{x}^t = q^t \cdot y^t = 1$ , i.e., the optimal bundle is on the budget plane and not just the budget set. Let  $\succsim$  denote the preference (i.e., the complete preorder) over  $\{x^t\}_{t \in T} \cup \{y^t\}_{t \in T}$  induced by  $u$ . Since it is generated by  $u$ , the notional data set  $\mathcal{N} = \{(x^t, p^t)\}_{t \in T} \cup \{(y^t, q^t)\}_{t \in T}$  is cyclically consistent in the sense of Afriat (1967) (equivalently, obeys the generalized axiom of revealed preference in the sense of Varian (1982)). The preference  $\succsim$  is a completion of the revealed preference relations generated by  $\mathcal{N}$  and defined on  $\{x^t\}_{t \in T} \cup \{y^t\}_{t \in T}$ ; by the Afriat's Theorem, there is a strictly increasing, continuous and *concave* utility function  $\hat{u}$  such that  $\hat{u}(x^t) \geq \hat{u}(x)$  for all  $x \in L(p^t)$ ,  $\hat{u}(y^t) \geq \hat{u}(x)$  for all  $x \in L(q^t)$ , and  $\hat{u}(x^t) \geq (>) u(y^t)$  if  $x^t \succsim (>) y^t$  (where  $\succ$  is the asymmetric part of  $\succsim$ ). Thus  $\hat{u}$  also rationalizes  $\mathcal{M}$ .

## C.8 Proof of Corollary 3.4.2

By Theorem 3.4.1, it suffices for us to show that  $L(p^t) \not\subseteq \bigcup_{t' \in T'} L(q^{t'})$  if and only if  $p^t \not\geq q$  for any  $q \in \text{conv}(\{q^s\}_{s \in T'})$ .

**(Only if-part)** Suppose that there exists some  $q \in \text{conv}(\{q^s\}_{s \in T'})$  such that  $p^t \geq q$ . Then for any  $x \in L(p^t)$ , we have  $p^t \cdot x \geq q \cdot x$ , which implies  $p^t \cdot x \geq q^{t'} \cdot x$  for some  $t' \in T'$ . Thus,  $x \in \bigcup_{s \in T'} L(q^s)$ . Since this is true for all any  $x \in L(p^t)$ , we have  $L(p^t) \subseteq \bigcup_{s \in T'} L(q^s)$ .

**(If-part)** Fix some nonempty  $T' \subseteq T$ . Let

$$Q := \{p \in \mathbb{R}_+^n : p \geq q \text{ for some } q \in \text{conv}(\{q^s\}_{s \in T'})\}.$$

Clearly,  $Q$  is closed and convex. Since there exists some  $t \in T'$  such that  $p^t \notin Q$ , by the Hyperplane Separating Theorem, there exists a vector  $r \in \mathbb{R}^n$  and a number  $b$  where  $r \neq 0$  such that  $p^t \cdot r = b < q \cdot r$  for any  $q \in Q$ . It is easy to verify that  $r \geq 0$ . Since  $r \neq 0$ , we have  $b = p^t \cdot r > 0$ . Let  $r' = \frac{r}{b}$ . We have  $p^t \cdot r' = 1 < q \cdot r'$  for all  $q \in Q$ . In words,  $r'$  is affordable under the price vector  $p^t$  but not under any  $q^s$ . Therefore,  $L(p^t) \not\subseteq \bigcup_{s \in T'} L(q^s)$ .

**Extension of Corollary 3.4.2** Consider a set of preferences over budget sets  $\mathcal{M} = \{(L(p^t), L(q^t))\}_{t \in T}$ . The data set  $\mathcal{M}$  is rationalizable by a strictly increasing, continuous, and concave utility function  $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$  if and only if the following property holds: for any nonempty  $T' \subseteq T$ ,

there exists  $\Phi \subsetneq T'$  and  $\epsilon \gg 0$  such that  $p \not\geq q$  whenever  $p \ll p^t$  for some

$$t \in T' \setminus \Phi \text{ and } q \in \text{conv}(\{q^s\}_{s \in T'} \cup \{q^r - \epsilon\}_{r \in \Phi \cup (T' \cap S)}).$$
<sup>2</sup>

---

<sup>2</sup> When  $T = S$ , the characterizing condition in this corollary reduces to that in Corollary 3.4.2. To see this, note that if  $T' \cap S = T'$ , then the condition in Corollary ?? is equivalent to saying that for any nonempty  $T' \subseteq T$ , there exists  $t \in T'$  and  $\epsilon \gg 0$  such that for any  $p \ll p^t$  and  $q \in \text{conv}(\{q^s\}_{s \in T'} \cup \{q^r - \epsilon\}_{r \in T'})$ ,  $p \not\geq q$ . By continuity, this is equivalent to the condition in Corollary 3.4.2.

**Proof.** By Theorem 3.4.1, it suffices to show that  $\mathcal{M}$  satisfies the never-covered property under  $\geq$  if and only if the property in the statement holds. The never-covered property under  $\geq$  requires that for any nonempty  $T' \subseteq T$ ,  $T' \setminus \Phi(T') \neq \emptyset$ . By the definition of  $\Phi(T')$ , if  $t \in T' \setminus \Phi(T')$ , then

$$L(p^t) \not\subseteq \left( \bigcup_{t' \in T'} L(q^{t'})^o \right) \cup \left( \bigcup_{t' \in (T' \cap S) \cup \Phi(T')} L(q^{t'}) \right).$$

We first prove the following lemma.

**Lemma C.8.1**  $L(p) \not\subseteq \left( \bigcup_{s=1}^n L(q^s) \right) \cup \left( \bigcup_{r=1}^m L(q^r)^o \right)$  if and only if there exists some  $\epsilon \gg 0$  such that for any  $\hat{p} \ll p$ ,  $L(\hat{p}) \not\subseteq \left( \bigcup_{s=1}^n L(q^s - \epsilon) \right) \cup \left( \bigcup_{r=1}^m L(q^r) \right)$ .

**Proof. (Only if-part)** Since  $L(p) \not\subseteq \left( \bigcup_{s=1}^n L(q^s) \right) \cup \left( \bigcup_{r=1}^m L(q^r)^o \right)$ , there exists some  $x \in L(p)$  such that  $x \notin \left( \bigcup_{s=1}^n L(q^s) \right) \cup \left( \bigcup_{r=1}^m L(q^r)^o \right)$ . Fix such an  $x$ . Since  $x \in L(p)$ , for any  $\hat{p} \ll p$ , there exists some  $\delta \gg 0$  such that  $x + \delta \in L(\hat{p})$ . Fix such a  $\delta$ . Since  $x \notin \bigcup_{s=1}^n L(q^s)$ , there exists some sufficiently small  $\epsilon \gg 0$  such that  $x \notin \bigcup_{s=1}^n L(q^s - \epsilon)$ , which further implies that  $x + \delta \notin \bigcup_{s=1}^n L(q^s - \epsilon)$ . Since  $x \notin \bigcup_{r=1}^m L(q^r)^o$ ,  $x + \delta \notin \bigcup_{r=1}^m L(q^r)$ . Since  $x + \delta$  is contained in  $L(\hat{p})$  but not in  $\left( \bigcup_{s=1}^n L(q^s - \epsilon) \right) \cup \left( \bigcup_{r=1}^m L(q^r) \right)$ , we have  $L(\hat{p}) \not\subseteq \left( \bigcup_{s=1}^n L(q^s - \epsilon) \right) \cup \left( \bigcup_{r=1}^m L(q^r) \right)$ .

**(If-part)** Fix  $\epsilon \gg 0$  such that for any  $\hat{p} \ll p$ ,  $L(\hat{p}) \not\subseteq \left( \bigcup_{s=1}^n L(q^s - \epsilon) \right) \cup \left( \bigcup_{r=1}^m L(q^r) \right)$ . Clearly, for each  $\theta \in (\frac{1}{2}, 1)$ ,  $\theta p \ll p$ , and there exists  $x^\theta \in L(\theta p)$  such that  $x^\theta \notin \left( \bigcup_{s=1}^n L(q^s - \epsilon) \right) \cup \left( \bigcup_{r=1}^m L(q^r) \right)$ . Note that  $\{x^\theta\}_{\theta \in (\frac{1}{2}, 1)} \in L(\frac{1}{2}p)$ , which is compact. Therefore,  $\{x^\theta\}_{\theta \in (\frac{1}{2}, 1)}$  has a convergent subsequence. Let  $x$  denote the limit of this subsequence. Since  $x^\theta \in L(\theta p)$ ,  $x \in L(p)$ . Furthermore, (1)  $x \notin \left( \bigcup_{s=1}^n L(q^s) \right)$ , otherwise  $x \in L(q^s - \epsilon)^o$  for some  $s$ , which implies that  $x^\theta \in L(q^s - \epsilon)^o$  for some  $\theta$  and some  $s$ ; and (2)  $x \notin \left( \bigcup_{r=1}^m L(q^r)^o \right)$ , otherwise  $x^\theta \in L(q^r)$  for some  $r$  and some  $\theta$ . Since  $x$  is contained in  $L(p)$  but not in  $\left( \bigcup_{s=1}^n L(q^s) \right) \cup \left( \bigcup_{r=1}^m L(q^r)^o \right)$ , we have  $L(p) \not\subseteq \left( \bigcup_{s=1}^n L(q^s) \right) \cup \left( \bigcup_{r=1}^m L(q^r)^o \right)$ . ■

It is easy to see that  $L(\hat{p}) \not\subseteq (\bigcup_{s=1}^n L(q^s - \epsilon)) \cup (\bigcup_{r=1}^m L(q^r))$  if and only if for all  $q \in \text{conv}(\{q^s - \epsilon\}_{s=1}^n \cup \{q^r\}_{r=1}^m)$ ,  $\hat{p} \not\geq q$  (using similar argument as in the proof of the if-part of Corollary 3.4.2). The desired result then follows from this and Proposition 3.3.6. ■

## C.9 Proof of Theorem 3.5.2

We skip the proof that Statement (i) implies (ii), which is straightforward and similar to the argument given in Section 3.4 for the claim that (i) implies (ii) in Theorem 3.4.1. It is also obvious that (iii) implies (i). So it remains for us to show that (ii) implies (ii). An appeal to Theorem 3.3.15 guarantees that there is  $x^t$  (for each  $t \in T$ ) and a strictly increasing and continuous utility function  $\tilde{u} : \mathbb{R}_+^n \rightarrow \mathbb{R}$  such that  $\tilde{u}(x^t) \geq \tilde{u}(x)$  for all  $x \in L(p^t)$ . Therefore the notional data set  $\{(x^t, p^t)\}_{t \in T}$  must satisfy cyclical consistency (equivalently GARP). By Afriat's Theorem, there is a strictly increasing, continuous, and concave utility function  $u$  such that  $u(x^t) \geq u(x)$  for all  $x \in L(p^t)$  for all  $t \in T$ .

## C.10 Proof of Theorem 3.6.2

(Statement (i)  $\Rightarrow$  Statement (ii)) Suppose that the data set  $(\Sigma, f)$  is rationalizable by multiple preferences. Fix an arbitrary nonempty  $\Sigma' \subseteq \Sigma$  and  $B \in \Sigma$ . Since  $(\Sigma, f)$  is rationalizable by multiple preferences,  $g(C) \subseteq g(\bigcup_{A \in \Sigma'} A)$  for all  $C \in \Sigma'$ . It follows that  $\bigcup_{A \in \Sigma'} g(A) \subseteq g(\bigcup_{A \in \Sigma'} A)$ . We then have

$$\begin{aligned} f(\bigcup_{A \in \Sigma'} A) &= \bigcup_{A \in \Sigma'} A \setminus g(\bigcup_{A \in \Sigma'} A) \\ &\subseteq \bigcup_{A \in \Sigma'} A \setminus \bigcup_{A \in \Sigma'} g(A) \\ &= \bigcup_{A \in \Sigma'} (f(A) \cup g(A)) \setminus \bigcup_{A \in \Sigma'} g(A) \end{aligned}$$

$$= \cup_{A \in \Sigma'} f(A) \setminus \cup_{A \in \Sigma'} g(A).$$

Thus, if  $(\cup_{A \in \Sigma'} f(A) \setminus \cup_{A \in \Sigma'} g(A)) \subseteq B$ , we must have that  $f(\cup_{A \in \Sigma'} A) \subseteq B$ . Since  $(\Sigma, f)$  is rationalizable by multiple preferences and  $f(\cup_{A \in \Sigma'} A) \subseteq B$ ,  $f(B) \cap g(\cup_{A \in \Sigma'} A) = \emptyset$ . Since  $\cup_{A \in \Sigma'} g(A) \subseteq g(\cup_{A \in \Sigma'} A)$ , we have  $f(B) \cap (\cup_{A \in \Sigma'} g(A)) = \emptyset$ .

**(Statement (ii)  $\Rightarrow$  Statement (iii))** We show that for each  $A \in \Sigma$  and  $x \in f(A)$ , the data set  $\mathcal{O}_{A,x}$  is rationalizable by a strict preference. Suppose to the contrary, for some  $A \in \Sigma$  and  $x \in f(A)$ , the data set  $\mathcal{O}_{A,x}$  is not rationalizable by a strict preference. By Corollary 3.3.12, there exists some nonempty  $\Sigma' \subseteq \Sigma$  such that

$$\cup_{B \in \Sigma'} f_{A,x}(B) \subseteq \cup_{B \in \Sigma'} g_{A,x}(B). \quad (\text{C.1})$$

Recall that in the main text, we show that the second statement implies that the data set  $\mathcal{O} = \{(f(A), g(A))\}_{A \in \Sigma}$  is rationalizable by a strict preference. By Corollary 3.3.12,

$$\cup_{B \in \Sigma'} f(B) \not\subseteq \cup_{B \in \Sigma'} g(B).$$

Since  $\mathcal{O}$  and  $\mathcal{O}_{A,x}$  only differ when the menu is  $A$ , we can conclude that  $A \in \Sigma'$ . Let  $\Sigma'' := \Sigma' \setminus \{A\}$ . Since  $\mathcal{O} = \{(f(A), g(A))\}_{A \in \Sigma}$  is rationalizable by a strict preference, by Corollary 3.3.12, we have

$$\cup_{B \in \Sigma''} f_{A,x}(B) \not\subseteq \cup_{B \in \Sigma''} g_{A,x}(B). \quad (\text{C.2})$$

It follows from (C.1) and (C.2) that  $(\cup_{B \in \Sigma''} f_{A,x}(B) \setminus \cup_{B \in \Sigma''} g_{A,x}(B)) \subseteq g_{A,x}(A) \subseteq A$  and  $x \in \cup_{B \in \Sigma''} g_{A,x}(B)$ . Therefore,  $x \in f(A) \cap (\cup_{B \in \Sigma''} g_{x,A}(B))$ . Since  $\mathcal{O}$  and  $\mathcal{O}_{A,x}$  coincide on  $\Sigma''$ , the second statement requires that  $f(A) \cap (\cup_{B \in \Sigma''} g_{x,A}(B)) = \emptyset$ . We arrive at a contradiction.

(**Statement (iii)  $\Rightarrow$  Statement (i)**) Suppose that for any  $A \in \Sigma$  and  $x \in f(A)$ ,  $\mathcal{O}_{A,x}$  is rationalizable by a strict preference. Let  $\succ_{A,x}$  denote a strict preference that rationalizes  $\mathcal{O}_{A,x}$ . By definition,  $\max(A; \succ_{A,x}) = x \in f(A)$  and  $\max(A'; \succ_{A,x}) \in f(A')$  for any  $A' \in \Sigma$  and  $A' \neq A$ . We claim that the set of strict preferences  $\{\succ_{A,x}\}_{A \in \Sigma, x \in f(A)}$  rationalizes  $(\Sigma, f)$  under the multiple preferences model. It suffices to show that

$$f_{\{\succ_{A,x}\}_{A \in \Sigma, x \in f(A)}}(A') = f(A')$$

for all  $A' \in \Sigma$ . This follows immediately from the construction of the set of strict preferences, since for each  $A' \in \Sigma$ , (1) each element  $y$  in  $f(A')$  is the maximal element in  $A'$  according to the strict preference  $\succ_{A',y}$ ; and (2) for any strict preference  $\succ$  in  $\{\succ_{A,x}\}_{A \in \Sigma, x \in f(A)}$ , the maximal element in  $A'$  according to  $\succ$  lies in  $f(A')$ .

## C.11 Proof of Corollary 3.6.3

The only if-part is trivial. We only prove the if-part below. We show that the second statement in Theorem 3.6.2 follows from the Chernoff axiom and the Aizerman axiom.

**Claim C.11.1** *For any  $A \in \mathcal{X}$ , if  $x \in g(A)$ , then  $f(f(A) \cup x) = f(A)$ .*

**Proof.** Since  $(f(A) \cup x) \subseteq A$ , the Chernoff axiom implies that  $f(A) \subseteq f(f(A) \cup x)$ . Since  $f(A) \subseteq (f(A) \cup x) \subseteq A$ , the Aizerman axiom implies that  $f(f(A) \cup x) \subseteq f(A)$ . Thus,  $f(f(A) \cup x) = f(A)$ . ■

**Claim C.11.2** *For any  $A, B \in \mathcal{X}$ , if  $f(A) \subseteq B$ , then  $f(B) \cap g(A) = \emptyset$ .*

**Proof.** Suppose that there exists some  $x \in (f(B) \cap g(A))$ . Since  $x \in g(A)$ , by Claim C.11.1,  $x$  is not chosen in  $f(A) \cup x$ . The Chernoff axiom implies that  $x$  is not chosen in any superset of  $f(A) \cup x$ . In particular,  $x \notin f(B)$ . We have a contradiction. ■

**Claim C.11.3** For any nonempty  $\mathcal{X}' \subseteq \mathcal{X}$  and  $B \in \mathcal{X}$ , if  $(\cup_{A \in \mathcal{X}'} f(A) \setminus \cup_{A \in \mathcal{X}'} g(A)) \subseteq B$ , then  $f(B) \cap (\cup_{A \in \mathcal{X}'} g(A)) = \emptyset$ .

**Proof.** Fix an arbitrary nonempty  $\mathcal{X}' \subseteq \mathcal{X}$  and  $B \in \mathcal{X}$ . By the Chernoff axiom,  $g(C) \subseteq g(\cup_{A \in \mathcal{X}'} A)$  for all  $C \in \mathcal{X}'$ . Therefore,  $\cup_{A \in \mathcal{X}'} g(A) \subseteq g(\cup_{A \in \mathcal{X}'} A)$ . We then have

$$\begin{aligned} f(\cup_{A \in \mathcal{X}'} A) &= \cup_{A \in \mathcal{X}'} A \setminus g(\cup_{A \in \mathcal{X}'} A) \\ &\subseteq \cup_{A \in \mathcal{X}'} A \setminus \cup_{A \in \mathcal{X}'} g(A) \\ &= \cup_{A \in \mathcal{X}'} (f(A) \cup g(A)) \setminus \cup_{A \in \mathcal{X}'} g(A) \\ &= \cup_{A \in \mathcal{X}'} f(A) \setminus \cup_{A \in \mathcal{X}'} g(A). \end{aligned}$$

Thus, if  $(\cup_{A \in \mathcal{X}'} f(A) \setminus \cup_{A \in \mathcal{X}'} g(A)) \subseteq B$ , we must have  $f(\cup_{A \in \mathcal{X}'} A) \subseteq B$ . By Claim C.11.2,  $f(B) \cap g(\cup_{A \in \mathcal{X}'} A) = \emptyset$ . The claim follows since  $\cup_{A \in \mathcal{X}'} g(A) \subseteq g(\cup_{A \in \mathcal{X}'} A)$ .

■

By Theorem 3.6.2, the data set is rationalizable by multiple preferences.

## C.12 Proof of Theorem 3.7.3

Let  $\bar{X} := \{(x, y) \in X \times X : x \neq y\}$ .

**(Only if-part)** Suppose that the data set  $(\Sigma, f)$  is rationalizable under the min-max regret model. By definition, there exists a finite set  $U$  of utility functions such that

$$\max_{y \in A \setminus f(A)} \phi_U(f(A), y) < \max_{y \in A \setminus x} \phi_U(x, y)$$

for all  $A \in \Sigma$  and  $x \in A \setminus f(A)$ . Since  $X$  is finite and these inequalities are strict, there exists a finite set  $U'$  of utility functions such that

$$\max_{y \in A \setminus f(A)} \phi_{U'}(f(A), y) < \max_{y \in A \setminus x} \phi_{U'}(x, y)$$

for all  $A \in \Sigma$  and  $x \in A \setminus f(A)$ , and  $\phi_{U'}(x, y) = \phi_{U'}(z, w)$  only if  $(x, y) = (z, w)$ . Define a strict partial order  $P$  such that  $(x, y) P (z, w)$  if and only if  $\phi_{U'}(x, y) > \phi_{U'}(z, w)$ . It follows from the Szpilrajn's extension theorem that we can extend  $P$  to a strict preference  $\succ$ . It is straightforward to verify that  $\bar{\mathcal{O}}$  can be rationalized by the strict preference  $\succ$ .

**(If-part)** Suppose that  $\bar{\mathcal{O}}$  is rationalizable by a strict preference, say  $\succ$ . It suffices to show that there exists a finite set  $U$  of utility functions such that for any  $(x, y), (z, w) \in \bar{X}$ ,  $\phi_U(x, y) > \phi_U(z, w)$  if and only if  $(x, y) \succ (z, w)$ .

Since  $\bar{X}$  is finite, we can construct a function  $\beta : \bar{X} \rightarrow (1, 2)$  such that  $\beta(x, y) > \beta(z, w)$  if and only if  $(x, y) \succ (z, w)$ . We now construct a finite set of utility functions

$$U = \{u_{x,y}\}_{(x,y) \in \bar{X}},$$

where the utility functions are indexed by  $(x, y) \in \bar{X}$ . For each  $(x, y) \in \bar{X}$ , let

$$u_{x,y}(z) = \begin{cases} 0, & \text{if } z = x; \\ \beta(x, y), & \text{if } z = y; \\ \frac{\beta(x,y)}{2}, & \text{otherwise.} \end{cases}$$

Obviously,  $u_{x,y}(y) - u_{x,y}(x) = \beta(x, y) > 1$ . We claim that  $u_{z,w}(y) - u_{z,w}(x) < 1$  if  $(z, w) \neq (x, y)$ . First consider the case in which  $w \neq y$ . It follows from the construction of the utility functions and the  $\beta$  function that

$$u_{z,w}(y) - u_{z,w}(x) \leq u_{z,w}(y) \leq \frac{\beta(z, w)}{2} < 1.$$

Next, we consider the case in which  $w = y$ . Since  $(z, w) \neq (x, y)$ , we have  $z \neq x$ . By the construction of the utility functions,

$$u_{z,w}(y) - u_{z,w}(x) = \beta(z, w) - \frac{\beta(z, w)}{2} = \frac{\beta(z, w)}{2} < 1.$$

Therefore, we can conclude that

$$\phi_U(x, y) = \max_{u \in U} \{u(y) - u(x)\} = u_{x,y}(y) - u_{x,y}(x) = \beta(x, y).$$

We have constructed a finite set  $U$  of utility functions such that for any  $(x, y), (z, w) \in \bar{X}$ ,

$$\phi_U(x, y) > \phi_U(z, w) \iff \beta(x, y) > \beta(z, w) \iff (x, y) \succ (z, w).$$

This completes the proof of the if-part.

# References

- AFRIAT, S. N. (1967): “The Construction of Utility Functions from Expenditure Data,” *International Economic Review*, 8(1), 67–77.
- AHN, D. S., R. IJIMA, Y. LE YAOUANQ, AND T. SARVER (2019): “Behavioural Characterizations of Naivete for Time-Inconsistent Preferences,” *The Review of Economic Studies*, 86(6), 2319–2355.
- AIZERMAN, M., AND A. MALISHEVSKI (1981): “General Theory of Best Variants Choice: Some Aspects,” *IEEE Transactions on Automatic Control*, 26(5), 1030–1040.
- AUGENBLICK, N., M. NIEDERLE, AND C. SPRENGER (2015): “Working over Time: Dynamic Inconsistency in Real Effort Tasks,” *The Quarterly Journal of Economics*, 130(3), 1067–1115.
- AUGENBLICK, N., AND M. RABIN (2018): “An Experiment on Time Preference and Misprediction in Unpleasant Tasks,” *The Review of Economic Studies*, 86(3), 941–975.
- BEAUCHÊNE, D., J. LI, AND M. LI (2019): “Ambiguous Persuasion,” *Journal of Economic Theory*, 179, 312–365.
- BERTSIMAS, D., AND J. N. TSITSIKLIS (1997): *Introduction to Linear Optimization*, vol. 6. Athena Scientific Belmont, MA.
- BLUME, A., AND O. BOARD (2014): “Intentional Vagueness,” *Erkenntnis*, 79(4), 855–899.
- BOSE, S., AND L. RENOU (2014): “Mechanism Design With Ambiguous Communication Devices,” *Econometrica*, 82(5), 1853–1872.
- BREIG, Z., M. GIBSON, AND J. SHRADER (2020): “Why do We Procrastinate? Present Bias and Optimism,” Working Paper, The University of Queensland, Williams College and Columbia University.
- CHAMBERS, C. P., AND F. ECHENIQUE (2016): *Revealed Preference Theory*, vol. 56. Cambridge University Press.

- CHANDRASEKHAR, M. (2018): “Informal Commitments in Planner–Doer Games,” *Journal of Economic Theory*, 173, 201–230.
- CHENG, X. (2021): “Relative Maximum Likelihood Updating of Ambiguous Beliefs,” Working Paper, Northwestern University.
- CHEW, S. H., AND J. S. SAGI (2008): “Small Worlds: Modeling Attitudes Toward Sources of Uncertainty,” *Journal of Economic Theory*, 139(1), 1–24.
- CRIPPS, M. W. (2018): “Divisible Updating,” Working Paper, University College London.
- CROSON, R. T. (1999): “The Disjunction Effect and Reason-based Choice in Games,” *Organizational Behavior and Human Decision Processes*, 80(2), 118–133.
- DE CLIPPEL, G., AND K. ROZEN (2012): “Bounded Rationality and Limited Datasets: Testable Implications, Identifiability, and Out-of-Sample Prediction,” mimeo, Brown University and Yale University.
- (forthcoming): “Bounded Rationality and Limited Datasets,” *Theoretical Economics*.
- DEKEL, E., AND B. L. LIPMAN (2012): “Costly Self-Control and Random Self-Indulgence,” *Econometrica*, 80(3), 1271–1302.
- DEKEL, E., B. L. LIPMAN, AND A. RUSTICHINI (2001): “Representing Preferences with A Unique Subjective State Space,” *Econometrica*, 69(4), 891–934.
- (2009): “Temptation-Driven Preferences,” *The Review of Economic Studies*, 76(3), 937–971.
- DELLAVIGNA, S., AND U. MALMENDIER (2006): “Paying Not to Go to the Gym,” *American Economic Review*, 96(3), 694–719.
- DEMPSTER, A. P. (1967): “Upper and Lower Probabilities Induced by a Multivalued Mapping,” *The Annals of Mathematical Statistics*, 38(2), 325–339.
- ELLSBERG, D. (1961): “Risk, Ambiguity, and the Savage Axioms,” *The Quarterly Journal of Economics*, 70(4), 643–669.
- EPSTEIN, L. G., Y. HALEVY, ET AL. (2019): “Hard-to-Interpret Signals,” Working Paper, Boston University and University of Toronto.
- ERGIN, H., AND T. SARVER (2010): “A Unique Costly Contemplation Representation,” *Econometrica*, 78(4), 1285–1339.
- FEDYK, A. (2018): “Asymmetric Naivete: Beliefs about Self-Control,” Working Paper, University of California, Berkeley.

- FISHBURN, P. (1976): “Representable Choice Functions,” *Econometrica*, 44(5), 1033–1043.
- FORGES, F., AND E. MINELLI (2009): “Afriat’s Theorem for General Budget Sets,” *Journal of Economic Theory*, 144(1), 135–145.
- GILBOA, I., AND D. SCHMEIDLER (1989): “Maxmin Expected Utility with Non-unique Prior,” *Journal of Mathematical Economics*, 18(2), 141–153.
- (1993): “Updating Ambiguous Beliefs,” *Journal of Economic Theory*, 59(1), 33–49.
- GUL, F., AND W. PESENDORFER (2001): “Temptation and Self-Control,” *Econometrica*, 69(6), 1403–1435.
- (2014): “Expected Uncertain Utility Theory,” *Econometrica*, 82(1), 1–39.
- (forthcoming): “Evaluating Ambiguous Random Variables and Updating by Proxy,” *Journal of Economic Theory*.
- HANANY, E., AND P. KLIBANOFF (2007): “Updating Preferences with Multiple Priors,” *Theoretical Economics*, 2(3), 261–298.
- HRISTOVA, E., AND M. GRINBERG (2008): “Disjunction Effect in Prisoner’s Dilemma: Evidences from an Eye-Tracking Study,” in *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, pp. 1225–1230.
- JACKSON, M. O. (1986): “Continuous Utility Functions in Consumer Theory: A Set of Duality Theorems,” *Journal of Mathematical Economics*, 15(1), 63–77.
- JAFFRAY, J.-Y. (1992): “Bayesian Updating and Belief Functions,” *IEEE Transactions on Systems, Man, and Cybernetics*, 22(5), 1144–1152.
- KAMENICA, E., AND M. GENTZKOW (2011): “Bayesian Persuasion,” *American Economic Review*, 101(6), 2590–2615.
- KELLNER, C., AND M. T. LE QUEMENT (2017): “Modes of Ambiguous Communication,” *Games and Economic Behavior*, 104, 271–292.
- (2018): “Endogenous Ambiguity in Cheap Talk,” *Journal of Economic Theory*, 173, 1–17.
- KELLNER, C., M. T. LE QUEMENT, AND G. RIENER (2019): “Reacting to Ambiguous Messages: An Experimental Analysis,” Working Paper, University of Southampton, University of East Anglia and Düsseldorf Institute for Competition Economics.
- KREPS, D. M. (1979): “A Representation Theorem for ‘Preference for Flexibility’,” *Econometrica*, 47(3), 565–577.

- (2013): *Microeconomic Foundations I: Choice and Competitive Markets*, vol. 1. Princeton University Press.
- KRISHNA, V., AND H. SONNENSCHN (1990): “Duality in Consumer Theory,” in *Preferences, Uncertainty, and Optimality*, ed. by J. S. Chipman, D. McFadden, and M. Richter, pp. 44–55. Boulder, CO: Westview Press.
- LARWOOD, L., AND W. WHITTAKER (1977): “Managerial Myopia: Self-Serving Biases in Organizational Planning,” *Journal of Applied Psychology*, 62(2), 194–198.
- LIANG, Y. (2021): “Learning from Unknown Information Sources,” Working Paper, briq Institute.
- MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006): “Ambiguity Aversion, Robustness, and the Variational Representation of Preferences,” *Econometrica*, 74(6), 1447–1498.
- MILNOR, J. (1954): “Games Against Nature,” in *Decision Processes*, ed. by R. M. Thrall, C. H. Coombs, and R. L. Davis, pp. 49–59. New York: Wiley.
- NISHIMURA, H., E. A. OK, AND J. K.-H. QUAH (2017): “A Comprehensive Approach to Revealed Preference Theory,” *American Economic Review*, 107(4), 1239–1263.
- O’DONOGHUE, T., AND M. RABIN (1999): “Doing It Now or Later,” *American Economic Review*, 89(1), 103–124.
- (2001): “Choice and Procrastination,” *The Quarterly Journal of Economics*, 116(1), 121–160.
- ORTOLEVA, P. (2012): “Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News,” *American Economic Review*, 102(6), 2410–36.
- PIRES, C. P. (2002): “A Rule for Updating Ambiguous Beliefs,” *Theory and Decision*, 53(2), 137–152.
- RENY, P. J. (2015): “A Characterization of Rationalizable Consumer Behavior,” *Econometrica*, 83(1), 175–192.
- RICHTER, M. K. (1966): “Revealed Preference Theory,” *Econometrica*, 34(3), 635–645.
- SALANT, Y., AND A. RUBINSTEIN (2008): “(A,f): Choice with Frames,” *The Review of Economic Studies*, 75(4), 1287–1296.
- SAVAGE, L. J. (1951): “The Theory of Statistical Decision,” *Journal of the American Statistical Association*, 46(253), 55–67.

- SCHMEIDLER, D. (1989): “Subjective Probability and Expected Utility without Additivity,” *Econometrica*, 57(3), 571–587.
- SHAFER, G. (1976): *A Mathematical Theory of Evidence*, vol. 42. Princeton University Press.
- SHISHKIN, D., AND P. ORTOLEVA (2019): “Ambiguous Information and Dilation: An Experiment,” Working Paper, University of California San Diego and Princeton University.
- STOVALL, J. E. (2010): “Multiple Temptations,” *Econometrica*, 78(1), 349–376.
- STOYE, J. (2011): “Statistical Decisions under Ambiguity,” *Theory and Decision*, 70(2), 129–148.
- STROTZ, R. H. (1955): “Myopia and Inconsistency in Dynamic Utility Maximization,” *The Review of Economic Studies*, 23(3), 165–180.
- SZPILRAJN, E. (1930): “Sur l’extension de l’ordre partiel,” *Fundamenta Mathematicae*, 16(1), 386–389.
- THALER, R. H., AND H. M. SHEFRIN (1981): “An Economic Theory of Self-Control,” *Journal of Political Economy*, 89(2), 392–406.
- TVERSKY, A., AND E. SHAFIR (1992): “The Disjunction Effect in Choice under Uncertainty,” *Psychological Science*, 3(5), 305–310.
- TYSON, C. J. (2018): “Rationalizability of Menu Preferences,” *Economic Theory*, 65(4), 917–934.
- VARIAN, H. R. (1982): “The Nonparametric Approach to Demand Analysis,” *Econometrica*, 50(4), 945–973.
- WALD, A. (1950): “Statistical Decision Functions,” Wiley.
- WASSERMAN, L. A., AND J. B. KADANE (1990): “Bayes’ Theorem for Choquet Capacities,” *The Annals of Statistics*, 18(3).