Essays in Behavioral Decision Theory

Thesis by
Gerelt Tserenjigmid

In Partial Fulfillment of the Requirements for the
degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2016
Defended (April 20, 2016)

Gerelt Tserenjigmid
ORCID: 0000-0003-1412-9692

I dedicate this project to my parents, my friends, and my fiancée Manal.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisors Federico Echenique and Pietro Ortoleva for their advice, guidance, and encouragement. I am also indebted to the other members of my committee, Leeat Yariv and Kota Saito. I also have benefited very much from insightful discussions with Colin Camerer, Jaksa Cvitanic, Lawrence Jin, Michihiro Kandori, Akihiko Matsui, and Jean-Laurent Rosenthal. I would also thank to my fellow classmates Ryo Adachi, Rahul Bhui, Nick Broten, Liam Clegg, Simon Dunne, Marcelo Fernandez, Matt Kovach, Sergio Montero, Euncheol Shin, Jay Viloria, and Jackie Zhang, for their support and friendship.

ABSTRACT

Many different behavioral phenomena that cannot be rationalized by standard models in economics have been well-documented both in the real world and in lab experiments. Motivated by these behavioral phenomena, the purpose of this dissertation is three-fold. First, I develop axiomatic models of individual decision-making to explain these well-documented phenomena. Second, I derive the implications and predictions of these axiomatic models for intertemporal choice, asset pricing, and other economic contexts. Third, I provide connections between these seemingly separate behavioral phenomena and widely-used properties of preferences in economics and psychology. This dissertation consists of five chapters. The first chapter studies dynamic choice under uncertainty. The second and third chapters study choice over multi-attribute alternatives. The fourth and fifth chapters study stochastic choice.

The first chapter studies history-dependent risk aversion and focuses on a behavioral phenomenon called the *reinforcement effect* (RE), which states that people become less risk-averse after a good history than after a bad history. The RE is well-documented in consumer choices, financial markets, and lab experiments. I show that this seemingly anomalous behavior occurs whenever risk preferences are history-dependent (in a nontrivial way) and satisfy *monotonicity with respect to first-order stochastic dominance*. To study history-dependent risk aversion and the RE formally, I develop a behaviorally-founded model of dynamic choice under risk that generalizes standard discounted expected utility. To illustrate the usefulness of my model, I apply it to the Lucas tree model of asset pricing and draw implications of the RE for asset price dynamics. I find that, compared to history-independent models, assets are overpriced when the economy is in a good state and are underpriced in a bad state. Moreover, my model generates high, volatile, and predictable asset returns, and low and smooth bond returns, consistent with empirical evidence.

In the second chapter, I develop an axiomatic model of reference-dependent preferences in which reference points are endogenous. In particular, I focus on choices from menus of two-attribute alternatives, and the reference point for a given menu is a vector that consists of the minimums of each dimension of the menu. I characterize this model by two weakenings of the *Weak Axiom of Revealed Preference* (WARP) in addition to standard axioms. My model is not just consistent with the *attraction effect* and the *compromise effect*, well-known preference reversals, but

it also provides a connection between these two effects and *diminishing sensitivity*, a widely used behavioral property in economics. The model also provides bounds on preference reversals. I apply the model to two different contexts, intertemporal choice and risky choice, and diminishing sensitivity has interesting implications. In intertemporal choice, the main implication of the model is that borrowing constraints produce a psychological pressure to move away from the constraints even if they are not binding. In risky choice, the model allows conflicting risk behaviors.

In the third chapter, I study choice over multidimensional alternatives. Making a choice between multidimensional alternatives is a difficult task. Therefore, a decision maker may adopt some procedure (heuristic) to simplify this task. I provide an axiomatic model of one such heuristic called the *Intra-Dimensional Comparison* (IDC) heuristic. The IDC heuristic is well-documented in the experimental literature on choice under risk. The IDC heuristic is a procedure in which a decision maker compares multidimensional alternatives dimension-by-dimension and makes a decision based on those comparisons. The model of the IDC heuristic provides a general framework applicable to many different contexts, including risky choice and social choice.

The fourth chapter is joint work with Federico Echenique and Kota Saito. We develop an axiomatic theory of random choice that builds on Luce's (1959) model to incorporate a role for perception. We capture the role of perception through *perception priorities;* priorities that determine whether an object or alternative is perceived sooner or later than other alternatives. We identify agents' perception priorities from their violations of Luce's axiom of *independence from irrelevant alternatives* (IIA). The direction of the violation of IIA implies an orientation of agents' priority rankings. We adjust choice probabilities to account for the effects of perception, and impose that adjusted choice probabilities satisfy IIA. So all violations of IIA are accounted for by the perception order. The theory can explain some very well-documented behavioral phenomena in individual choice. We can also explain the effects of forced choice and choice overload in experiments.

The fifth chapter studies how the ordering of alternatives (e.g., the location of products in a grocery store, the order of candidates on a ballot) affects a decision maker's choices. I develop an axiomatic model of random choice that builds on Luce's (1959) and incorporates the effect of the ordering of alternatives on choice frequencies. When the ordering of alternatives is observed, I characterize the model by two weakenings of IIA. When the ordering of alternatives is not observed, I can identify it from choice data. The model can accommodate the similarity, compromise, and attraction effects, violations of stochastic transitivity, and the choice overload, which are well-known behavioral phenomena in individual choice.

PUBLISHED CONTENT AND CONTRIBUTIONS

As the date of April 20th, 2016, I declare that (i) Chapter III in this dissertation, "Theory of decisions by intra-dimensional comparisons," has been published in the *Journal of Economic Theory*, and (ii) Chapter IV in this dissertation, "The Perception-Adjusted Luce Model," has been revised and resubmitted to the *Theoretical Economics*.

Tserenjigmid, G. (2015): "Theory of decisions by intra-dimensional comparisons," *Journal of Economic Theory*, 159, 326–338.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

HISTORY-DEPENDENT RISK AVERSION AND THE REINFORCEMENT EFFECT

1.1 Introduction

Empirical evidence suggests that risk preferences evolve over time with personal experiences. This paper studies history-dependent risk aversion and focuses on a well-documented behavioral phenomenon called the *reinforcement effect* (henceforth, RE). The RE says that people become less risk-averse after a good history than after a bad history. I show that this seemingly anomalous behavior occurs whenever a risk preference is history-dependent (in a nontrivial way) and satisfies *monotonicity with respect to first-order stochastic dominance*. Since monotonicity is a reasonable condition, my result provides a theoretical justification for the RE.[1]

To study history-dependent risk aversion and the RE formally, I develop a behaviorally founded model of dynamic choice under risk. The model generalizes standard discounted expected utility in two ways. First, risk preferences are allowed to reflect past risky choices and their payoffs. For example, if an agent is an expected utility maximizer, then the model is a generalization of discounted expected utility, in which the concavity of the agent's utility function changes with her past risky choices and their payoffs. In Section 1.2, I informally introduce this example and illustrate the main result. In this example, the RE arises if the utility function after a good history is less concave than the utility function after a bad history.

Second, risk preferences are allowed to violate expected utility and be non-expected utility preferences such as rank-dependent utility preferences (Quiggin 1982 and Tversky and Kahneman 1992) or disappointment aversion theory preferences (Gul 1991). For example, consider a disappointment aversion theory agent, who distorts probabilities by a real number called a *disappointment parameter*. Then the model is a dynamic version of disappointment aversion theory, in which the disappointment parameter is history-dependent. It turns out that, for a fixed

[1] The RE is documented in the lab experiments of Thaler and Johnson (1990), Ackert et al. (2006), Harrison (2007), and Peng et al. (2013), and in the field studies of Massa and Simonov (2005), Kaustia and Knüpfer (2008), Liu et al. (2010), Malmendier and Nagel (2011), Guiso et al. (2013), and Knüpfer et al. (2014). I carefully discuss three well-known examples of the RE in Section 1.1.1.

utility function, the disappointment parameter dictates the agent's degree of risk aversion. Therefore, the RE arises if the disappointment parameter is smaller after a good history than a bad history. The behavioral foundations of my model and the two special cases above (dynamic versions of expected utility and disappointment aversion theory) are provided.

To illustrate the usefulness of the model, I apply it to the classical Lucas tree model of asset pricing (Lucas 1978) and draw implications of the RE on the dynamics of asset prices. I find that, compared to history-independent models, assets are overpriced when the economy is in a good state, but they are underpriced when the economy is in a bad state. Moreover, I relate the predictions of my model to empirical facts on asset prices. Specifically, my model generates high, volatile, and predictable asset returns, and low and smooth bond returns, consistent with empirical evidence. Indeed, these results are consistent with the *equity premium puzzle* of Mehra and Prescott (1985) and the *risk-free rate puzzle* of Weil (1989).

Let me illustrate the key intuition behind the main result. Suppose there are two periods, today and tomorrow. Today an agent compares two lotteries, a dominant lottery and a dominated lottery. The dominant lottery returns high payoffs with high probability, while the dominated lottery returns high payoffs with low probability. Then *monotonicity* requires that the dominant lottery must be preferred to the dominated one. Tomorrow the agent receives a lottery after choosing one of today's two lotteries. *Dynamic monotonicity*, an extension of monotonicity to the dynamic environment, requires that the dominant lottery is still preferred to the dominated one, independent of tomorrow's lottery.

Suppose the agent's risk preference is history-dependent; that is, the utility of tomorrow's lottery depends on today's payoffs. The RE says that the agent is less risk averse after a good history (high payoff) than after a bad history (low payoff). Since a less risk-averse agent values risky lotteries more than a risk-averse agent does, the RE is equivalent to requiring that tomorrow's lottery is more valuable after a good history than after a bad history.

Now suppose the RE is violated. This implies that tomorrow's lottery is less valuable after a good history than after a bad history. Since the dominant lottery returns high payoffs with high probability, it generates good histories more often than the dominated one does. In other words, tomorrow's lottery is less valuable after the dominant lottery than after the dominated one because good histories make tomorrow's less valuable.

Therefore, when the RE is violated, there is a tradeoff between today and tomorrow: although the dominant lottery is preferred to the dominated one today (the dominant lottery's advantage today), it makes tomorrow's lottery less valuable (disadvantage tomorrow). Therefore, if the dominant lottery's disadvantage tomorrow exceeds its advantage today, then dynamic monotonicity is violated. It turns out that, when tomorrow's lottery is significantly more valuable than today's two lotteries, the dominant lottery's disadvantage tomorrow can be greater than its advantage today.

1.1.1 Examples of the Reinforcement Effect

Using data from the Survey of Consumer Finances from 1960-2007, Malmendier and Nagel (2011) show that individuals' experiences of macroeconomic shocks affect their financial risk taking, consistent with the RE. The authors find that individuals who have experienced low stock market returns throughout their lives report lower willingness to take financial risks, are less likely to participate in the stock market, and invest a lower fraction of their liquid assets in stocks if they participate. These results are robust to controlling for age, year effects, and household characteristics such as wealth, income, and education.

The second example of the RE comes from the experimental study of Thaler and Johnson (1990). The authors run the following experiment, which involves two choice scenarios. In the first scenario, the subjects are asked to choose between a risky lottery $(q, \$y_1, 1-q, \$y_2)$ and a sure outcome $\$x$, right after winning $\$z$ from a lottery Z.[2] In the second scenario, the subjects are asked to choose between a risky lottery $(q, \$y_1+z, 1-q, \$y_2+z)$ and a sure outcome $\$x+z$. According to expected utility theory and considering the final wealth of each situation, there is no difference between the two scenarios. However, Thaler and Johnson found that when $z = \$15$ (a good history), 77% of the subjects prefer the risky option in the first scenario, but only 44% of subjects prefer the risky option in the second scenario.[3] By contrast, when $z = -\$4.50$ (a bad history), only 32% of the subjects prefer the

[2]The vector $(q, \$y_1, 1-q, \$y_2)$ is a lottery that gives $\$y_1$ with probability q and $\$y_2$ with probability $1-q$.

[3]Thaler and Johnson phrase their questions in the following way (see p. 652): "You won x, now choose between a gamble A and a sure outcome B." In the followup experiment by Peng et al. (2013), they phrase their questions in the following two ways (p. 154): i) "You won x from from a gamble X, now choose between a gamble A and a sure outcome B" or ii) "You will get allowance x, now choose between gamble A and a sure outcome B." It turns out that, significantly more subjects choose the gamble A in i) compared to that in ii). Therefore, it is important that the subjects know that they won x by chance.

risky option in the first scenario, but 57% of subjects prefer the risky option in the second scenario.[4] Therefore, the subjects become less risk-averse after $z = \$15$ (a good history) than after $z = -\$4.50$ (a bad history).

The third example of the RE is from the empirical finance literature. It is well known that some market variables move countercyclically in a way that is consistent with the RE (see Cochrane 2011). For example, consider the stock market Sharpe ratio (sometimes called the "price of risk") – the expected excess return of an asset divided by the standard deviation of return. The Sharpe ratio is an important indicator for risk aversion because the first-order condition of intertemporal utility maximization gives that

$$\text{Sharpe ratio} = \text{degree of risk aversion} \times \text{std.dev}(\Delta c) \times cov(\Delta c; R),$$

where Δc is consumption growth and R is asset return. Empirical evidence (e.g., Tang and Whitelaw 2011) suggests that the Sharpe ratio is countercyclical: when the economy is good, the Sharpe ratio is low and when the economy is bad, the Sharpe ratio is high. Therefore, if the standard deviation of consumption growth std.dev(Δc) and the covariance between consumption growth and asset return $cov(\Delta c; R)$ do not vary much over time, the countercyclicity of the Sharpe ratio suggests the presence of the RE. That is, when the economy is good, the Sharpe ratio is low, which implies low risk aversion, but when the economy is bad, the Sharpe ratio is high, which implies high risk aversion. Probably for that reason, some of the most successful models of asset pricing (e.g., Campbell and Cochrane 1999 and Barberis et al. 2001) use countercyclical risk aversion, which is consistent with the RE.[5]

The three examples above illustrate that the RE is a robust and economically relevant notion. The first example illustrates the robustness of the RE to different measures of risk aversion, i.e., willingness to take financial risk, stock market participation, and so on. My explanation of the RE depends on changing preferences for risk, but it is also possible to explain the RE through changes in wealth or beliefs. The first example rules out these two explanations of the RE. First, it unambiguously rules out the wealth effect explanation of the RE (wealthy people are less risk-averse than poor people) since the two scenarios in the experiment generate the same

[4] In the case of $z = \$15$, Thaler and Johnson use numbers $x = \$0$, $q = 0.5$, $y_1 = \$4.5$, and $y_2 = -\$4.5$, and in the case of $z = -\$4.50$, they use numbers $x = \$5$, $q = 0.33$, $y_1 = \$15$, and $y_2 = \$0$.

[5] In fact, empirical finance research directly finds countercyclical risk aversion using an estimator for time-varying risk aversion (e.g., Kim 2014).

final wealth, but the subjects exhibit the RE. Second, it rules out a belief-based explanation (optimistic people are less risk-averse than pessimistic people) since in the experiment, subjects are asked to compare objective lotteries.[6] The third example illustrates that the RE might have market level implications.

1.1.2 Related Literature

Three well-known classes of models generate history-dependent behavior: the Kreps-Porteus model (Kreps and Porteus 1978 and Selden 1978), the Epstein-Zin model (Epstein and Zin 1989 and Weil 1989), and the habit-formation model (Pollak 1970, Constantinides 1990, and Campbell and Cochrane 1999). In all of these models, an agent's current preference is affected by past outcomes or consumption. The main difference between my model and the above three models is that I allow the agent's current preference to be affected by past outcomes and their distributions. Therefore, my model allows the following history-dependent behavior: an agent becomes more risk averse after winning \$10 from a lottery $(\frac{1}{2}, \$10, \frac{1}{2}, \$20)$ and becomes less risk averse after winning \$10 from a lottery $(\frac{1}{2}, \$10, \frac{1}{2}, \$0)$. Moreover, an additive version of the Kreps-Porteus model is a special case of my model (see Section 1.3.3).

The closest paper is Dillenberger and Rozen (2015), which models choice over multi-stage compound lotteries and studies the RE. A two-stage compound lottery is a lottery over simple lotteries, and a t-stage compound lottery is a lottery over $(t-1)$-stage compound lotteries. In Dillenberger and Rozen (2015), risk preferences are affected by a realized $(t-1)$-stage compound lottery as well as unrealized $(t-1)$-stage compound lotteries.[7] Since each multi-stage compound lottery corresponds to a distribution over final outcomes in the future, one key difference is that in their model, risk preferences are affected by past distributions over future outcomes, while in my model, risk preferences are affected by past risky choices and their outcomes. In other words, in their model, agents care about "what might have been" in the future, but in my model, agents care about "what might have been" in the past. Interestingly, Dillenberger and Rozen prove that the RE is a result of internal consistency of changes in risk preferences[8], while I prove that the RE is

[6] Although I consider objective probabilities, in some special cases, the way my model explains the RE is similar to the belief-based explanation. See Section 1.6.4 for more details.

[7] Note that any compound lottery returns a single outcome at the end and so it does not allow intermediate consumptions.

[8] In their model, internal inconsistency may occur because the evaluations of final outcomes affect the current risk preference, while the current risk preference also affects the evaluations of

a result of dynamic monotonicity. Thus, we focus on two different channels that may give rise to the RE. Arguably, the first two examples of the RE indicate that the RE is caused by past risky choices rather than future outcomes. Moreover, their model violates monotonicity while my model is built around monotonicity.[9] Dillenberger and Rozen also apply their model to asset pricing and find volatile and history-dependent prices.

The remainder of the paper is organized as follows. First, I outline the main result and the application to asset pricing in Section 1.2. In Section 1.3, I introduce the model and state the main result. In Section 1.4, I then apply my model to the classical Lucas tree model of asset pricing and draw implications of the RE on the dynamics of asset prices. Two different behavioral foundations for my model are provided in Section 1.5. In Section 1.6, I introduce three special cases of my model (including dynamic versions of expected utility and disappointment aversion theory) and discuss the RE. The proofs are collected in Appendix A.1. Behavioral foundations for dynamic versions of expected utility and disappointment aversion are provided in Appendix A.2.

1.2 Overview of the Main Result and Its Applications

1.2.1 Main Result

In this subsection, I introduce a special case of my model and illustrate the main result in this simple case. Suppose the agent receives a bundle $(X; Z)$, which gives a lottery X today and a lottery Z tomorrow. Suppose the lottery X gives \$$x$ with probability p and \$$y$ with probability $1 - p$, denoted by $X = (p, x, 1 - p, y)$, and the lottery Z gives \$$z$ with probability r and nothing with probability $1 - r$, denoted by $Z = (r, z, 1 - r, 0)$. See Figure 1.1.

Suppose the agent is an expected utility maximizer, and her risk preference is history-dependent. Suppose that the agent has a CRRA utility function $u(t) = t^{\mu}$ where $1 - \mu$ is the agent's degree of risk aversion. Tomorrow the agent's risk attitude changes with today's outcome. In particular, the agent's degree of aversion $1 - \mu$ changes to $1 - \mu(x)$ when the outcome x is realized from the lottery X, and the

final outcomes. Dillenberger and Rozen (2015) show that internal consistency also implies a behavior called the *primacy effect*: an agent who experiences a bad history today and good histories forever is more risk averse than an agent who experiences a good history today and bad histories forever. However, the primacy effect seems not consistent with the following findings of Malmendier and Nagel (2011): most recent experiences have a stronger effect on risk preferences than those in early in life.

[9]See p. 461 of Dillenberger and Rozen 2015, where they explain why monotonicity is violated in their model.

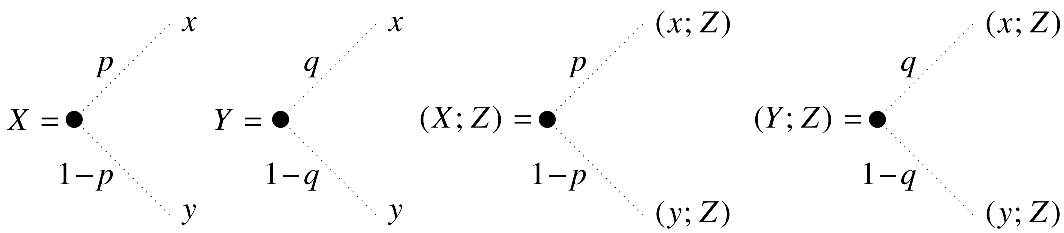

Figure 1.1: Lotteries X and Y and bundles $(X;Z)$ and $(Y;Z)$

certainty equivalent of Z is $\mu_x(Z) = (r\, z^{\mu(x)} + (1-r)\,0)^{\frac{1}{\mu(x)}}$. Therefore, in my model, the utility of the bundle $(X;Z)$ is

$$p\, x^{\mu} + (1-p)\, y^{\mu} + \beta \Big(p\, (\mu_x(Z))^{\mu} + (1-p)\, (\mu_y(Z))^{\mu}\Big),$$

where β is the discount factor. Note that the utility of a sure bundle $(x;z)$ is $x^{\mu} + \beta\, z^{\mu}$.

The key assumption will be a form of monotonicity of risk preferences. Take a lottery $Y = (q, x, 1-q, y)$ with $x > y$ and $q < p$. *Monotonicity* requires that the lottery X must preferred to the lottery Y because X first-order stochastically dominates Y. I define the following extension of monotonicity to the dynamic environment called *dynamic monotonicity*: the bundle $(X;Z)$ must be preferred to the bundle $(Y;Z)$. Intuitively, since the two bundles provide a common lottery Z tomorrow and X first-order stochastically dominates Y, the first bundle must be preferred to the second. In my model, dynamic monotonicity is equivalent to the following inequality:

$$x^{\mu} - y^{\mu} \geq \beta \Big((\mu_y(Z))^{\mu} - (\mu_x(Z))^{\mu}\Big) = \beta\, \big(r^{\frac{\mu}{\mu(y)}} - r^{\frac{\mu}{\mu(x)}}\big)\, z^{\mu}. \tag{1.1}$$

The RE states that the agent is less risk-averse after a good history (when x is realized) than after a bad history (when y is realized); i.e., $\mu(x) \geq \mu(y)$.[10] Equivalently, $\mu_x(Z) \geq \mu_y(Z)$. Therefore, if the RE is violated, when z is large enough, then the RHS of (1.1) exceeds the LHS of (1.1); i.e., dynamic monotonicity is violated.

As I show later, the above argument is true under an assumption called *non-triviality*. It requires that if the agent's risk preference is history-dependent; that is, $\mu_x(Z) \neq \mu_y(Z)$ for some Z, then there exists Z^* such that $|u(\mu_x(Z^*)) - u(\mu_y(Z^*))|$

[10]Note that in this definition, the RE does not rule out the standard case $\mu = \mu(x) = \mu(y)$.

is large enough. In three special cases of my model discussed in Section 1.6, non-triviality is equivalent to the condition that the agent's Bernoulli utility function u is unbounded. In the above example, nontriviality is satisfied since $u(z) = z^{\mu}$ is unbounded.

I prove the result in a general case in Section 1.3.2. In particular, the certainty equivalents $\mu_x(Z)$ and $\mu_y(Z)$ of Z not only depend on the outcomes x and y, but also depend on lotteries X and Y. Moreover, the agent's risk preference can be a non-expected utility preference such as a rank-dependent utility preference or disappointment aversion theory preference.

1.2.2 Application to Asset Pricing

I apply my model to the Lucas tree model of asset pricing (Lucas 1978) and show that the RE has important implications for asset prices. In Section 1.4, I show that I can obtain predictable, high, and volatile asset returns, and low and smooth bond returns, consistent with empirical data. I use a simple model of an economy that follows a two-state Markov process in which the states, high and low, are persistent. The price of an asset increases when agents become less risk-averse, since less risk-averse agents value the asset more than risk-averse agents. Therefore, the first implication of the RE is as follows: the asset is overpriced in a high state, but it is underpriced in a low state, compared to models with history-independent risk aversion. In the LHS of Figure 1.2, the price-dividend ratio is plotted against the degree of risk aversion. The blue (dotted) lines illustrate the price-dividend ratio in my history-dependent model (discounted expected utility). Note that in a high state, the price-dividend ratio is above the dotted line; i.e., assets are overpriced, while in a low state, the price-dividend ratio is below the dotted line; i.e., assets are underpriced.

The second implication of the RE is that there is asymmetry between high and low states. This implication relies on the following property of the model: the degree of mispricing (overpricing and underpricing) increases as agents become more risk-averse. Therefore, as illustrated in the LHS of Figure 1.2 in high states, the price-dividend ratio increases as agents become more risk-averse because the degree of overpricing increases. However, in low states, the price-dividend ratio decreases as agents become more risk-averse because the degree of underpricing increases.

I now discuss the dynamics of asset prices using the RHS of Figure 1.2. Since

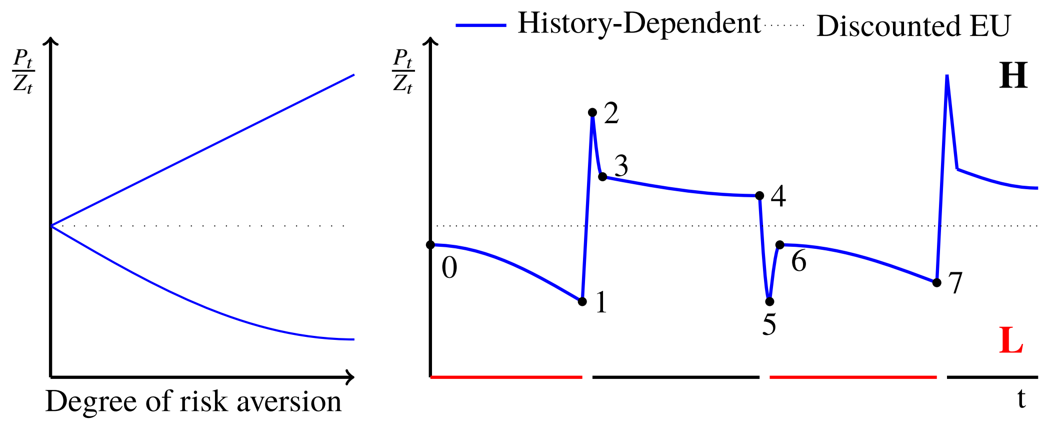

Figure 1.2: The Price-Dividend Ratio

states are persistent, low states continue for a while; after that, high states continue for a while, and so on. Suppose the economy is in a low state first. By the second implication of the RE, in a low state, asset prices decrease as agents become more risk-averse (Point 0 to Point 1). But, when the economy recovers, asset prices overshoot because underpricing turns to overpricing as the state changes (Point 1 to Point 2). Moreover, the overshooting is large because the degrees of risk aversion and overpricing are very high after a long period of low states.

Now the economy is in a high state. By the second implication of the RE, asset prices decrease because the degrees of risk aversion and overpricing decrease (Point 2 to Point 4). Moreover, when the economy declines, asset prices drop because overpricing turns to underpricing as the state changes (Point 4 to Point 5). However, the drop is not as large as the overshooting since the degree of risk aversion decreased after a long period of high states (between Point 3 to Point 4).

1.3 Model

1.3.1 Basic Setup and Model

I now introduce a model of dynamic choice under risk that generalizes standard discounted expected utility. There are two periods, today and tomorrow.[11] An agent evaluates intertemporal consumption lotteries, which gives a lottery today and another lottery tomorrow depending on the realization of today's lottery. First, I define what a lottery is. Let $\mathbb{R}_+$ be the set of all (monetary) outcomes. For any set

[11]The number of periods is not important. In fact, I use an infinite horizon version of the model for the application to asset pricing in Section 1.4.

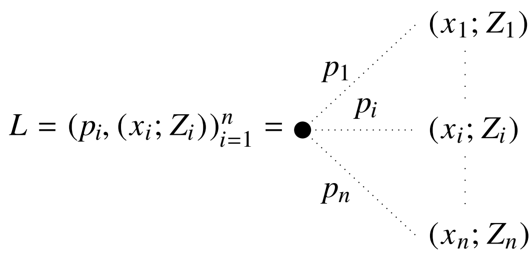

Figure 1.3: Intertemporal Consumption Lotteries

$\mathcal{A}$, I define the set of all finite lotteries over $\mathcal{A}$, denoted by $\Delta(\mathcal{A})$. That is,

$$\Delta(\mathcal{A}) \equiv \{A = (p_1, a_1, \ldots, p_n, a_n) \in ([0,1] \times \mathcal{A})^n \text{ with } \sum_{i=1}^{n} p_i = 1 \text{ for some } n\},$$

where a lottery $A = (p_i, a_i)_{i=1}^n$ gives a_i with probability p_i for each i. I call elements of $\Delta(\mathbb{R}_+)$ simple lotteries.

I focus on elements of $\mathscr{L} = \Delta(\mathbb{R}_+ \times \Delta(\mathbb{R}_+))$ called intertemporal consumption lotteries. Figure 1.3 illustrates an intertemporal consumption lottery $L = (p_i, (x_i; Z_i))_{i=1}^n$, an element of $\mathscr{L} = \Delta(\mathbb{R}_+ \times \Delta(\mathbb{R}_+))$. The lottery L gives an intertemporal consumption bundle $(x_i; Z_i)$ with probability p_i for each i, and the bundle $(x_i; Z_i)$ gives an outcome x_i today and gives a simple lottery Z_i tomorrow. In other words, the agent receives a simple lottery $X = (p_i, x_i)_{i=1}^n$ today and receives another simple lottery Z_i tomorrow after an outcome x_i is realized from the lottery X. I call a pair (x_i, X) a history when $x_i \in supp(X) \equiv \{x_1, \ldots, x_n\}$, the support of X. So the agent receives a simple lottery Z_i after the history (x_i, X).

The main focus of this paper is to study how the value of a lottery Z_i is affected by the history (x_i, X). I assume that there exists a utility function $W\colon \mathscr{L} \to \mathbb{R}_+$ for intertemporal consumption lotteries. After specifying W, I analyze and compare risk preferences after different histories (x, X) and (x', X) with $x > x'$.

The benchmark model is *discounted expected utility* (henceforth, DEU). In DEU, the utility of an intertemporal consumption lottery $(p_i, (x_i; Z_i))_{i=1}^n \in \mathscr{L}$ is

$$W\left((p_i, (x_i; Z_i))_{i=1}^n\right) = \mathbb{E}[u(X)] + \beta\, \mathbb{E}_X\left[\mathbb{E}[u(Z_i)]\right] = \sum_{i=1}^{n} p_i\left(u(x_i) + \beta\, \mathbb{E}[u(Z_i)]\right) \tag{1.2}$$

where $u\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a Bernoulli utility function, $\beta \in (0,1)$ is a discount factor, $\mathbb{E}$ is the standard expectation operator, and $\mathbb{E}_X$ is the expectation operator with respect to the distribution of X. Note the following four properties of DEU:

1. *Simple Expected Utility:* the agent uses expected utility theory when she evaluates simple lotteries;

2. *History Independence:* tomorrow's risk preference is the same as today's risk preference; i.e, she uses the same Bernoulli utility function u to calculate expected utilities $\mathbb{E}[u(X)]$ and $\mathbb{E}[u(Z_i)]$;

3. *Discounted Utility:* the agent uses discounted utility theory when she aggregates utilities of today and tomorrow; i.e., the utility of $(x; z)$ is $u(x) + \beta\, u(z)$;

4. *Expected Utility Aggregator:* once tomorrow's lotteries are evaluated ($\mathbb{E}[u(Z_i)]$ is calculated for each i), she aggregates them using expected utility theory ($\mathbb{E}_X\left[\mathbb{E}[u(Z_i)]\right] = \sum_{i=1}^{n} p_i\, \mathbb{E}[u(Z_i)]$).

I generalize DEU by weakening the first two of the above four properties of DEU while retaining the latter two. Indeed, weakening the second property, history independence, is essential for analyzing history-dependent risk preferences. However, I weaken the first property, simple expected utility, not only to demonstrate the generality of the main result, but also to include well-known non-expected utility models such as rank-dependent utility theory (Quiggin 1982 and Tversky and Kahneman 1992) and disappointment aversion theory (Gul 1991). In fact, to demonstrate the usefulness of the model, I apply a dynamic version of disappointment aversion theory to asset pricing in Section 1.4. The role of the third property, discounted utility, will be discussed in Sections 1.3.2-3.

Therefore, in my model, the agent uses a function $V_0 : \Delta(\mathbb{R}_+) \to \mathbb{R}_+$ to evaluate simple lotteries today, but she uses a history-dependent function $V_{(x,X)} : \Delta(\mathbb{R}_+) \to \mathbb{R}_+$ to evaluate simple lotteries tomorrow after a history (x, X). Once simple lotteries are evaluated, the agent aggregates them using the last two of the above four properties of DEU. Formally, I study the following model.

Definition 1 (History-Dependent Model) *A utility W is a* history-dependent model *if there exists a triplet $(V_0, \beta, \{V_{(x,X)}\})$ such that the utility of an intertemporal consumption lottery $(p_i, (x_i; Z_i))_{i=1}^n \in \mathscr{L}$ can be represented as*

$$W\left((p_i, (x_i; Z_i))_{i=1}^n\right) = V_0(X) + \beta \sum_{i=1}^{n} p_i\, V_{(x_i, X)}(Z_i). \tag{1.3}$$

Indeed, the history-dependent model (1.3) reduces to DEU when $V_0(Z) = V_{(x,X)}(Z) = \mathbb{E}[u(Z)]$ for some u. Two different behavioral foundations for the history-dependent model (1.3) are provided in Section 1.5. For notational simplicity, when there is no danger of confusion, I also call $\{V_{(x,X)}\}$ a history-dependent model.[12]

In order to discuss a relation between the RE and monotonicity, I assume the following property of continuity of $V_{(x,X)}$ with respect to X.

Definition 2 *(Right-continuity) A history-dependent model $\{V_{(x,X)}\}$ is right-continuous if for any lotteries $\{X^n\}_{n=1}^{\infty}$ and X^* such that $supp(X^n) = supp(X^*)$ and X^n first-order stochastically dominates X^* for each n, $X^n \xrightarrow{W} X^*$ implies $V_{(x,X^n)} \xrightarrow{U} V_{(x,X^*)}$.*[13]

Roughly speaking, the above assumes that changes in risk preferences caused by a lottery X are not so extreme as long as x is fixed. Indeed, right-continuity is satisfied when $V_{(x,X)}$ is independent of X. Moreover, the following special case of (1.3), in which $\{V_{(x,X)}\}$ is represented by two step functions, satisfies right-continuity: for each $x \in \mathbb{R}_+$, there are two functions $\{\mathcal{V}_x^0, \mathcal{V}_x^1\}$ such that

$$V_{(x,X)} = \begin{cases} \mathcal{V}_x^1 \text{ when } x > \mu_0(X) \\ \mathcal{V}_x^0 \text{ when } x \leq \mu_0(X), \end{cases} \tag{1.4}$$

where $\mu_0(X) = V_0^{-1}(V_0(X))$ is the certainty equivalent of X. In (1.4), when the outcome x is greater than the certainty equivalent of the lottery X, the agent is content and uses $\mathcal{V}_x^1$, but when the outcome is equal or less than the certainty equivalent, the agent is disappointed and uses $\mathcal{V}_x^0$. In Section 1.4, I apply a special case of (1.4) to asset pricing. I can also consider general cases of (1.4) in which

[12] One might find it conflicting since I assume that risk preferences over compound lotteries are expected utility preferences (expected utility aggregator) while risk preferences over simple lotteries are possibly non-expected utility preferences (V_0). Non-expected utility preferences are motivated by violations of the independence axiom such as the Allais Paradox. However, it turns out that, the compound independence axiom, a counterpart of the independence axiom for compound lotteries, is less likely to be violated compared to the independence axiom (Segal 1990, Luce 1990, and Camerer and Ho 1994). For that reason, I use expected utility aggregator. Moreover, the main result is still true if I relax expected utility aggregator to non-expected utility aggregators such as disappointment aversion aggregator or rank-dependent utility aggregator.

[13] The notation $X^n \xrightarrow{\mathrm{W}} X^*$ denotes the weak convergence; that is, $\mathbb{E}f(X^n) \to \mathbb{E}f(X^*)$ for any bounded continuous function f. Moreover, $V_{(x,X^n)} \xrightarrow{\mathrm{U}} V_{(x,X^*)}$ denotes the uniform convergence; that is, for any $\delta > 0$, there exists n^* such that for any $n > n^*$, $|V_{(x,X^n)}(Z) - V_{(x,X^*)}(Z)| < \delta$ for any $Z \in \Delta(\mathbb{R}_+)$. I also can assume left-continuity instead of right-continuity.

there are more than 2 step functions. Note that such models generalize a model with state-dependent risk aversion, in which states are endogenously determined by histories.

1.3.2 The Reinforcement Effect and Monotonicity

I now turn to the discussion of the RE and monotonicity. The RE states that the agent becomes less risk-averse after a good history (x, X) than after a bad history (x', X) where $x > x'$. Since I only focus on changes in risk preferences (not on changes in time preferences), I assume that the utility of money does not change over time. In other words, I assume *time consistency*; that is, $V_0(z) = V_{(x,X)}(z)$ for any $z \in \mathbb{R}_+$. Note that time consistency generalizes the third property, discounted utility, of the aforementioned four properties of DEU. To see this, let $u(z) \equiv V_0(z)$ for any $z \in \mathbb{R}_+$. Then in the history-dependent model (1.3), the utility of a deterministic bundle $(x; z)$ is $u(x) + \beta u(z)$.

Under time consistency, comparisons among V_0, $V_{(x,X)}$, and $V_{(x',X)}$ capture changes in risk attitude. Specifically, the RE is equivalent to stating $V_{(x,X)}$ values risky lotteries more than $V_{(x',X)}$ does. Formally,

Definition 3 (Reinforcement Effect) *A history-dependent model $\{V_{(x,X)}\}$ exhibits the* reinforcement effect *if for any lottery $X \in \Delta(\mathbb{R}_+)$ and $x, x' \in supp(X)$ with $x > x'$, $V_{(x,X)}(Z) \geq V_{(x',X)}(Z)$ for any $Z \in \Delta(\mathbb{R}_+)$.*

The first-order stochastic dominance is a fundamental concept in risky choice. It defines when one lottery is unambiguously better than another lottery. A lottery X first-order stochastically dominates a lottery Y if for any $z \in \mathbb{R}_+$, the probability that the agent receives at least z from X is not smaller than that from Y; i.e.,

$$\sum_{i=1}^{n} p_i \, \mathbb{1}\{x_i \geq z\} \geq \sum_{j=1}^{m} q_j \, \mathbb{1}\{y_j \geq z\}.$$

Monotonicity of risk preferences requires that if X first-order stochastically dominates Y, then X must be preferred to Y. Monotonicity is appealing because it is rarely violated (e.g., see Hey 2001) and its violation may lead to the "Dutch book" argument (e.g., see Machina 1989). Moreover, in expected utility, monotonicity is equivalent to the monotonicity of the agent's Bernoulli utility function. I now extend monotonicity to the dynamic environment.[14]

[14] A necessary and sufficient condition for a general (differentiable) non-expected utility function to satisfy monotonicity is given in Machina (1982).

Definition 4 (Dynamic Monotonicity) *For any $X, Y \in \Delta(\mathbb{R}_+)$, if X first-order stochastically dominates Y, then $(X; Z)$ must be preferred to $(Y; Z)$ for any $Z \in \Delta(\mathbb{R}_+)$.*

Dynamic monotonicity says that if two intertemporal consumption lotteries $(X; Z)$ and $(Y; Z)$ share a common lottery Z tomorrow, then monotonicity in the static environment must be satisfied. Dynamic monotonicity is a weak version of monotonicity in which intertemporal consumption lotteries give different lotteries after different histories (in line with Segal 1990), but dynamic monotonicity will be enough for my purpose.

Another key assumption is *nontriviality*, which requires that if two functions $V_{(x,X)}$ and $V_{(x',X)}$ are different, then they must be significantly different. Specifically, if $V_{(x,X)}(Z) > V_{(x',X)}(Z)$ for some Z, then there exists Z^* such that $V_{(x,X)}(Z^*) - V_{(x',X)}(Z^*)$ is large enough. Formally,

Assumption 1 (Nontriviality) *Take any lottery $X \in \Delta(\mathbb{R}_+)$ and $x, x' \in supp(X)$. If there exists $Z \in \Delta(\mathbb{R}_+)$ such that $V_{(x,X)}(Z) > V_{(x',X)}(Z)$, then for any $M > 0$, there exists $Z^* \in \Delta(\mathbb{R}_+)$ such that*

$$V_{(x,X)}(Z^*) - V_{(x',X)}(Z^*) > M.$$

I briefly argue that dynamic monotonicity rather than nontriviality is mostly responsible for the RE for two reasons. First, nontriviality treats good and bad histories symmetrically. Second, in the three examples of the history-dependent model (1.3) discussed in Section 1.6, I show that nontriviality is essentially unrelated to history dependence. Specifically, in these three examples, I show that nontriviality is equivalent to the condition $u(+\infty) = +\infty$, which is unrelated to histories (x, X) and (x', X). I now state the main result.

Theorem 1 (Dynamic Monotonicity Implies the Reinforcement Effect) *If a history-dependent model $\{V_{(x,X)}\}$ satisfies right-continuity, dynamic monotonicity, and nontriviality, then it exhibits the reinforcement effect.*

In the introduction, I argued that the RE is well documented in empirical studies. Since dynamic monotonicity is a natural behavior, Theorem 1 provides a justification for why the RE is well observed. The proof of Theorem 1 is in Appendix A.1. In the proof, in fact, I only use a weak version of dynamic monotonicity, in which X dominates Y in the following obvious manner: $X = \big(p_i, x_i, p_j, x_j, (p_k, x_k)_{k\neq i,j}\big)$ and $Y = \big(p_i - \epsilon, x_i, p_j + \epsilon, x_j, (p_k, x_k)_{k\neq i,j}\big)$ where $x_i > x_j$.

The main idea behind Theorem 1 is that when the RE is violated, bad histories generate higher utilities than good histories do. Hence, since Y generates bad histories more often than X does, the utility of Z is higher after Y than after X. Therefore, if there exists Z such that the advantage of Z of $(Y; Z)$ over Z of $(X; Z)$ exceeds the advantage of X over Y, then dynamic monotonicity is violated. In fact, nontriviality guarantees that such Z exists. Therefore, under nontriviality, dynamic monotonicity implies the RE.

1.3.3 Time Consistency and Relation to Existing Models

My model (1.3) satisfies time consistency: $V_0(z) = V_{(x,X)}(z)$ for any $z \in \mathbb{R}_+$. Because of the additive structure of (1.3), time consistency has the following interpretation: the utility $V_{(x,X)}(Z)$ of Z is history-independent when Z is a deterministic lottery (i.e., $\frac{\partial V_{(x,X)}(z)}{\partial x} = \frac{\partial V_0(z)}{\partial x} = 0$).[15] As I argue in this subsection, time consistency is an important difference between my model and other history-dependent models. In the habit-formation model (e.g., Pollak 1970 and Constantinides 1990), the utility of a deterministic bundle $(x; z)$ is

$$u(x) + \beta\, u(z - \alpha\, x) \text{ for some } \alpha \in (0, 1).$$

The habit formation model not only violates time consistency ($\frac{\partial u(z-\alpha x)}{\partial x} \neq 0$), but also deviates from discounted utility.

The Kreps-Porteus model (Kreps and Porteus 1978 and Selden 1978) does not necessarily violate time consistency. In the Kreps-Porteus model, the utility of $(p_i, (x_i; Z_i))_{i=1}^n$ is

$$\mathbb{E}_X\, V(x_i, u_{x_i}^{-1}(\mathbb{E}\, u_{x_i}(Z_i))), \tag{1.5}$$

where V is a time aggregator (see Koopmans 1960) and $u_{x_i}^{-1}(\mathbb{E}\, u_{x_i}(Z_i))$ is the certainty equivalent of Z_i after x_i. In fact, my model generalizes the following additive version of the Kreps-Porteus model (1.5), in which $V(x, z) = u(x) + \beta\, u(z)$: the utility of $(p_i, (x_i; Z_i))_{i=1}^n$ is

$$\mathbb{E}_X \Big(u(x_i) + \beta\, u(u_{x_i}^{-1}(\mathbb{E}\, u_{x_i}(Z_i)))\Big). \tag{1.6}$$

[15] This separability property is captured by a behavioral axiom called *weak separability between today and tomorrow*. See Section 1.5 for the behavioral foundations of (1.3).

Since the additive Kreps-Porteus model (1.6) is a special case of the history-dependent model (1.3), I can derive the following corollary of Theorem 1 for (1.6).

Corollary 1 *Suppose dynamic monotonicity and nontriviality are satisfied. If there are utility functions* $\{u_x\}_{x\in\mathbb{R}_+}$ *that satisfy (1.6), then for any* $x, x' \in \mathbb{R}_+$ *with* $x > x'$, $u_{x'}$ *is more concave than* u_x*; that is, there exists a concave function* $f_{x,x'}$ *such that* $u_{x'} = f_{x,x'} \circ u_x$.

However, the Kreps-Porteus model (1.5) does not nest my model (1.3) because I allow i) u_{x_i} to be dependent on X and ii) risk preferences to violate expected utility.

The Epstein-Zin model (Epstein and Zin 1989 and Weil 1989) extends the Kreps-Porteus model (1.5) by allowing risk preferences to violate expected utility, but they are history-independent. Hence, for example, the utility of a bundle $(x; Z)$ is $V(x, \mu(Z))$ where the certainty equivalent $\mu(Z)$ does not necessarily follow expected utility, but is independent of x. The Epstein-Zin model can generate history-dependent behavior because of the time aggregator V. In other words, if V is additive-separable, then the Epstein-Zin model is history-independent. For that reason, the only intersection between my model (1.3) and the Epstein-Zin model is DEU. For example, consider the most popular version of the Epstein-Zin model, in which $V(x, z) = (x^\rho + \beta z^\rho)^{\frac{\alpha}{\rho}}$ and $\mu(Z) = (\mathbb{E}Z^\alpha)^{\frac{1}{\alpha}}$ for some α, ρ. That is, the utility of $(p_i, (x_i; Z_i))_{i=1}^n$ is

$$\mathbb{E}_X [x_i^\rho + \beta(\mathbb{E}[Z_i^\alpha])^{\frac{\rho}{\alpha}}]^{\frac{\alpha}{\rho}}. \tag{1.7}$$

When $\rho = \alpha$, (1.7) reduces to DEU. However, when $\rho \neq \alpha$, Z_i is not separable from x_i even if Z_i is a deterministic lottery (i.e., $\frac{\partial V(x_i, \mu(Z_i))}{\partial x_i} \neq 0$).[16]

1.4 Application: The Lucas Tree Model with HDDA Agents

In this section, I study the implications of the RE on the dynamics of asset prices. I consider a special case of my model, a dynamic version of the disappointment aversion theory of Gul (1991). I apply this model to the classical Lucas tree model of asset pricing.[17] I first describe the economy and preferences.

[16]In fact, the Epstein-Zin model (1.7) violates the aforementioned *weak separability between today and tomorrow* when $\rho \neq \alpha$.

[17]A dynamic version of expected utility is directly applicable to this environment, but finding a closed-form solution is difficult because the Bernoulli utility function is history-dependent. As I show, finding a closed-form solution is easy in the case of disappointment aversion theory because the Bernoulli utility function is history-independent while the disappointment parameter is history-dependent.

The Economy: There is one unit of identical agents who live forever. An asset produces a stochastic dividend stream $\{Z_t\}$. There are two states, H and L. At each date $t \in \{1, 2, \ldots\}$, for a given state $s_{t-1} \in \{H, L\}$ and dividend z_{t-1}, Z_t takes the value of z_t^H with probability $\rho(H|s_{t-1})$ and z_t^L with probability $\rho(L|s_{t-1})$ where $z_t^H > z_t^L$. That is,

$$Z_t = (\rho(H|s_{t-1}), z_t^H, \rho(L|s_{t-1}), z_t^L).$$

I assume that states are persistent; i.e., $\rho(H|H) = \rho(L|L) = \rho > \frac{1}{2}$. In each period, after z_t is realized from Z_t, agents trade the consumption good, c_t, and the asset in a competitive spot market at price p_t. For a given state s_t, each agent faces the following budget constraint:

$$c_t^s + p_t\, x_{t+1}^s = (z_t^s + p_t)\, x_t,$$

where x_t^s is the asset demand at date t ($x_0 = 1$).

History-Dependent Disappointment Aversion: The preferences of agents in this Lucas economy are defined by a dynamic version of the disappointment aversion theory of Gul (1991).[18]

The disappointment aversion theory is a one-parameter generalization of expected utility theory. In the disappointment aversion theory, an agent overweights probabilities of small outcomes and underweights probabilities of large outcomes. The degree of such probability distortion is summarized by a single parameter δ_0, the *disappointment parameter*. Namely, the utility of $C = (\rho, c^H, 1 - \rho, c^L)$ with $c^H > c^L$ is

$$u(\mu(C|\delta_0)) = \frac{\rho\, u(c^H) + (1 - \rho)(1 + \delta_0)\, u(c^L)}{\rho + (1 - \rho)(1 + \delta_0)}, \tag{1.8}$$

where $\mu(C|\delta_0)$ is the certainty equivalent of C. For a fixed utility function u, the disappointment parameter captures the degree of risk aversion; that is, a higher disappointment parameter implies a higher degree of risk aversion.

In *history-dependent disappointment aversion* (HDDA), the agent's disappointment parameter changes with her experiences. In particular, tomorrow, the agent's initial disappointment parameter δ_0 changes to $\delta(x, X)$ after a history (x, X). Then the utility of C after (x, X) is

$$u(\mu(C|\delta(x, X))) = \frac{\rho\, u(c^H) + (1 - \rho)(1 + \delta(x, X))\, u(c^L)}{\rho + (1 - \rho)(1 + \delta(x, X))}. \tag{1.9}$$

[18] See Section 1.6 for a more detailed discussion of disappointment aversion theory.

In other words, the agent's distorted probabilities change with her experiences. Since the disappointment parameter captures the degree of risk aversion, the RE is equivalent to the following simple condition:

$$\delta(x, X) \leq \delta(x', X) \text{ when } x > x'.$$

Preferences: In this Lucas economy, each agent has a HDDA preference with an initial disappointment parameter $\delta_0 > 0$. Consider a special case of the representation with step functions (1.4), in which the disappointment parameter $\{\delta_t\}$ follows the following transition law: for some $\alpha \in (0, 1)$,

$$\delta_{t+1}^s = \begin{cases} \alpha \cdot \delta_t & \text{when } s_t = H \\ \frac{\delta_t}{\alpha} & \text{when } s_t = L. \end{cases}$$

Since I fix the utility function u, δ_t dictates the agents' degree of risk aversion and the transition law captures the RE.

At each date t, for a given (z_{t-1}, s_{t-1}), the disappointment parameter δ_t, and the asset demand x_t, the agents' continuation value of the asset is $V(z_{t-1}, s_{t-1}, \delta_t, x_t)$. So agents solve the following Bellman equation for given $\{Z_t\}$, $\{p_t\}$, and $\{x_t\}$:

$$\begin{aligned} V(z_{t-1}, s_{t-1}, \delta_t, x_t) &= \max_{C_t}\{u(\mu(C_t|\delta_t)) + \beta E_{Z_t} V(Z_t, S_t, \delta_{t+1}^s, x_{t+1}^s)\} \\ \text{s.t} \quad & \quad c_t^s + p_t^s\, x_{t+1}^s = (z_t^s + p_t^s)\, x_t, \end{aligned}$$

where $\mu(C_t|\delta_t)$ is the certainty equivalent of $C_t = (\rho(H|s_{t-1}), c_t^H, \rho(L|s_{t-1}), c_t^L)$ for a given disappointment parameter δ_t. In order to find a closed-form solution, I assume $u(x) = \log(x)$.

To emphasize history dependence and the RE, I compare HDDA with history-independent disappointment aversion (henceforth, HIDA, in which the disappointment parameter is constant) and discounted expected utility (DEU).

1.4.1 Optimal Consumption Profile

Let me start with the optimal consumption path for given processes $\{Z_t\}$ and $\{p_t\}$. For given s_t, agents solve the following problem:

$$\max_{c_t^s} \pi(s_t) \log(c_t^s) + \beta\, \rho(s_t)\, V(z_t^s, s_t, \delta_{t+1}^s, \frac{(z_t^s + p_t^s)x_t - c_t^s}{p_t^s}),$$

where $\pi(s_t)$ is the distorted probability of $\rho(s_t)$ such that

$$\pi(s_t) = \begin{cases} \frac{\rho(H)}{1+(1-\rho(H))\,\delta_t} & \text{when } s_t = H \\ \frac{\rho(L)(1+\delta_t)}{1+\rho(L)\delta_t} = \frac{(1-\rho(H))(1+\delta_t)}{1+(1-\rho(H))\delta_t} & \text{when } s_t = L, \end{cases}$$

and $V(z_t^s, s_t, \delta_{t+1}^s, \frac{(z_t^s+p_t^s)x_t - c_t^s}{p_t^s})$ is the continuation value of the asset for a given z_t^s and s_t.

Note that when $s_t = H$, the probability $\rho(H)$ is underweighted; that is, $\pi(H) < \rho(H)$, but when $s_t = L$, the probability $\rho(L)$ is overweighted; that is, $\pi(L) > \rho(L)$. Let me denote the size of the probability distortion by

$$\lambda_t^s = \begin{cases} \frac{\rho(H)}{\pi(H)} = 1 + (1 - \rho(H))\delta_t & \text{when } s_t = H \\ \frac{\rho(L)}{\pi(L)} = \frac{1+(1-\rho(H))\delta_t}{1+\delta_t} & \text{when } s_t = L. \end{cases}$$

One important feature of HDDA is that agents use distorted probabilities $\pi(s_t)$ when they evaluate the expected utility of today's consumption, but they use objective probabilities $\rho(s_t)$ when they evaluate the expected continuation value of the asset. Therefore, the size of the probability distortion λ_t plays an important role in my analysis. Since δ_t dictates the degree of risk aversion, $\lambda_t^H = 1 + (1 - \rho(H))\delta_t$ increases and $\lambda_t^L = \frac{1+(1-\rho(H))\,\delta_t}{1+\delta_t}$ decreases as agents become more risk-averse (as δ_t increases). I can interpret δ_t as the degree of pessimism because as δ_t increases, agents pay more attention to bad states. As δ_t increases, agents become more pessimistic; consequently, they think a high state will occur with very low probability. Moreover, the less likely the high state is, the larger the distortion is; that is, both $\lambda_t^H = 1 + (1 - \rho(H))\delta_t$ and $\lambda_t^L = \frac{1+(1-\rho(H))\,\delta_t}{1+\delta_t}$ are decreasing in $\rho(H)$. I now solve the optimal consumption path.

Optimal Consumption: For given $\{Z_t\}$ and $\{p_t\}$, the optimal consumption level at time t is

$$c_t^{*s} = \frac{(p_t^s + z_t^s)x_t}{1 + \frac{\beta}{1-\beta} \cdot \lambda_t^s}.$$

Note that $(p_t^s + z_t^s)x_t$ is the agents' wealth at time t. In DEU, agents consume $\frac{1}{1+\frac{\beta}{1-\beta}}$ fraction of their wealth independently of the current state since $\lambda_t = 1$. However, in HDDA, agents consume $\frac{1}{1+\frac{\beta}{1-\beta}\cdot\lambda_t^s}$ fraction of their wealth. Since $\lambda_t^H > 1 > \lambda_t^L$, I have

$$c_t^H(DEU) = \frac{(p_t^H + z_t^H)x_t}{1 + \frac{\beta}{1-\beta}} > c_t^{*H} = \frac{(p_t^H + z_t^H)x_t}{1 + \frac{\beta}{1-\beta} \cdot \lambda_t^H} >$$

$$> c_t^{*L} = \frac{(p_t^L + z_t^L)x_t}{1 + \frac{\beta}{1-\beta} \cdot \lambda_t^H} > c_t^L(DEU) = \frac{(p_t^L + z_t^L)x_t}{1 + \frac{\beta}{1-\beta}}.$$

Therefore, in HDDA, for given $\{p_t\}$, the optimal consumption path is much smoother than that in DEU. The main reason is that the HDDA agents are pessimistic in general and so they will save more in high states.

1.4.2 Market Clearing and Equilibrium Asset Price

I now solve the equilibrium price $\{p_t\}$. In order for the prices to clear the asset market, at each time t, I must have the conditions:

Market Clearing: $x_t = x^* = 1$ and $C_t = Z_t$.

Now it is simple to find the equilibrium asset prices.

Equilibrium Asset Price: For given process $\{Z_t\}$, the equilibrium asset price at time t is

$$p_t^{*s} = \frac{\beta}{1-\beta} \cdot \lambda_t^s \cdot z_t^s.$$

In DEU, the price-dividend ratio $\frac{p_t}{z_t}$ is constant $\frac{\beta}{1-\beta}$. In HDDA, the price-dividend ratio is $\frac{\beta}{1-\beta}$ times the size of the probability distortion, so

$$p_t^{*H} = \frac{\beta}{1-\beta} \cdot \lambda_t^H \cdot z_t^H > p_t^H(DEU) = \frac{\beta}{1-\beta} \cdot z_t^H > p_t^L(DEU) = \frac{\beta}{1-\beta} \cdot z_t^L > p_t^{*L} = \frac{\beta}{1-\beta} \cdot \lambda_t^L \cdot z_t^L.$$

In HDDA, the asset is mispriced by the size of the probability distortion. In a high state, the asset is overpriced since agents undervalue the expected utility of today's consumption compared to the expected continuation value of the asset. In a low state, the asset is underpriced since agents overvalue the expected utility of today's consumption compared to the expected continuation value of the asset. In Figure 1.4, the price-dividend ratio $\frac{p_t}{z_t}$ is plotted against δ_t. As I described earlier, in high states the price-dividend ratio in HDDA is greater than the price-dividend ratio in DEU (the dotted line), and in low states the price-dividend ratio in HDDA is smaller than the price-dividend ratio in DEU. In a high state, the asset price can take two forms because $\rho(H)$ is either ρ or $1-\rho$ depending on the previous state s_{t-1}: $p_t^H(L)$ is the price in a high state after a low state and $p_t^H(H)$ is the price in a high state after a high state. Similarly, the asset price in a low state can take two forms, $p^L(L)$ and $p^L(H)$. Figure 1.4 illustrates that the asset prices for high states increase while the asset prices for low states decrease as agents become more risk-averse (as δ_t increases). Now I turn to the dynamics of the price-dividend ratio in HDDA and compare it with HIDA.

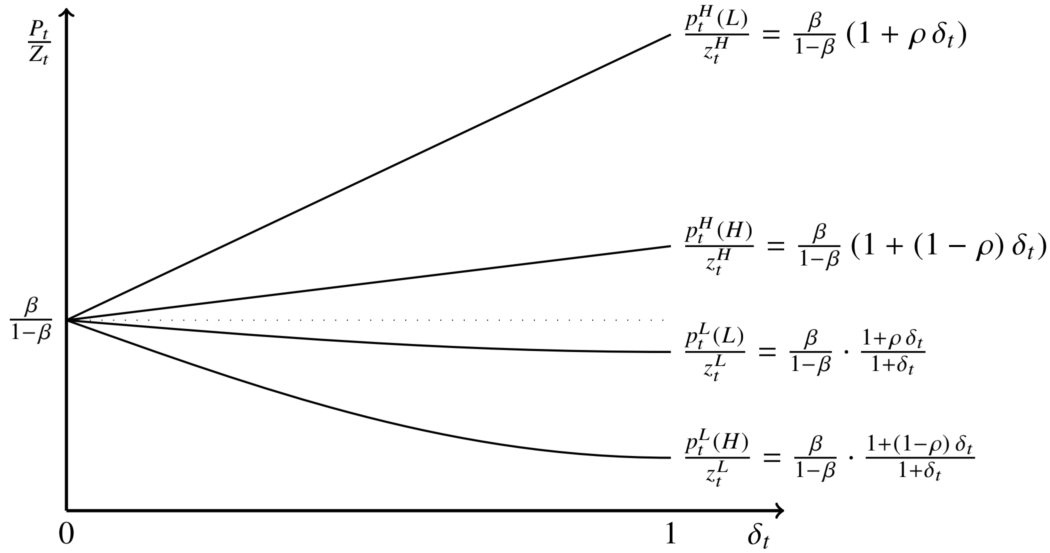

Figure 1.4: The Price-Dividend Ratio and the Disappointment Parameter

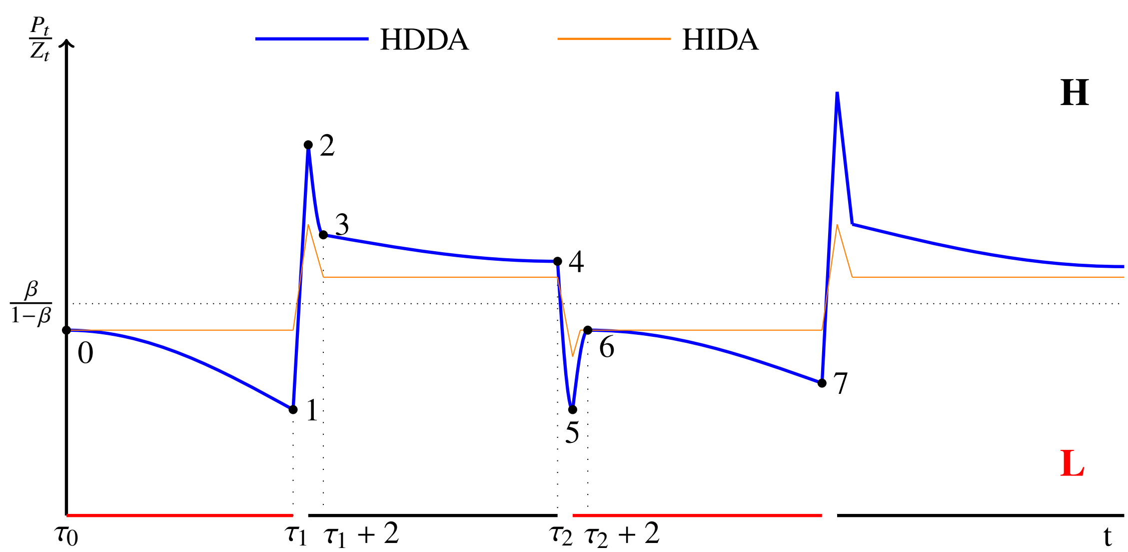

Figure 1.5: The Dynamics of the Price-Dividend Ratio

1.4.3 The Dynamics of the Price-Dividend Ratio

Figure 1.5 illustrates the dynamics of the price-dividend ratio. Since states are persistent, as illustrated in Figure 1.5, low states continue for a while, then high states continue for a while. Suppose we start with a low state at time τ_0 (Point 0). By the RE, as low states continue to τ_1, agents become more and more risk-averse (δ_t increases). Hence, the price-dividend ratio $\frac{p_t^L(L)}{z_t^L} = \frac{1+\rho\delta_t}{\delta_t}$ decreases between τ_0 and τ_1 (Point 0 to Point 1). However, in HIDA illustrated by the orange lines, since the disappointment parameter is constant, the price-dividend ratio is constant between τ_0 and τ_1. Once the economy starts to recover, the asset price overshoots (Point 1 to Point 2). The overshooting in HDDA is larger than in HIDA because agents with HDDA preferences are more risk-averse at $\tau_1 + 1$ compared to τ_0. At $\tau_1 + 2$, the asset price will be adjusted down (Point 2 to Point 3) because the transition probability $\rho(H|L)$ switches to $\rho(H|H) > \rho(H|L)$ (in Figure 1.4, the first line $\frac{p_t^H(L)}{z_t^H}$ switches to the second line $\frac{p_t^H(H)}{z_t^H}$).

Now high states continue for a while, then by the RE, agents become less risk-averse (δ_t decreases). Hence, the price-dividend ratio $\frac{p_t^H(H)}{z_t^L} = 1 + (1-\rho)\,\delta_t$ decreases between $\tau_1 + 2$ and τ_2 (Point 3 to Point 4). However, in HIDA, the price-dividend ratio is constant between $\tau_1 + 2$ and τ_2. Once the economy starts to decline, the price-dividend ratio drops below the dotted line (Point 4 to Point 5). But the drop is not as large as the overshooting at τ_1 because agents are already less risk-averse after a long period of high states. At $\tau_2 + 2$, the price-dividend ratio will be adjusted up (Point 5 to Point 6) because the transition probability $\rho(H|H)$ switches to $\rho(H|L) < \rho(H|H)$ (in Figure 1.4, the fourth line $\frac{p_t^L(H)}{z_t^L}$ switches to the third line $\frac{p_t^L(L)}{z_t^L}$). So the state is low again.

A sharp prediction that can be spotted easily from Figure 1.5 is that the price-dividend ratio decreases most of the time to correct the overshooting that happens during state changes.

1.4.4 Empirical Regularities: High, Volatile, and Predictable Asset Returns, and Low and Smooth Bond Returns

Finally, I relate the predictions of the model to four stylized facts about the stock market: i) high excess returns (the *equity premium puzzle* of Mehra and Prescott 1985), ii) low bond returns (the *risk-free rate puzzle* of Weil 1989), iii) volatile asset returns and smooth bond returns (the *equity volatility puzzle* of Campbell 2003),

Table 1.1: Simulated Asset and “Shadow" Bond Returns

	$\delta_0=0.15$		$\delta_0=0.2$		Empirical	
	Asset	Bond	Asset	Bond	Asset	Bond
HDDA						
Mean	6.03%	1.53%	7.59%	1.46%	6.98%	0.80%
SD	11.4%	3.2%	15.6%	3.75%	16.54%	5.67%
HIDA						
Mean	2.1%	1.7%	2.23%	1.65%		
SD	2.1%	0.98%	2.6%	1.3%		
DEU						
Mean	1.9%	1.86%				
SD	1.03%	0.82%				

- Empirical values are from Mehra and Prescott (1985) (annualized).
- 1 period = 1 month, β = 0.999 (annual 0.988), ρ = 0.9, and α = 0.97.

and iv) the predictability of asset returns.[19]

One of the most well-known empirical facts about the stock market is that asset returns are high and volatile compared to bonds (Mehra and Prescott 1985, Campbell 2003). For example, Mehra and Prescott (1985) find that the historical average of asset returns in the U.S. is 6.98% with a standard deviation of 16.54%, while the historical average of relatively riskless security returns is 0.80% with the standard deviation of 5.67%. The challenge in obtaining a high excess return (6.18% = 6.98 − 0.80) in standard models is that consumption growth is too smooth, so unreasonably high risk aversion is required (the equity premium puzzle). For example, in the standard discounted expected utility model with CRRA utility, a relative risk aversion of 50 is required to obtain a 6% excess return, while a relative risk aversion of 10 is the maximum level considered plausible by Mehra and Prescott (1985). Moreover, in order to obtain a low bond return with a high degree of risk aversion, the discount factor must be very close to one or even greater than one (the risk-free rate puzzle). A high risk aversion also implies volatile bond returns (the equity volatility puzzle). My model can generate high and volatile asset returns and low and smooth bond returns with a reasonable degree of risk aversion and a reasonable discount factor.

To emphasize that, I choose the dividend process $\{z_t\}$ (equal to $\{c_t\}$ in equilibrium) in a way that the first and second moments of the consumption growth match historical U.S. data. In particular, I choose growth rates $g_H = \frac{z_t^H}{z_{t-1}}$ and $g_L = \frac{z_t^L}{z_{t-1}}$ to

[19] See Campbell (2003) for a survey of important stylized facts about the aggregate stock market.

satisfy $E(g) = 1.018$ and $\sigma(g) = 0.036$ (Mehra and Prescott 1985). Therefore, I have $g_H = 1.054$ and $g_L = 0.982$. Then I run simulations with values $\beta = 0.999$ (annual 0.988), $\rho = 0.9$, $\alpha = 0.97$, and $\delta_0 = 0.2$, where each period is one month. The simulations are illustrated in Figure 1.6. Table 1.1 shows the average asset return and the standard deviation calculated from the simulated data. For example, for HDDA, I find an average return of 7.59% with a standard deviation of 15.6% when the initial disappointment parameter δ_0 is 0.2. However, for HIDA and DEU, the average returns are only 2.23% and 1.9%, respectively (no significant difference between HIDA and DEU). Although there is no one-to-one translation between the disappointment parameter and the degree of relative risk aversion, the initial disappointment parameter $\delta_0 = 0.2$ corresponds to a very low degree of relative risk aversion.[20]

Although bonds are not traded in the Lucas economy, the shadow bond return can be calculated. In the stochastic discount factor framework of asset pricing, any asset return should satisfy the following equation:

$$1 = \mathbb{E}_t[M_{t+1}\, R_{t+1}]$$

where M_{t+1} is the stochastic discount factor and R_{t+1} is the asset return. In my model, the stochastic discount factor is given by

$$M_{t+1} = \beta\, \frac{u'(c_{t+1})/\tilde{\lambda}_{t+1}}{u'(c_t)/\lambda_t} = \beta\, \frac{\lambda_t\, c_t}{\lambda_{t+1}\, c_{t+1}}.$$

Then I can find the shadow bond return by

$$R^f_{t+1} = \frac{1}{\mathbb{E}_t[M_{t+1}]} = \frac{1}{\beta\, \lambda_t} \frac{1}{\pi_{t+1}\, \frac{c_t}{c^H_{t+1}} + (1 - \pi_{t+1})\, \frac{c_t}{c^L_{t+1}}}.$$

From the simulation, I find an average bond return of 1.46% (consistent with the risk-free rate puzzle) with a standard deviation of 3.75% when the initial disappointment parameter δ_0 is 0.2. Consequently, I obtain a high excess return of 6.13% (the equity premium puzzle), and volatile stock returns and smooth bond returns (the equity volatility puzzle) with reasonable parameters.

The main reason HDDA generates high average returns is that agents undervalue high returns because of the probability distortion, so they demand very high returns.

[20]For example, Epstein and Zin (2001) note that, for a range of moderate to large gambles, the disappointment parameter $\delta = 1.63$ corresponds to a relative risk aversion 10 or less. Moreover, Routledge and Zin (2010) introduce a one-parameter generalization of Gul's disappointment aversion theory and applied it to the Epstein-Zin framework of asset pricing. They use the disappointment parameter $\delta = 9$ in order to generate stylized facts I discussed above.

Moreover, when returns are very high, agents happen to be very risk-averse (Point 1 to Point 2 in Figure 1.5), so they underappreciate high returns.

Another empirical fact about the stock market is the predictability of asset returns, which refers to predictable dynamics of asset returns in the following sense: when asset prices are high, subsequent long-horizon asset returns tend to be low. To check the predictability, the following simple regression is usually used:

$$\log(R_{t+1}) = a + b \log\left(\frac{p_t}{z_t}\right) + \epsilon_{t+1},$$

where $R_{t+1} = \frac{p_{t+1}+z_{t+1}}{p_t}$ is the asset return. The predictability of asset returns means that empirically b is negative, since a high $\frac{p_t}{z_t}$ must be followed by a low R_{t+1}. From the simulation, I find that $\hat{a} = 4.9$ and $\hat{b} = -0.7$ (with standard error 0.06). This regression suggests that a 10% increase in the price-dividend ratio implies a 7% decrease in the next period return.

1.5 Behavioral Foundations of the History-Dependent Model

In this section, I provide behavioral foundations for the history-dependent model (1.3) with axioms on the primitive: a preference relation $\succeq$ on the set of intertemporal consumption lotteries $\mathscr{L} = \Delta(\mathbb{R}_+ \times \Delta(\mathbb{R}_+))$. There are two different approaches to characterizing (1.3). First, since the history-dependent model (1.3) is a generalization of DEU and a characterization of DEU is standard, I can obtain (1.3) from DEU by dropping some restrictions of DEU. Second, I can directly characterize (1.3) by imposing axioms on $\succeq$.

I start with the first approach. For notational simplicity, I write $(p_i, (x_i; Z_i); (p_k, (x_k; 0))_{k\neq i})$ as $(p_i, (x_i; Z_i); (X_{-i}; 0))$ when $X = (p_k, x_k)_{i=1}^n$. First, recall the four properties of DEU discussed in Section 1.3.1: i) simple expected utility, ii) history independence, iii) discounted utility, and iv) expected utility aggregator. Then recall that the history-dependent model (1.3) weakens the first two properties, simple expected utility and history independence. Therefore, I introduce two axioms that capture the above two properties.

First, I state an axiom for history independence, which states that if μ is the certainty equivalent of Z today, then μ is the certainty equivalent of Z after any history (x_i, X). Formally,

Axiom 1 (Axiom for History Independence) *For any $X, Z \in \Delta(\mathbb{R}_+)$ and $\mu \in \mathbb{R}_+$,*

$$(Z; 0) \succeq (\mu; 0) \text{ iff } (p_i, (x_i; Z), (X_{-i}; 0)) \succeq (p_i, (x_i; \mu), (X_{-i}; 0)) \text{ for any } i.$$

Second, I state an axiom for simple expected utility. Simple expected utility is captured by an axiom on $\succeq$ called *strong independence*, which states that if a lottery X today is equivalent to an outcome z tomorrow and a lottery Y today is equivalent to an outcome t tomorrow; then a mixture $\alpha X + (1-\alpha)Y$ is equivalent to a compound lottery $(\alpha, (0; z), 1-\alpha, (0; t))$. Formally,

Axiom 2 (Strong Independence) *For any $X, Y \in \Delta(\mathbb{R}_+)$, $z, t \in \mathbb{R}_+$, and $\alpha \in (0, 1)$,*

if $(X; 0) \succeq (0; z)$ and $(Y; 0) \succeq (0; t)$, then $(\alpha X + (1-\alpha)Y; 0) \succeq (\alpha, (0; z), 1-\alpha, (0; t))$.

Strong independence is slightly stronger than the independence axiom of expected utility theory. I now can state the first characterization result.

Theorem 2 (Discounted Expected Utility) *A continuous preference relation $\succeq$ on $\mathscr{L}$ is represented by a history-dependent model $\{V_0, \beta, V_{(x,X)}\}$ and satisfies time consistency; that is, $V_0(z) = V_{(x,X)}(z)$ for any $z \in \mathbb{R}_+$, the axiom for history independence, and strong independence if and only if there exists a continuous function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that for any $L = (p_i, (x_i; Z_i))_{i=1}^n, L' = (p'_k, (x'_k; Z'_k))_{k=1}^m \in \mathscr{L}$,*

$$L \succeq L' \text{ iff } \sum_{i=1}^n p_i \left(u(x_i) + \beta\, \mathbb{E}u(Z_i)\right) \geq \sum_{k=1}^m p'_k \left(u(x'_k) + \beta\, \mathbb{E}u(Z'_k)\right).$$

Theorem 2 formally shows that the history-dependent model (1.3) is a result of dropping two properties of DEU, history-independence and simple expected utility.

Now I turn to the second approach: imposing three axioms on $\succeq$. The first axiom is called *regularity*, a collection of four standard postulates.

Axiom 3 (Regularity) *A preference relation $\succeq$ on $\mathscr{L}$ satisfies the following four conditions.*

1. *The preference relation $\succeq$ is complete, transitive, and continuous.*
2. *(Deterministic Monotonicity) For any $z, z' \in \mathbb{R}_+$ with $z > z'$ and $X \in \Delta(\mathbb{R}_+)$,*

$$(z; 0) \succ (z'; 0) \text{ and } (p_i, (x_i; z), (X_{-i}; 0)) \succ (p_i, (x_i; z'), (X_{-i}; 0)).$$

3. *(Discounted Utility) There exist a utility function* $u : \mathbb{R}_+ \to \mathbb{R}_+$ *and a discount factor* $\beta \in (0, 1)$ *such that for any* $(x; z), (x'; z') \in \mathbb{R}^2_+$.

$$(x; z) \geq (x'; z') \text{ iff } u(x) + \beta\, u(z) \geq u(x') + \beta\, u(z'), \tag{1.10}$$

4. *(Expected Utility Aggregator) There exists* $U_2 : \mathbb{R}_+ \to \mathbb{R}$ *such that for any* $Z = (r_k, z_k)_{k=1}^m, Z' = (r'_k, z'_k)_{k'=1}^{m'} \in \Delta(\mathbb{R}_+)$,

$$(r_k, (0; z_k))_{k=1}^m \geq (r'_k, (0; z'_k))_{k=1}^{m'} \text{ iff } \mathbb{E}U_2(Z) \geq \mathbb{E}U_2(Z'). \tag{1.11}$$

The first part of regularity collects completeness, transitivity, and continuity. I also assume a very weak form of monotonicity called *deterministic monotonicity*. The second half of regularity imposes the other two properties of DEU, discounted utility and expected utility aggregator. Specifically, the third part states that the agent uses discounted utility theory when she aggregates utilities of today and tomorrow; i.e., the utility of $(x; z)$ is $u(x) + \beta\, u(z)$ (discounted utility). The fourth part states that compound lotteries are evaluated by expected utility theory; i.e, the utility of a compound lottery $(r_k, (0; z_k))_{k=1}^m$ is $\sum_{k=1}^m r_k U_2(z_k)$ (expected utility aggregator).

The next two axioms are novel axioms. The second axiom (Axiom 4) is called *separability*, which consists of two properties of separability. Separability is essential for studying history-dependent risk aversion because it allows me to define a risk preferences for each history independently of other histories. Separability also implies time consistency.

Axiom 4 (Separability) *A preference relation* $\geq$ *on* $\mathscr{L}$ *satisfies the following two properties.*

1. *(Separability between Parallel Histories) For any* $(p_i, (x_i; Z_i))_{i=1}^n \in \mathscr{L}$ *and* $Y, Y' \in \Delta(\mathbb{R}_+)$,

$$\big(p_i, (x_i; Y), (X_{-i}; 0)\big) \geq \big(p_i, (x_i; Y'), (X_{-i}; 0)\big)$$

$$\text{iff } \big(p_i, (x_i; Y), (p_k, (x_k; Z_k))_{k \neq i}\big) \geq \big(p_i, (x_i; Y'), (p_k, (x_k; Z_k))_{k \neq i}\big).$$

2. *(Weak Separability between Today and Tomorrow) For any* $(p_i, (x_i; z_i))_{i=1}^n \in \Delta(\mathbb{R}_+ \times \mathbb{R}_+)$ *and* $\mu, z \in \mathbb{R}_+$,

   i) $(X; 0) \geq (\mu; 0)$ *if and only if* $(X; z) \geq (\mu; z)$ *and* *ii)* $(p_i, (0; z_i))_{i=1}^n \succeq (0; \mu)$ *if and only if* $(p_i, (x_i; z_i))_{i=1}^n \succeq (X; \mu)$.

Suppose the agent receives either of two simple lotteries Y and Y' after a history (x_i, X). Separability between parallel histories requires that a comparison between the two simple lotteries Y and Y' cannot be affected by what she would receive in histories other than (x_i, X). This axiom is essentially a dynamic version of an axiom called *replacement separability*, introduced in Machina (1989). Weak separability between today and tomorrow essentially requires that the utility of a deterministic outcome z is history-independent (recall time consistency); i.e., the utility of z is not affected by a lottery X and a deterministic outcome μ. Specifically, i) states that if μ is the certainty equivalent of a simple lottery X, then μ is still the certainty equivalent of X even if the agent will receive a deterministic outcome z in the second period. Moreover, ii) states that if μ is the certainty equivalent of a compound lottery $(p_i, (0; z_i))_{i=1}^n$, then $(0; \mu)$ is still the certainty equivalent of $(p_i, (0; z_i))_{i=1}^n$ even if the agent receives a lottery X in the first period.

The third axiom (Axiom 5) is called *additivity*.[21] Suppose a risky option that gives z with probability p tomorrow is equally preferred to a safe option μ today. Similarly, suppose a risky option that gives z' with probability $1 - p$ tomorrow is equally preferred to a safe option λ tomorrow. Additivity requires that a combination of the two risky options is equally preferred to a combination of the two safe options; that is, receiving z with probability p and z' with probability $1-p$ tomorrow is equally preferred to receiving μ today and λ tomorrow.

Axiom 5 (Additivity) *For any* $(p, (0; z), 1 - p, (0; z')) \in \Delta(\mathbb{R}_+ \times \mathbb{R}_+)$ *and* $\lambda, \mu \in \mathbb{R}_+$*, and for any* i, j*,*

$$\text{if } (p, (0; z), 1 - p, (0; 0)) \sim (\mu; 0) \text{ and } (p, (0; 0), 1 - p, (0; z')) \sim (0; \lambda),$$

$$\text{then } (p, (0; z), 1 - p, (0; z')) \sim (\mu; \lambda).$$

Finally, I can state the second characterization theorem.

Theorem 3 (History-Dependent Model) *A preference relation* $\succeq$ *on* $\mathscr{L}$ *satisfies regularity, separability, and additivity if and only if there are strictly increasing*

[21] Additivity is not essential to Theorem 1. I can relax it and Theorem 1 can be modified accordingly.

continuous functions $V_0 : \Delta(\mathbb{R}_+) \to \mathbb{R}_+$ *and* $V_{(x,X)} : \Delta(\mathbb{R}_+) \to \mathbb{R}_+$ *such that for any* $L = (p_i, (x_i; Z_i))_{i=1}^n, L' = (p'_k, (x'_k; Z'_k))_{k=1}^m \in \mathscr{L}$,

$$L \succeq L' \text{ iff } V_0(X) + \beta \sum_{i=1}^n p_i\, V_{(x_i, X)}(Z_i) \geq V_0(X') + \beta \sum_{k=1}^m p'_k\, V_{(x'_k, X')}(Z'_k), \quad (1.12)$$

and $V_0(z) = V_{(x,X)}(z)$ *for each* $z \in \mathbb{R}_+$.

I also have a strong uniqueness result.

Proposition 1 (Uniqueness) *Take any preference relation* $\succeq$ *on* $\mathscr{L}$ *that satisfies regularity 1-2. If it is represented by triplets* $(V_0, \beta, \{V_{(x,X)}\})$ *and* $(V'_0, \beta', \{V'_{(x,X)}\})$ *that satisfy time consistency,* $V_0(0) = V'_0(0)$, *and* $V_0(1) = V'_0(1)$, *then the two triplets are identical.*

I conclude this section by illustrating that Theorem 1 can be stated in terms of axioms on $\succeq$ instead of using the history-dependent model $\{V_{(x,X)}\}$. Since dynamic monotonicity is already defined on the primitive $\succeq$, and the history-dependent model (1.3) is characterized by Theorem 3, it is sufficient to define the RE in terms of conditions on $\succeq$.[22] It turns out, the RE is equivalent to the following condition.

Definition 5 (Reinforcement Effect) *For any simple lottery* $X \in \Delta(\mathbb{R}_+)$ *and* $x_i, x_j \in supp(X)$ *with* $x_i > x_j$, *for any* $Z \in \Delta(\mathbb{R}_+)$ *and* $\mu \in \mathbb{R}_+$,

$$\text{if } \big(p_j, (x_j; Z), (X_{-j}; 0)\big) \succeq \big(p_j, (x_j; \mu), (X_{-j}; 0)\big), \text{ then}$$

$$\big(p_i, (x_i; Z), (X_{-i}; 0)\big) \succeq \big(p_i, (x_i; \mu), (X_{-i}; 0)\big).$$

1.6 Specifying Risk Preferences: Expected Utility, Disappointment Aversion, and Rank-Dependent Utility

So far, I have discussed a relation between risk preferences for different histories, but I have not focused on specific functional forms for V_0 and $V_{(x,X)}$. In this section, I obtain three special cases of the history-dependent model (1.3) by specifying V_0 and $V_{(x,X)}$. Specifically, I apply the history-dependent model (1.3) to three well-known models of choice under risk: expected utility, the disappointment aversion theory of Gul (1991), and the rank-dependent utility of Quiggin (1982) (which

[22] I can also state right-continuity and nontriviality in terms of axioms on the primite $\succeq$.

subsumes the cumulative prospect theory of Tversky and Kahneman 1992). By doing so, I demonstrate the generality of Theorem 1 in the following two ways. First, three special cases demonstrate the generality of the history-dependent model (1.3). Second, for these three special cases, I show that nontriviality, the main assumption of Theorem 1, is equivalent to the condition $u(+\infty) = +\infty$ where u is a Bernoulli utility function. This illustrates that nontriviality is equivalent to a condition that is unrelated to history dependence; i.e., dynamic monotonicity is almost fully responsible for the RE.

Specifying V_0 and $V_{(x,X)}$ is useful for the following three reasons. First, I can derive a simple condition that is equivalent to the RE in each special case, which allows me to provide further interpretations for the RE. Second, although I consider objective lotteries, I can demonstrate that the way my model explains the RE is similar to a belief-based explanation (Section 1.6.4). Lastly, these specifications make the history-dependent model (1.3) more tractable and applicable (demonstrated in Section 1.4).

I call the three special cases the *history-dependent expected utility* (HDEU), *history-dependent disappointment aversion* (HDDA), and *history-dependent rank-dependent utility* (HDRDU), respectively. Behavioral foundations for HDEU and HDDA are provided in Appendix A.2. Let me start with the definition of HDEU.

1.6.1 History-Dependent Expected Utility (HDEU)

Here I introduce a dynamic version of expected utility theory in which the concavity of an agent's Bernoulli utility function changes with her experiences. More specifically, today, the agent uses expected utility theory with some Bernoulli utility function u; i.e., the utility of a simple lottery X is

$$V_0(X) = \mathbb{E}[u(X)] = \sum_{i=1}^{n} p_i\, u(x_i).$$

Tomorrow, u becomes $u^{\mu(x,X)}$ after a history (x, X) for some positive real number $\mu(x, X)$. In other words, the utility of Z after the history (x, X) is

$$V_{(x,X)}(Z) = \Big(\mathbb{E}[(u(Z))^{\mu(x,X)}]\Big)^{\frac{1}{\mu(x,X)}} = \Big(\sum_{k=1}^{m} r_k\, (u(z_k))^{\mu(x,X)}\Big)^{\frac{1}{\mu(x,X)}}.$$

Therefore, in the *history-dependent expected utility* (HDEU), the utility of an intertemporal consumption lottery $(p_i, (x_i; Z_i))_{i=1}^n \in \mathscr{L}$ is

$$W((p_i, (x_i; Z_i))_{i=1}^n) = \mathbb{E}[u(X)] + \beta \sum_{i=1}^{n} p_i \Big(\mathbb{E}[(u(Z_i))^{\mu(x_i,X)}]\Big)^{\frac{1}{\mu(x_i,X)}}. \qquad (1.13)$$

Since $\mu(x, X)$ dictates the concavity of the utility function $u^{\mu(x,X)}$, the RE is equivalent to the condition:

$$\mu(x, X) \geq \mu(x', X) \text{ when } x > x'.$$

1.6.2 History-Dependent Disappointment Aversion (HDDA)

I now introduce a dynamic version of the disappointment aversion theory of Gul (1991). Disappointment aversion theory is a one-parameter generalization of expected utility theory. In disappointment aversion theory, an agent overweights the probabilities of small outcomes and underweights the probabilities of large outcomes. The degree of such probability distortion is summarized by a single parameter δ_0, a *disappointment parameter*. In particular, for a given lottery X, the agent overweights the probabilities of outcomes that are not larger than the certainty equivalent of X by $1 + \delta_0$. Then the certainty equivalent of the lottery X, denoted by $\mu(X|\delta_0)$, is a unique solution of the following implicit formula where μ is variable:

$$u(\mu) = \frac{\sum_{i=1}^{n} p_i \, (1 + \delta_0 \, \mathbb{1}\{x_i \leq \mu\}) \, u(x_i)}{\sum_{i=1}^{n} p_i \, (1 + \delta_0 \, \mathbb{1}\{x_i \leq \mu\})}.^{23} \tag{1.14}$$

Note that the disappointment parameter dictates the degree of risk aversion (for a fixed u); that is, a higher δ_0 implies a higher degree of risk aversion.

In *history-dependent disappointment aversion* (HDDA), the agent's disappointment parameters change with her experiences. In particular, tomorrow, the agent's initial disappointment parameter δ_0 changes to $\delta(x, \ X)$ after a history $(x, \ X)$. Then the certainty equivalent of a lottery Z after (x, X), denoted by $\mu(Z|\delta(x, X))$, is a unique solution to

$$u(\mu) = \frac{\sum_{k=1}^{m} r_k \, (1 + \delta(x, \ X) \, \mathbb{1}\{z_k \leq \mu\}) \, u(z_k)}{\sum_{k=1}^{m} r_k \, (1 + \delta(x, \ X) \, \mathbb{1}\{z_k \leq \mu\})}.$$

Therefore, in HDDA, the utility of $(p_i, (x_i; Z_i))_{i=1}^{n} \in \mathscr{L}$ is

$$u(\mu(X|\delta_0)) + \beta \sum_{i=1}^{n} p_i \, u\big(\mu(Z_i|\delta(x_i, X))\big). \tag{1.15}$$

Since the disappointment parameter dictates the degree of risk aversion, the RE is equivalent to the following simple condition:

$$\delta(x, X) \leq \delta(x', X) \text{ when } x > x'.$$

[23] Note that (1.8) is a special case of this formula.

1.6.3 History-Dependent Rank-Dependent Utility (HDRDU)

I then introduce a dynamic version of rank-dependent utility in which distorted probabilities change with histories. The rank dependent utility of Quiggin (1982) is a modification of the prospect theory of Kahneman and Tversky (1979) that satisfies monotonicity in the static environment.

In rank-dependent utility, the agent distorts probabilities by some function π and the utility of X is

$$V_0(X) = \sum_{i=1}^{n} \left(\pi(P_i) - \pi(P_{i+1})\right) u(x_i),$$

where $x_1 > \ldots > x_n$ and $P_i = \sum_{k=i}^{n} p_k$. In the *history-dependent rank-dependent utility*, π becomes $\pi^{\mu(x,\, X)}$ after a history (x, X). In other words, the utility of Z after (x, X) is

$$V_{(x,\, X)}(Z) = \sum_{k=1}^{m} \left[(\pi(R_k))^{\mu(x,\, X)} - (\pi(R_{k+1}))^{\mu(x,\, X)}\right] u(z_k),$$

where $z_1 > \ldots > z_m$ and $R_k = \sum_{s=k}^{m} r_s$. Therefore, in HDRDU, the utility of an intertemporal consumption lottery $(p_i, (x_i; Z_i))_{i=1}^{n} \in \mathscr{L}$ is

$$\sum_{i=1}^{n}(\pi(P_i)-\pi(P_{i+1}))u(x_i)+\beta\sum_{i=1}^{n}p_i\Big(\sum_{k=1}^{m_i}\Big[(\pi(R_{i,k}))^{\mu(x_i,\, X)}-(\pi(R_{i,k+1}))^{\mu(x_i,\, X)}\Big]u(z_{i,k})\Big), \tag{1.16}$$

where $z_{i,1} > \ldots > z_{i,m_i}$ and $R_{i,k} = \sum_{s=k}^{m_i} r_{i,s}$ for each i.

It turns out that, the concavity of the distortion function dictates the degree of risk-aversion. Therefore, in HDRDU, the RE is equivalent to the following simple condition:

$$\mu(x,\, X) \geq \mu(x',\, X) \text{ when } x > x'.$$

1.6.4 Probability Distortion, Belief Change, and Nontriviality

My model provides the preference-based explanation of the RE; that is, an agent becomes less risk-averse after a good history than after a bad history. However, the belief-based explanation is another obvious possibility: the agent becomes less pessimistic after a good history and acts as if she is less risk-averse because she thinks probabilities of high outcomes are higher than she previously thought. I have two remarks on the belief-based explanation. First, the belief-based explanation is not consistent with the experiment of Thaler and Johnson (1990) because in their experiment, the subjects are asked to compare objective lotteries. Second, the way the above three special cases of the history-dependent model explain the RE is very similar to the belief-based explanation. To illustrate, consider HDDA, in which the utility of $Z = (r, z, 1 - r, 0)$ after (x, X) is

$$V_{(x,X)}(Z) = \frac{r}{1 + (1 - r)\,\delta(x, X)}\, u(z).$$

Note that the distorted probability $\frac{r}{1+(1-r)\,\delta(x,X)}$ can be interpreted as the agent's subjective belief. Therefore, the agent exhibits the RE because she acts as if her belief changes with her experiences. In particular, when $\delta(x, X) \leq \delta(x', X)$, I have

$$\frac{r}{1 + (1 - r)\,\delta(x, X)} \geq \frac{r}{1 + (1 - r)\,\delta(x', X)}.$$

Therefore, the agent acts as if she becomes less pessimistic after a good history (x, X) than a bad history (x', X). Similar to HDDA, in HDRDU, the utility of $Z = (r, z, 1 - r, 0)$ after (x, X) is

$$V_{(x,X)}(Z) = \left[1 - (\pi(1 - r))^{\mu(x,X)}\right] u(z)$$

and the distorted probability $1 - (\pi(1 - r))^{\mu(x,X)}$ can be interpreted as the agent's subjective belief. In HDEU, the utility of $Z = (r, z, 1 - r, 0)$ after (x, X) is

$$V_{(x,X)}(Z) = r^{\frac{1}{\mu(x,X)}}\, u(z),$$

and $r^{\frac{1}{\mu(x,X)}}$ can be interpreted as the agent's subjective belief.[24]

Now it is easy see that why nontriviality is equivalent to the condition $u(+\infty) = +\infty$. Let $\pi_{(x,X)}(r)$ be a distorted version of probability r after a history (x, X), which can be obtained in all three cases. Note that if there exists Z such that $V_{(x,X)}(Z) > V_{(x,X)}(Z)$, then there is some $r \in (0, 1)$ such that $\pi_{(x,X)}(r) > \pi_{(x',X)}(r)$. Therefore, nontriviality is satisfied when $u(+\infty) = +\infty$ since

$$V_{(x,X)}(r, z, 1-r, 0) - V_{(x',X)}(r, z, 1-r, 0) = \left(\pi_{(x,X)}(r) - \pi_{(x',X)}(r)\right) u(z) \to +\infty \text{ as } z \to +\infty.$$

[24] In HDEU, the belief-based interpretation is limited to binary lotteries while in HDDA and HDRDU, the belief-based interpretation works for any lottery Z.

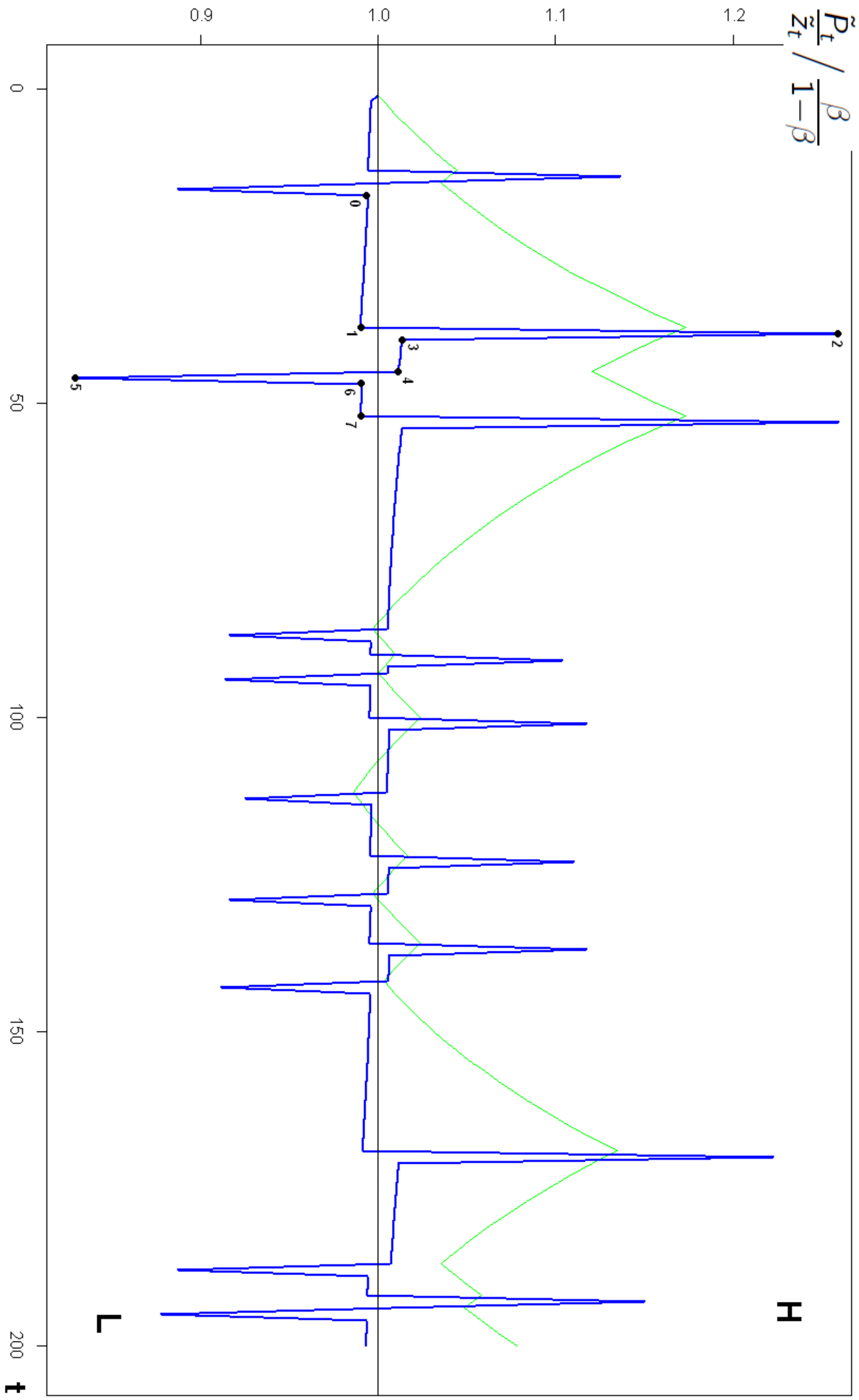

Figure 1.6: The Price-Dividend Ratio (blue) and the Disappointment Parameter (green)

Chapter 2

CHOOSING WITH THE WORST IN MIND: A REFERENCE-DEPENDENT MODEL

2.1 Introduction

Rational choice theory does not allow for *preference reversals*, which occur when an agent chooses alternative **x** over alternative **y** in some cases but she chooses **y** over **x** in other cases. However, many experimental, marketing, and field studies suggest that the presence of an irrelevant third alternative could cause a preference reversal. Preference reversals are typically explained by models of reference-dependent behavior in which the agent's basis for decision making, her *reference point*, changes depending on the situations that she faces. In this paper, we axiomatically develop a reference-dependent model with endogenous reference points that is consistent with two well-known preference reversals, the *compromise effect* and the *attraction effect* (to be defined below).

The notion of reference dependence is first introduced to economics by Markowitz (1952) and is formalized by Kahneman and Tversky (1979) in the context of risky choice. Later, Tversky and Kahneman (1991) provide the first explicit model of reference-dependent preferences on riskless choice by extending Kahneman and Tversky (1979). In their model, they treat reference points as exogenous and explain many anomalies of the standard model using the following two properties: i) people are more sensitive to losses than gains (*loss aversion*) and ii) the marginal values of both gains and losses decrease with their distance from the reference point (*diminishing sensitivity*). Despite their explanatory power, reference-dependent models with exogenous reference points can be consistent with essentially any choice behavior by freely choosing reference points. In order to have a model that makes testable predictions for observed behavior, we need to explicitly model reference points in a way that they are determined by observable factors. Therefore, in this paper we propose a model in which reference points endogenously arise from the menu that the agent faces. Our model will be flexible enough to allow for the compromise and attraction effects, but restrictive enough to give bounds on preference reversals. Moreover, the explicit modeling of reference points is useful when we apply the model to different contexts such as intertemporal choice and risky choice.

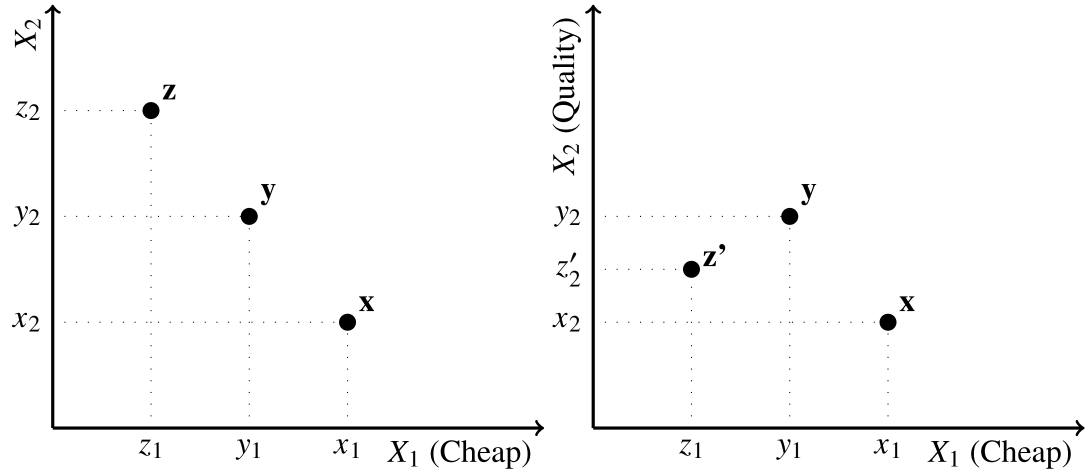

Figure 2.1: Compromise and Attraction Effects

Before introducing our modeling approach, let us discuss the compromise and attraction effects, two well-documented deviations from rational choice theory.[1] Throughout the paper, we focus on alternatives with two attributes and the agent's preference is increasing in both attributes. For example, the first dimension could be the inverse of the price (cheapness) of the good and the second dimension could be the quality of the good. In both the compromise and attraction effects, the agent first compares two alternatives $\mathbf{x} = (x_1, x_2)$, a cheap and low quality good, and $\mathbf{y} = (y_1, y_2)$, a medium price and medium quality good. Since $x_1 > y_1$ and $x_2 < y_2$, the agent faces a tradeoff between price and quality. Suppose **x** is chosen over **y**. Then the two effects relate to the consequences of adding a very expensive third alternative.[2]

In the compromise effect, the introduction of a very expensive high quality good $\mathbf{z} = (z_1, z_2)$ causes a preference reversal; that is, **y** is preferred over **x** in the presence of **z** (the left hand side of Figure 2.1). The common explanation is that the very expensive high quality good **z** makes the cheap low quality good **x** seems like an extreme alternative and people compromise by choosing **y** since they have a

[1] The compromise effect and the attraction effect are first documented in the experimental studies of Simonson (1989) and Huber et al. (1982), respectively, and confirmed by many studies in consumer choice (e.g., Simonson and Tversky 1992, Tversky and Simonson 1993, Ariely and Wallsten 1995, Herne 1998, Doyle et al. 1999, Chernev 2004, and Sharpe et al. 2008). These effects are also demonstrated in the contexts of choice over risky alternatives (Herne 1999), choice over policy issues (Herne 1997), and choice over political candidates (Sue O'Curry and Pitts 1995), among others.

[2] Here x_1 is the lowest price and z_1 is the highest price while x_2 is the lowest quality and z_2 is the highest quality.

tendency to avoid extremes.

In the attraction effect (sometimes called the asymmetric dominance effect or the decoy effect), the introduction of a very expensive, but medium quality good **z'** $= (z_1, z_2')$ causes a preference reversal; that is, **y** is preferred over **x** in the presence of **z'** (the right hand side of Figure 2.1). The common explanation is that since the medium quality but expensive good **z'** is dominated by the medium quality, medium price good **y**, but not by **x**, the third alternative makes **y** seem more attractive than **x**. In other words, **z'** works as a decoy for **y**.

The common understanding among psychologists is that the attraction and compromise effects are not separate phenomena, but rather two manifestations of the same behavior. Indeed, in the seminal paper by Simonson (1989), the same group of subjects exhibited the two effects in roughly the same magnitude. The aim of this paper is to develop a model that can provide a unified explanation for the two effects using the aforementioned diminishing sensitivity property of Tversky and Kahneman (1991). To illustrate the idea, recall that in both effects the introduction of a very expensive good causes a preference reversal. Then note that the two effects can be explained by the following rationale: *adding* z_1 *hurts* x_1 *over* y_1. In our model, this rationale will be an implication of diminishing sensitivity. More specifically, we develop a model in which the reference point for the given menu is a vector that consists of the minimums (e.g., the worst price and the worst quality) of each dimension of the menu. Therefore, since adding z_1 decreases the minimum of the first dimension of the menu $\{\mathbf{x}, \mathbf{y}\}$, the marginal value of x_1 over y_1 decreases by diminishing sensitivity. Intuitively, adding a very expensive good makes a moderately expensive good seems like relatively cheaper, consequently the advantage of **x** in price is less important. In other words, z_1 hurts x_1 over y_1. We call the above rationale *weak diminishing sensitivity*, and it not only rationalizes the two effects, but also has three additional implications that provide bounds on preference reversals.

Our explanation of the compromise and attraction effects is consistent with the common explanations of the two effects since adding a very expensive alternative makes the advantage of **y** in quality more important (attraction effect) and makes people to avoid low quality alternatives (compromise effect).

We behaviorally characterize both diminishing sensitivity and weak diminishing sensitivity. In particular, we show that weak diminishing sensitivity is equivalent to observing the attraction effect, while diminishing sensitivity is equivalent to observing the attraction effect in a more limited way (Section 2.3.2). In other words, we characterize diminishing sensitivity by bounds on the attraction effect. This result illustrates that diminishing sensitivity is not only sufficient for the compromise and attraction effects, but also necessary.

We focus on a choice theoretic environment in which an agent makes choices from menus of two-attribute alternatives. To introduce our model, we use the following formulation of Tversky and Kahneman (1991): the total utility of a two-attribute alternative $\mathbf{x}$ for a given reference point $\mathbf{r} = (r_1, r_2)$ is

$$U(\mathbf{x}|\mathbf{r}) = f_1(u_1(x_1) - u_1(r_1)) + f_2(u_2(x_2) - u_2(r_2)). \tag{2.1}$$

where u_1 and u_2 are strictly increasing utility functions and f and g are strictly increasing distortion functions. In other words, the utility of the i-th dimension x_i is evaluated with respect to the reference for the i-th dimension r_i and distorted by the distortion function f_i. Note that when f_1 and f_2 are linear, (2.1) reduces to the standard additive utility model.

Now we turn to our model in which we specify (2.1) in the following two ways. First, we explicitly model reference points in order to obtain predictive power, while Tversky and Kahneman treated $\mathbf{r}$ as exogenous. In particular, the agent uses a vector that consists of the minimum of each dimension of the menu as a reference point; that is, $\mathbf{m}^A$ is the reference point of a menu A where m_i^A is the minimum of the i-th dimension of A. Since utility functions are strictly increasing, m_i^A is the worst attribute of i-th dimension of A. Second, we also require $f_1 = f_2$ in order to rule out violations of transitivity (see Remark 1 in Section 2.2). Then in our model, the total utility of an alternative $\mathbf{x}$ of a given menu A is

$$U_A(\mathbf{x}) = f(u_1(x_1) - u_1(m_1^A)) + f(u_2(x_2) - u_2(m_2^A)).$$

We take this particular approach using the minimums as references for several reasons. First, as we discussed earlier, in order to be consistent with the two effects, the reference point should decrease when (z_1, z_2) or (z_1, z_2') (i.e., a very expensive good) is added. Second, if we consider more general models in which reference points are determined by several factors in the menu, then the predictive power of the model will be much weaker. For example, if we allow any kind of menu dependence for reference points, then the model will not have any testable restriction on observed behavior (Section 2.5). So we need to restrict our attention to a particular kind of menu dependence such as dimension-by-dimension minimums.[3] Lastly, our model is one of the most restrictive specifications of (2.1) because it has only two additional "parameters" f and $\mathbf{m}^A$ compared to the standard additive utility model. Without one of the above two "parameters", the model reduces to the standard model. Therefore, our model is restrictive enough to make strong predictions (Sections 2.2.3, 2.3.2, and 2.4) but flexible enough to capture well-known deviations from the standard choice theory (Section 2.2.2).

One of the main contributions of the paper is that we axiomatically characterize the model by two novel axioms called *independence from non-extreme alternatives* (INEA) and *reference translation invariance* (RTI) in addition to standard axioms (Section 2.3). Both INEA and RTI are weakenings of the weak axiom of revealed preference (WARP). The essence of the model will be captured by INEA, and INEA can be tested quite easily in experiments.

We apply the model to two different contexts in Section 2.4: intertemporal choice and risky choice. In intertemporal choice, the main implication of the model is that borrowing constraints produce a psychological pressure to move away from the constraints even if they are not binding. To illustrate the intuition, consider a model with two periods, today and tomorrow, and suppose there is a non-binding borrowing constraint today. In the standard model, non-binding borrowing constrain would not affect the optimal consumption levels. However, in our model, the constraint decreases the maximum level for today's consumption, and it increases the minimum level for tomorrow's consumption. By diminishing sensitivity, tomorrow's consumption becomes more valuable as the reference (i.e., the minimum level) for tomorrow's consumption increases. Therefore, the consumer increases tomorrow's consumption and decreases today's consumption.

It is common to explain deviations from the standard model of consumer choice such as *excess sensitivity of consumption to income* and a *hump-shaped consumption profile* with liquidity or borrowing constraints (see Attanasio 1999 and Attanasio and Weber 2010). But empirically it cannot be determined whether constraints

[3]It is very natural to think that people care about how attractive a given attribute is compared to the worst and best attributes. In this paper, we focus on the worst (min) attribute instead of the the best (max) attribute for two reasons. First, a model in which references are determined only by maximums cannot be consistent with the attraction effect. Second, when an alternative that dominates all alternatives in a menu is added to the menu, comparisons between alternatives in the menu cannot be observed. In Appendix B.3, we will consider a model in which reference points depend on both the maximums and the minimums. The main implications stay the same, but behavioral predictions are weaker.

were binding at the time of the decision. Our model gives a justification for those explanations with constraints since liquidity constraints can have an effect on choices even if they are not binding.

In risky choice, our model allows contradicting risk behaviors: an agent makes a risky choice in case 1 and makes a safe choice in case 2 and the estimated degrees of risk aversion (in a sense of EUT) from two cases contradict each other. In both applications, the driving force of the results is diminishing sensitivity.

We discuss existing models of reference-dependent preferences in the related literature. To the best of our knowledge, Ok et al. (2014) is the only paper that axiomatically develops a model with purely endogenous reference points. Ok et al. (2014) also weaken WARP and can accommodate the attraction effect. The main difference is that we explicitly model reference points, while they identify reference points from observed choices in the spirit of classical revealed preference analysis. The explicit modeling gives us more flexibility to apply our model to many different contexts and to use diminishing sensitivity to provide bounds on preference reversals. Moreover, Ok et al. (2014) cannot capture the compromise effect.[4]

A challenge to many existing reference-dependent models that can rationalize the compromise and attraction effects is that they cannot predict that a preference reversal is unlikely when the third alternative is obviously dominated by both **x** and **y** (Masatlioglu and Uler 2013). However, our model can predict that a preference reversal is unlikely when the third alternative is dominated by both **x** and **y**. Moreover, our model predicts the following implication of the attraction effect: when there is the fourth alternative which is a decoy for **x**, then **x** should chosen over **y**. Intuitively, the decoy effect of **z'** for **y** as in the attraction effect should be canceled out by the decoy effect of the fourth alternative for **x**. See Section 2.2.3 for more detailed discussions of the above predictions and additional predictions.

The paper is organized as follows. In Section 2.2.2, we formally define the model and diminishing sensitivity. We demonstrate that our model is not only consistent with the compromise and attraction effects, but it also connects the two effects through weak diminishing sensitivity. Section 2.3 discusses a behavioral foundation of the model. Section 2.4 discusses applications to intertemporal and risky choice. Section 2.5 shows that models with general menu-dependent reference

[4] In their model, an alternative that causes a preference reversal (e.g., $\mathbf{z}$ of $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ in the compromise effect) should be unambiguously worse than the chosen alternative ($\mathbf{y}$ of $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$). See Proposition 2 of Ok et al. (2014).

points have no testable implications on observed behavior. Section 2.6 discusses the related literature. The proofs and three extensions of the model are collected in Appendix B.

2.2 Model

Let $X = X_1 \times X_2 \equiv \mathbb{R}^2_+$ be the set of all alternatives and $\mathscr{A} \subset 2^X \setminus \{\emptyset\}$ be a collection of finite subsets (menus) of X. An alternative with two attributes can have many different interpretations: i) a single consumer good with two attributes (price and quality), ii) a consumption bundle of two goods, iii) an allocation to two agents, and iv) a state-contingent prospect.[5]

Let $\mathbf{m}^A = (m^A_1, m^A_2) \equiv (\min_{\mathbf{x}\in A} x_1, \min_{\mathbf{x}\in A} x_2)$ for each $A \in \mathscr{A}$ (i.e., the meet of A). We denote generic members of X by $\mathbf{x}, \mathbf{y}$, $\mathbf{z}$, and $\mathbf{t}$. We write $\mathbf{x} > \mathbf{y}$ if $\mathbf{x} \neq \mathbf{y}$, $x_1 \geq y_1$, and $x_2 \geq y_2$, and $\mathbf{x} \gg \mathbf{y}$ if $x_1 > y_1$ and $x_2 > y_2$.

The primitive is a choice correspondence $C : \mathscr{A} \rightrightarrows X$ where $C(A)$ is non-empty and $C(A) \subseteq A$ for each $A \in \mathscr{A}$. First, we define our model.

Definition 6 *A choice correspondence C is an* additive reference dependent choice *(ARDC) if there are strictly increasing functions $f, u_1, u_2 : \mathbb{R}_+ \to \mathbb{R}_+$ such that $f(0) = u_1(0) = u_2(0) = 0$ and for any menu $A \in \mathscr{A}$,*

$$C(A) = \arg\max_{\mathbf{x}\in A} \{f(u_1(x_1) - u_1(m^A_1)) + f(u_2(x_2) - u_2(m^A_2))\}. \tag{2.2}$$

Denote an ARDC with functions f, u_1, u_2 by $C_{(f,u_1,u_2)}$. Throughout the paper, we assume that f, u_1, u_2 are continuous and surjective. Note that when f is linear, ARDC reduces to the standard additive utility model:

$$C(A) = \arg\max_{\mathbf{x}\in A} \{u_1(x_1) + u_2(x_2)\}.$$

Now we define two induced preferences for a given C. The first preference $\succsim$ represents binary comparisons. For example, if $\mathbf{x}$ is chosen over $\mathbf{y}$ from the binary menu $\{\mathbf{x}, \mathbf{y}\}$, then we have $\mathbf{x} \succ \mathbf{y}$. The second preference is defined from choices on tripleton menus. For example, for a given alternative $\mathbf{t}$, if $\mathbf{x}$ is chosen over $\mathbf{y}$ in the presence of $\mathbf{t}$, then we have $\mathbf{x} \succ_{\mathbf{t}} \mathbf{y}$.

Definition 7 (Induced Preferences) *For any $\mathbf{x}, \mathbf{y}, \mathbf{t} \in X$,*

[5] It is not difficult to extend our model to the context of choices from menus with n-attribute alternatives. See Appendix B.2.

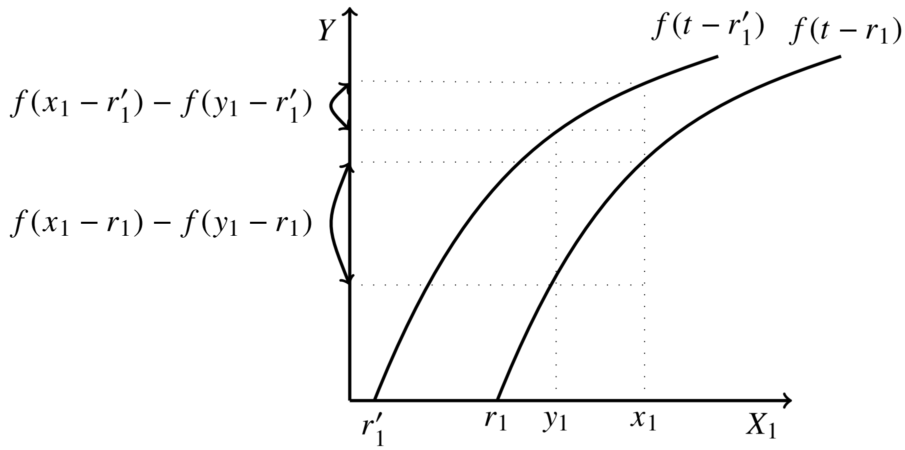

Figure 2.2: Diminishing Sensitivity When $u_1(x_1) = x$

i) $\boldsymbol{x} \succeq \boldsymbol{y}$ *if and only if* $\{\boldsymbol{x}\} \in C(\{\boldsymbol{x}, \boldsymbol{y}\})$ *and*

ii) $\boldsymbol{x} \succeq_{\boldsymbol{t}} \boldsymbol{y}$ *if and only if* $\{\boldsymbol{x}\} \in C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}\})$.

Rational choice theory relies on the weak axiom of revealed preference (WARP), which requires that if an alternative is not chosen from a menu, then eliminating that alternative does not effect the choice from the menu. Under WARP, $\mathbf{x} \succeq_{\mathbf{t}} \mathbf{y}$ implies $\mathbf{x} \succeq \mathbf{y}$ since WARP requires that the irrelevant third alternative $\mathbf{t}$ should not affect the comparison between $\mathbf{x}$ and $\mathbf{y}$. In this paper, we weaken WARP and allow for preference reversals such as $\mathbf{x} \succ \mathbf{y}$ and $\mathbf{y} \succ_{\mathbf{t}} \mathbf{x}$.

Finally, note that for any ARDC $C_{(f, u_1, u_2)}$, $\succeq$ satisfies transitivity ($\mathbf{x} \succeq \mathbf{y}$ and $\mathbf{y} \succeq \mathbf{z}$ imply $\mathbf{x} \succeq \mathbf{z}$) as illustrated in the following remark.

Remark 1. Take any binary menu $\{\mathbf{x}, \mathbf{y}\}$ with $x_1 > y_1$ and $x_2 < y_2$. Note that the reference point of $\{\mathbf{x}, \mathbf{y}\}$ is (y_1, x_2). Therefore, we have

$$\begin{aligned} \mathbf{x} \succeq \mathbf{y} \quad &\Leftrightarrow \quad f(u_1(x_1) - u_1(y_1)) + \cancel{f(u_2(x_2) - u_2(x_2))} \geq \\ &\qquad \cancel{f(u_1(y_1) - u_1(y_1))} + f(u_2(y_2) - u_2(x_2)) \\ &\Leftrightarrow \quad u_1(x_1) + u_2(x_2) \geq u_1(y_1) + u_2(y_2). \end{aligned}$$

2.2.1 Diminishing Sensitivity

Diminishing sensitivity is a widely used behavioral property in economics and will be a driving force of our results. We define it in terms of functional properties on f as in Tversky and Kahneman (1991). It requires that the marginal increase of utility is decreasing in the distance from the reference point.

Definition 8 (Diminishing Sensitivity) *An ARDC $C_{(f,u_1,u_2)}$ satisfies* diminishing sensitivity *if for any x_1, y_1, r_1, r'_1 such that $x_1 > y_1$ and $r_1 > r'_1$,*

$$f(u_1(x_1)-u_1(r_1))-f(u_1(y_1)-u_1(r_1)) > f(u_1(x_1)-u_1(r'_1))-f(u_1(y_1)-u_1(r'_1)).^6 \tag{2.3}$$

Diminishing sensitivity is equivalent to the strict concavity of f and it is illustrated in Figure 2.2.[7]

We can phrase diminishing sensitivity in the following way: the relative value of the i-th dimension increases when its reference increases. This interpretation will be useful when we discuss the compromise and attraction effects. Figure 2.3 demonstrates how indifference curves change as reference points change under diminishing sensitivity. Here curves $I_{\mathbf{r}}$ represent indifference curves for the reference point $\mathbf{r}$. For example, $I_{(1,1)}$ (solid curve) and $I_{(3,1)}$ (dashed curve) are indifference curves such that $\sqrt{x_1-1}+\sqrt{x_2-1}=2.7$ and $\sqrt{x_1-3}+\sqrt{x_2-1}=2$, respectively. As the reference for the first dimension 1 increases to 3 (the reference point $(1,1)$ shifts to $(3,1)$), the indifference curve $I_{(1,1)}$ (solid) becomes steeper since the first dimension is more valuable now and shifts to $I_{(3,1)}$ (dashed). On the other hand, as the reference for the second dimension 1 increases to 3 (the reference point $(1,1)$ shifts to $(1,3)$), the indifference curve $I_{(1,1)}$ becomes flatter since the second dimension is more valuable now and shifts to $I_{(1,3)}$ (dotted). This intuition will be the driving force of our results in Section 2.2.2 and Section 2.2.4.

Finally, note that the relative value of x_1 over y_1 in the presence of r_1 decreases when r'_1 is added since $f(u_1(x_1)-u_1(r_1))-f(u_1(y_1)-u_1(r_1)) > f(u_1(x_1)-u_1(r'_1))-f(u_1(y_1)-u_1(r'_1))$. Therefore, we can rephrase diminishing sensitivity in the following way: *adding r'_1 hurts x_1 over y_1 in the presence of* $r_1 \in (r'_1, y_1]$. In the introduction, we suggested a rationale that can explain both the compromise and

[6]By the additive nature of the model, diminishing sensitvity can be defined only using f and u_1. Indeed, diminishing sensitvity is equivalent to requiring $f(u_2(y_2)-u_2(r_2))-f(u_2(x_2)-u_2(r_2)) > f(u_2(y_2)-u_2(r'_2))-f(u_2(x_2)-u_2(r'_2))$ for any $y_2 > x_2$ and $r_2 > r'_2$.

[7]Diminishing sensitivity may be an implication of a general law of human perception in psychology called the Weber-Fechner law. The law states that perceived intensity is proportional to the logarithm of the stimulus (or to more general concave functions). In other words, the law states that people exhibit diminishing sensitivity to stimuli in general. This connection is suggested by Bruni and Sugden (2007) and also discussed in Bordalo et al. (2012) and Bordalo et al. (2013).

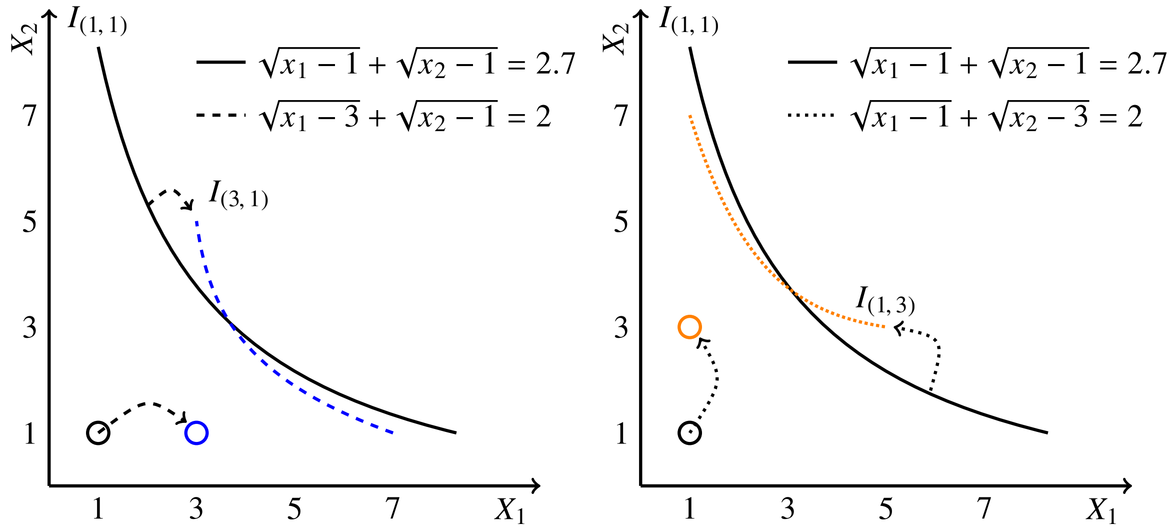

Figure 2.3: The Effect of a Shift in Reference Points When $f(t) = \sqrt{t}$.

attraction effects: *adding* z_1 *hurts* x_1 *over* y_1. This rationale can be obtained from diminishing sensitivity by setting $r_1 = y_1$ and $r_1' = z_1$. We call this rationale *weak diminishing sensitivity* and define it formally in the following way.

Definition 9 (Weak Diminishing Sensitivity) *An ARDC* $C_{(f,u_1,u_2)}$ *satisfies* weak diminishing sensitivity *if for any* x_1, y_1, *and* z_1 *with* $x_1 > y_1 > z_1$,

$$f(u_1(x_1) - u_1(y_1)) + f(u_1(y_1) - u_1(z_1)) > f(u_1(x_1) - u_1(z_1)). \qquad (2.4)$$

Now we will discuss the compromise and attraction effects and implications of diminishing sensitivity. Section 2.3.2 provides two behavioral postulates (on choices) equivalent to two diminishing sensitivities, and argues that diminishing sensitivity is not just sufficient, but also necessary for the two effects.

2.2.2 Implications of Diminishing Sensitivity

In this subsection, weak diminishing sensitivity plays an important role. First, we show the equivalence between the attraction and compromise effects and weak diminishing sensitivity (Proposition 2). Second, we show that more concavity (of f) means that we are more likely to observe the compromise and attraction effects (Proposition 3). Third, we discuss additional implications of weak diminishing sensitivity that are testable in lab experiments (Observation 1 and Proposition 4). In Section 2.2.3, we discuss the effects of adding a fourth alternative and symmetrically dominated third alternative, and their relation to the compromise and attraction effects. Explicit modelling of reference points (especially using minimums) will be useful to have predictions which are either consistent with experimental results or can be tested easily.

Although weak diminishing sensitivity is enough to rationalize the compromise and attraction effects, the intuition of diminishing sensitivity will be more useful to interpret the two effects (e.g., Figure 2.4). Moreover, in section 2.4, we will apply our model to two different contexts of decision making, and diminishing sensitivity will be necessary to obtain our results. Now we turn to the first implication of weak diminishing sensitivity.

Proposition 2 *Suppose $C = C_{(f,u_1,u_2)}$. Take any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$, and $\boldsymbol{z'} = (z_1, z_2') \in X$ such that $x_1 > y_1 > z_1$ and $x_2 < z_2' < y_2 < z_2$. Suppose z_2 is small enough such that $\boldsymbol{z}$ cannot be chosen from $\{\boldsymbol{x}, \boldsymbol{z}\}$. Then the following statements are equivalent.*

i) (Compromise effect) The alternative $\boldsymbol{x}$ is chosen over $\boldsymbol{y}$ from menu $\{\boldsymbol{x}, \boldsymbol{y}\}$, but $\boldsymbol{y}$ is chosen over $\boldsymbol{x}$ from $\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}$.

ii) (Attraction effect) The alternative $\boldsymbol{x}$ is chosen over $\boldsymbol{y}$ from menu $\{\boldsymbol{x}, \boldsymbol{y}\}$, but $\boldsymbol{y}$ is chosen over $\boldsymbol{x}$ from $\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z'}\}$.

iii) (Weak diminishing sensitivity) C exhibits weak diminishing sensitivity at x_1, y_1, z_1 in the following way:

$$f(u_1(x_1) - u_1(y_1)) > f(u_2(y_2) - u_2(x_2)) > f(u_1(x_1) - u_1(z_1)) - f(u_1(y_1) - u_1(z_1)).$$

The first two statements i) and ii) are the formal definitions of the compromise and attraction effects, respectively. Note that in the first case, adding **z** to $\{\mathbf{x}, \mathbf{y}\}$ makes **x** an extreme option and **y** a compromise option. In the second case, $\mathbf{z'} = (z_1, z_2')$ is dominated by **y** since $y_1 > z_1$ and $y_2 > z_2'$.

The intuitive argument behind Proposition 2 is given in Figure 2.4. Here $I_{(y_1, x_2)}$ (solid curve) is the indifference curve for the reference point (y_1, x_2) of $\{\mathbf{x}, \mathbf{y}\}$ and $I_{(z_1, x_2)}$ (dashed curve) is the indifference curve for the reference point (z_1, x_2) of $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ (also $\{\mathbf{x}, \mathbf{y}, \mathbf{z'}\}$). In both the compromise and attraction effects, adding a third alternative changes the reference point (y_1, x_2) to (z_1, x_2). Then by diminishing sensitivity, the second dimension becomes more valuable since the reference of the first dimension decreased. Therefore, the indifference curve $I_{(y_1, x_2)}$ (solid curve)

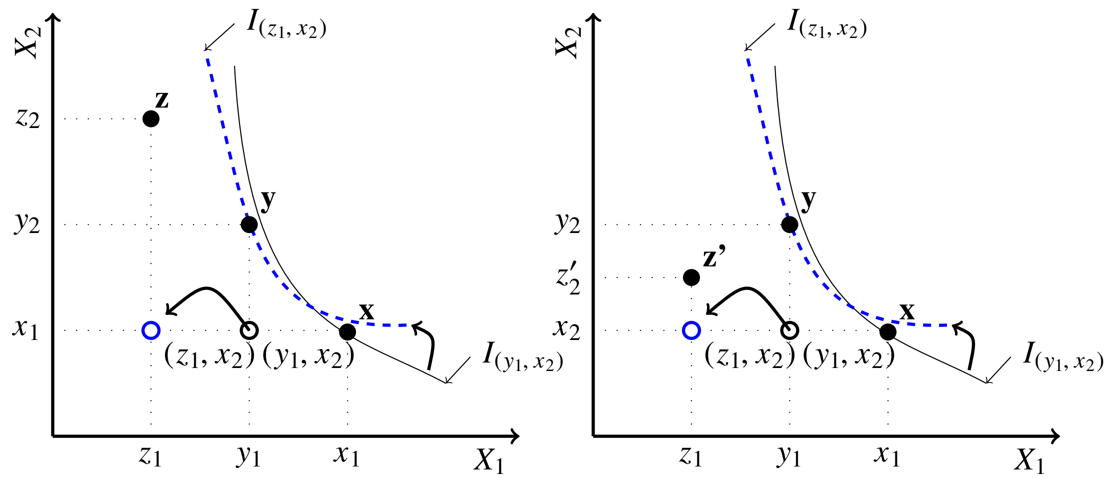

Figure 2.4: Compromise Effect, Attraction Effect, and Diminishing Sensitivity

becomes flatter and should shift left to $I_{(z_1,x_2)}$ (dashed curve), which helps **y** to be chosen over **x**.

Note that our explanations are consistent with common rationales for the two effects. A common rationale for the compromise effect is that people avoid extreme options. Similarly, in our model, the agent also avoids extremes because of the strict concavity of f. A rationale for the attraction effect is that a decoy **z'** makes **y** more attractive, i.e., the *decoy effect*. However, it must be that **y** becomes more attractive because of the second dimension rather than the first dimension since $y_1 < x_1$. Similarly, in our model, the second dimension becomes more valuable compared to the first dimension because of diminishing sensitivity. Next, we will show that the compromise and attraction effects are more likely to be observed when f is more concave.

Proposition 3 (More Concave, More Preference Reversals) *Take any two ARDCs $C_{(f,u_1,u_2)}$ and $C_{(f',u_1,u_2)}$ and $\{\succeq_t\}$ and $\{\succeq'_t\}$ are induced preferences, respectively. Suppose f' is more concave than f; that is, there is some strictly concave and increasing function h such that $f' = h(f)$. Then for any $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \in X$ and $z'_2 \in X_2$ such that $x_1 > y_1 > z_1$ and $x_2 < z'_2 < y_2 < z_2$, $\boldsymbol{x} \prec_{(z_1,z'_2)} \boldsymbol{y}$ implies $\boldsymbol{x} \prec'_{(z_1,z'_2)} \boldsymbol{y}$ and $\boldsymbol{x} \prec_z \boldsymbol{y}$ implies $\boldsymbol{x} \prec'_z \boldsymbol{y}$.*

This proposition shows that the concavity of f measures how likely we are to observe preference reversals. Indeed, when f is linear, our model reduces to the

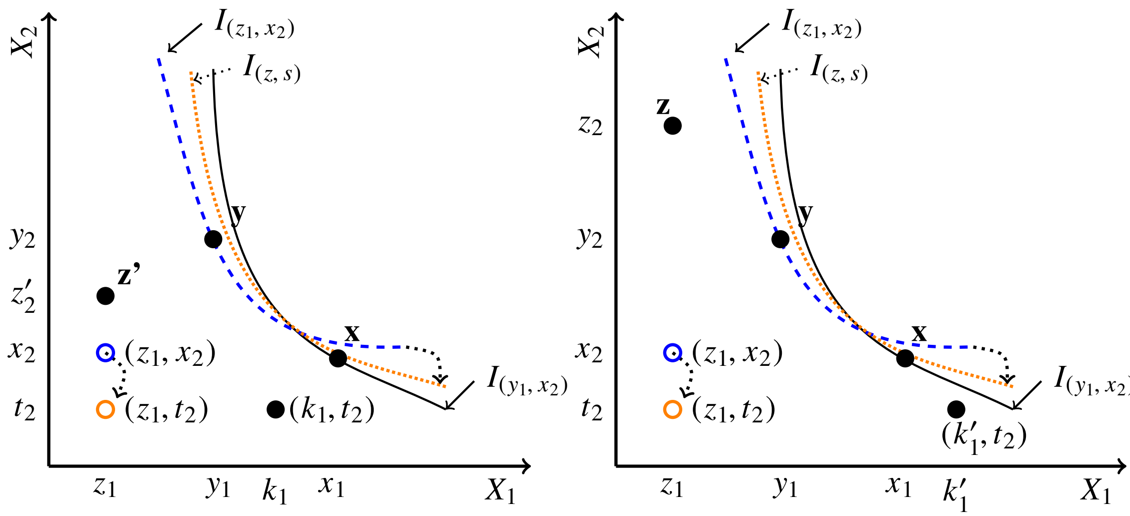

Figure 2.5: Two Decoy Effect and Two Compromise Effect

standard additive utility which does not allow preference reversals.

2.2.3 Implications of the Compromise and Attraction Effects, and Bounds on Preference Reversals

When can we say our model is a good model of the compromise and attraction effects? We argue that our model needs to i) be consistent with common explanations of the compromise and attraction effects; and ii) predict implications of the compromise and attraction effects. In the previous subsection, we argued that our explanations are consistent with common explanations of the compromise and attraction effects. We now will argue that predictions of our model are consistent with implications of the compromise and attraction effects.

Implication of the Attraction Effect – Two Decoy Effect: The usual explanation of the attraction effect says that **z'** works as a decoy for **y**, so **y** is chosen over **x** in the presence of **z'**. Then we can consider the following obvious implication of the attraction effect: If we introduce a fourth alternative which is a decoy for **x**, then the decoy effect of **z'** for **y** should be cancelled out by the decoy effect of the fourth alternative for **x**. We call it the *two decoy effect*, which is confirmed by the experimental result of Teppan and Felfernig (2009). In fact, our model is consistent with the two decoy effect, and the intuition is given in the left hand side of Figure 2.5.

To illustrate, suppose we add (k_1, t_2) to $\{\mathbf{x}, \mathbf{y}, \mathbf{z'}\}$ where (k_1, t_2) is dominated by

x (i.e., $k_1 < x_1$ and $t_2 < x_2$), but not by **y** (i.e., $y_1 < k_1$). The model predicts that the decoy effect of **z'** to **y** should be cancelled out with the decoy effect of (k_1, t_2) to **x**. Note that when (k_1, t_2) is added, the reference point shifts down to (z_1, t_2) since $s < p$. So the indifference curve $I_{(z_1, x_2)}$ (dashed curve) becomes steeper and shifts to $I_{(z, s)}$ (dotted curve) as the reference point (z_1, x_2) shifts to (z_1, t_2).[8] In other words, although z_1 hurts x_1 over y_1, it should be cancelled out by t_2 since t_2 also hurts y_2 over x_2. Therefore, it is more likely that **x** is chosen from {**x**, **y**, **z'**, (k_1, t_2)}.

Implication of the Compromise Effect – Two Compromise Effect: The usual explanation of the compromise effect says that **z** makes **x** an extreme option, so people compromise to **y**. Consider the following obvious implication of the compromise effect: If we add a fourth alternative which is more extreme than **x**, then **x** is not an extreme option anymore, so people would not avoid **x**. We call it the *two compromise effect*. In fact, our model is consistent with the two compromise effect, and the intuition is given in the right hand side of Figure 2.5.

To illustrate, suppose we add (k_1', t_2) to {**x**, **y**, **z**} where $x_1 < k_1'$ and $t_2 < x_2$. We will obtain an argument which is very similar to the two decoy effect. Note that when (k_1', t_2) is added, the reference point shifts down to (z_1, t_2) since $t_2 < x_2$. Note that the three indifference curves in the right hand side of Figure 2.5 are identical to the three indifference curves in the left hand side of Figure 2.5. Therefore, it is more likely that **x** is chosen from {**x**, **y**, **z**, (k_1', t_2)}.

Symmetric Dominance: The above arguments also suggest that if we add a symmetrically dominated alternative (z_1, t_2) to the binary menu {**x**, **y**}, we are less likely to observe a preference reversal, compared to the compromise and attraction effects (see Figure 2.6). In fact, this prediction is also consistent with the experimental results of Masatlioglu and Uler (2013). Observation 1 shows this argument formally.

Observation 1: Suppose $C_{(f, u_1, u_2)}$ satisfies weak diminishing sensitivity. For any **x**, **y** and z_1, z_2', and t_2 with $x_1 > y_1 > z_1$ and $y_2 > z_2' > x_2 > t_2$, if $\mathbf{x} \prec_{(z_1, t_2)} \mathbf{y}$, then $\mathbf{x} \prec_{(z_1, z_2')} \mathbf{y}$.

[8] Note that $I_{(z_1, t_2)}$ (dotted curve) is very similar to $I_{(y_1, x_2)}$ (solid curve) compared to $I_{(z_1, x_2)}$ (dashed curve).

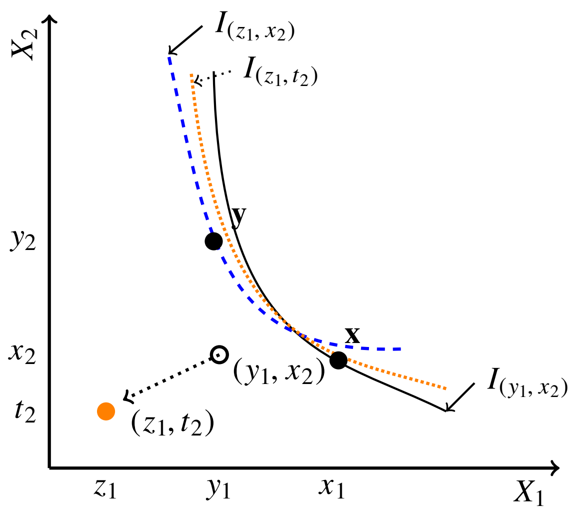

Figure 2.6: Symmetric Dominance

Intuitively, Observation 1 states that we are more likely to observe the attraction effect than to observe a preference reversal when the third alternative is symmetrically dominated.

Now let us summarize Proposition 2, the two decoy effect, the compromise effect, and the symmetric dominance (or Observation 1) in the following way (where $\dot{\geq}$ represents the likelihood of observing a preference reversal):

Compromise Effect $\doteq$ Attraction (Decoy) Effect $\dot{>}$

Two Decoy Effect $\doteq$ Two Compromise Effect $\doteq$ Symmetric Dominance.

Let us conclude the discussion of Symmetric Dominance by arguing that Observation 1 can be strengthen in some cases. The idea is that when the attributes of the third alternative z_1 and t_2 are small enough, the utility differences $u_1(x_1) - u_1(z_1)$ and $u_1(y_1) - u_1(z_1)$ and $u_2(y_2) - u_2(t_2)$ and $u_2(x_2) - u_2(t_2)$ are large enough, and f is close to linear at large values. Therefore, similar the fact that we cannot obtain a preference reversal when f is linear, we cannot obtain a preference reversal when z_1 and t_2 is small enough because f is close to linear.

We now show that weak diminishing sensitivity provides an additional behavioral prediction that can be easily tested in experimental settings. Roughly speaking, Proposition 4 provides an observable restriction on pairs of alternatives that can be involved in preference reversals.

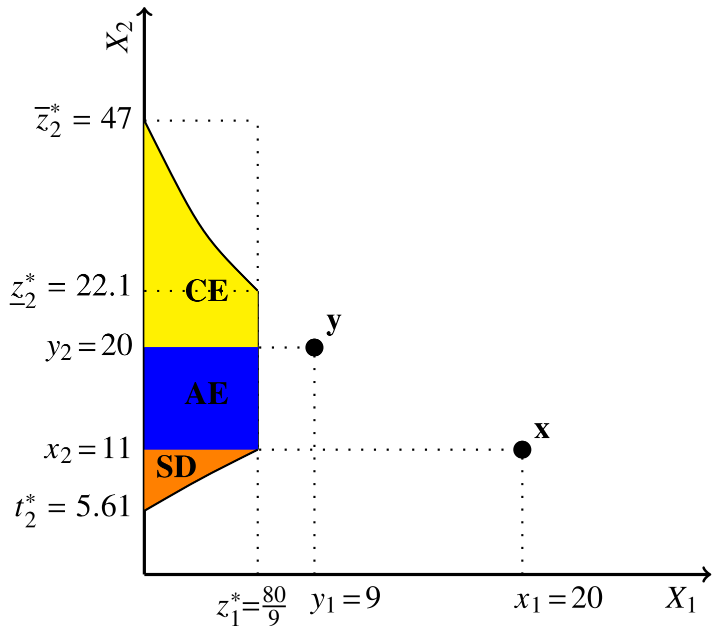

Figure 2.7: All Possible Third Alternatives That Can Cause a Preference Reversal

Proposition 4 (Preference Reversal) *Suppose $C_{(f,u_1,u_2)}$ satisfies weak diminishing sensitivity. For any alternatives $\boldsymbol{x}, \boldsymbol{y} \in X$ and menus $A, B \in \mathscr{A}$ such that $\boldsymbol{x}, \boldsymbol{y} \in A \cap B$, if $\boldsymbol{x} \in C(A)$ and $\boldsymbol{y} \notin C(A)$, but $\boldsymbol{y} \in C(B)$ and $\boldsymbol{x} \notin C(B)$, then either $\boldsymbol{x}$ is a non-extreme option of A or $\boldsymbol{y}$ is a non-extreme option of B.*[9]

We can check whether the attraction and compromise effects are consistent with Proposition 4. In both effects, we observe a preference reversal for a pair **x** and **y**. The two effects are consistent with Proposition 4 since **y** is a non-extreme option of {**x**, **y**, **z**} and {**x**, **y**, **z'**}.

We conclude Section 2.2.3 with the following numerical example that demonstrates the predictive power of the model. Suppose $u_1(t) = u_2(t) = t$, $f(t) = \sqrt{t}$, $\mathbf{x} = (20, 11)$, and $\mathbf{y} = (9, 20)$. Note that $\mathbf{x} > \mathbf{y}$. Then Figure 2.7 illustrates all possible third alternatives that can cause a preference reversal for the pair **x** and **y**. Note that in order to cause a preferences reversal, the first dimension of the third alternative should be smaller than $z_1^* = \frac{80}{9}$ and the second dimension of the third alternative should be in the interval $(t_2^*, \bar{z}_2^*) \approx (5.61, 47)$. In particular, the yellow shaded area is the set of **z** that can cause the compromise effect. The blue shaded area is the set of all asymmetrically dominated alternatives **z'** that can cause the

[9]We say an alternative **x** is an extreme option of $A \cup \{\mathbf{x}\}$ if either $x_1 < m_1^A$ or $x_2 < m_2^A$. So $\mathbf{x} \in A$ is a non-extreme option of $A \cup \{\mathbf{x}\}$ if $\mathbf{x} \geq \mathbf{m}^A$.

attraction effect. Finally, the orange shaded area is the set of all symmetrically dominated alternatives (z_1, t_2) that can cause a preference reversal. In other words, any alternative in the non-shaded area cannot cause a preference reversal.[10]

2.3 Behavioral Foundation

In this section, we axiomatically characterize our model by two novel axioms called *independence of non-extreme alternatives* (INEA) and *reference translation invariance* (RTI) in addition to three standard axioms (Theorem 4). The role of axioms and the sketch of the proof of the main theorem will be discussed in Section 2.3.3. We also characterize diminishing sensitivity by another novel axiom called *bound on the attraction effect* (BAE) (Proposition 5) in Section 2.3.2.

2.3.1 Axioms and Representation Theorem

We impose five axioms, and the first three of them are standard axioms. The first axiom, *regularity*, is a collection of three properties that guarantee a well-behaved representation.

Axiom 6 *(Regularity) A choice correspondence C satisfies the following three properties.*

i) (Monotonicity) For any $A \in \mathscr{A}$ and $\boldsymbol{x}, \boldsymbol{x'} \in A$, if $\boldsymbol{x'} > \boldsymbol{x}$, then $\boldsymbol{x} \notin C(A)$.

ii) (Continuity) $\succeq$ and $\{\succeq_{\boldsymbol{t}}\}_{\boldsymbol{t} \in X}$ are continuous; that is, $\{y \in X : y \succeq x\}$ and $\{y \in X : x \succeq y\}$ are closed, and $\{y \in X : y \succeq_{\boldsymbol{t}} x\}$ and $\{y \in X : x \succeq_{\boldsymbol{t}} y\}$ are closed for each $\boldsymbol{t}$.

iii) (Solvability) For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t} \in X$, for each $i \in \{1, 2\}$, there exists $x'_i \in \mathbb{R}_+$ such that $(x'_i, x_j) \succ \boldsymbol{y}$ and $(x'_i, x_j) \succ_{\boldsymbol{t}} \boldsymbol{y}$.

Monotonicity and continuity are standard properties. Solvability requires that (x'_i, x_j) can be preferred over $\mathbf{y}$ as long as x'_i is large enough. Solvability is rather technical, but frequently used in the literature. We also require transitivity of binary comparisons.

[10] All the implications of diminishing sensitivity we have discussed are true even if we weaken the additive structure of the model. In section 2.3.2, we characterize diminishing sensitivity by bounds on the attraction effect. If we weaken the additive structure of the model, bounds on the attraction effect will not be enough to characterize diminishing sensitivity. However, bounds on preference reversals will be sufficient to characterize diminishing sensitivity.

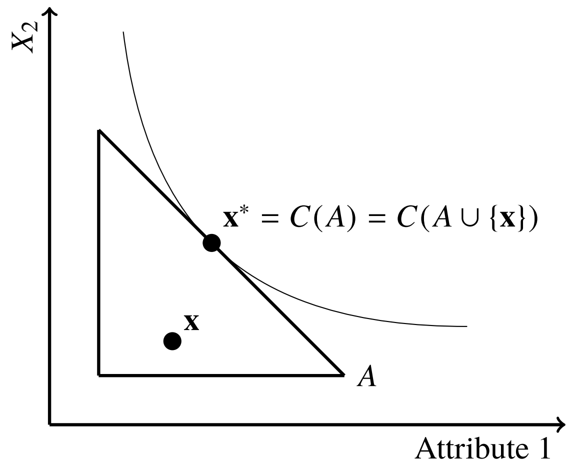

Figure 2.8: Independence of Non-Extreme Alternatives (INEA)

Axiom 7 (Transitivity) *For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$, if $\boldsymbol{x} \succeq \boldsymbol{y}$ and $\boldsymbol{y} \succeq \boldsymbol{z}$, then $\boldsymbol{x} \succeq \boldsymbol{z}$.*[11]

The next two axioms, INEA and RTI, are our main axioms and weakenings of the weak axiom of revealed preference (WARP). It is well known that WARP is necessary and sufficient for choice correspondence to be consistent with utility maximization. We can phrase WARP in the following way: for any A and $\mathbf{x} \notin A$, if $\{\mathbf{x}\} \neq C(A \cup \{\mathbf{x}\})$, then $C(A) = C(A \cup \{\mathbf{x}\}) \setminus \{\mathbf{x}\}$.[12] INEA is a postulate that is very similar to WARP and can be stated in the following way: for any non-extreme $\mathbf{x}$'s, WARP is satisfied, but for extreme $\mathbf{x}$'s, WARP can be violated.

Axiom 8 (Independence of Non-Extreme Alternatives (INEA)) *For any $A \in \mathscr{A}$ and $\boldsymbol{x} \notin A$,*

$$\text{if } \{\boldsymbol{x}\} \neq C(A \cup \{\boldsymbol{x}\}) \text{ and } \boldsymbol{x} \geq \boldsymbol{m}^A, \text{ then } C(A) = C(A \cup \{\boldsymbol{x}\}) \setminus \{\boldsymbol{x}\}.$$

INEA is illustrated in Figure 2.8. Suppose A is a triangle and $\mathbf{x}$ is inside the triangle. Suppose $\mathbf{x}^*$ is chosen from $A \cup \{\mathbf{x}\}$. Then INEA requires that $\mathbf{x}^*$ is also chosen from A. Under, regularity and INEA, we will obtain the following representation: there exists $\mathcal{W}$ such that

[11] In Appendix B.4, we weaken transitivity and obtain a general representation with two different distortion functions.

[12] Arrow (1959) stated WARP in the following way: for any $A, B \in \mathscr{A}$ with $B \subset A$, if $C(A) \cap B \neq \emptyset$, then $c(A) \cap B = C(B)$. If we state INEA in this way, then our representation theorem also works for all compact menus.

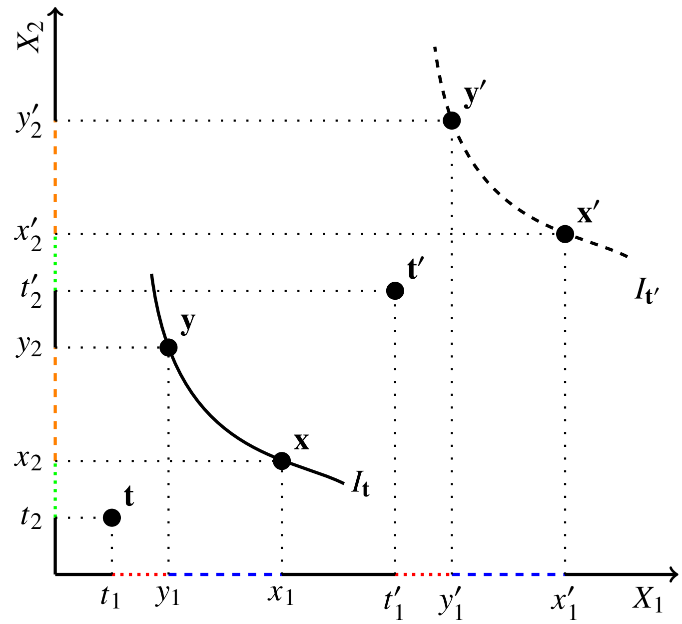

Figure 2.9: Reference Translation Invariance (RTI)

$$C(A) = \arg\max_{\mathbf{x}\in A} \mathcal{W}(\mathbf{x}, \mathbf{m}^A).$$

Therefore, the last two axioms impose more structure on $\mathcal{W}$.

We now define RTI, which is a modification of the standard *translation invariance*. The main idea of RTI is to connect two different preferences $\succeq_{\mathbf{t}}$ and $\succeq_{\mathbf{t}'}$ in the following way: if we can get $(\mathbf{x}', \mathbf{y}', \mathbf{t}')$ from $(\mathbf{x}, \mathbf{y}, \mathbf{t})$ by some distance-preserving shift (to be defined later), then we have $\mathbf{x} \succeq_{\mathbf{t}} \mathbf{y}$ iff $\mathbf{x}' \succeq_{\mathbf{t}'} \mathbf{y}'$. When we have linear utility functions, the above simply means that if $(\mathbf{x}', \mathbf{y}', \mathbf{t}') = (\mathbf{x} + \Delta, \mathbf{y} + \Delta, \mathbf{t} + \Delta)$ for some $\Delta \in \mathbb{R}^2$, then $\mathbf{x} \succeq_{\mathbf{t}} \mathbf{y}$ iff $\mathbf{x}' \succeq_{\mathbf{t}'} \mathbf{y}'$.

When we have nonlinear utility functions, distance-preserving shift should take account for preferences. So we formally define a notion of relative distance which takes account for nonlinear utilities.

Relative Distance: For any x_i, y_i, x_i', and y_i', we say a *relative distance between x_i and y_i is equivalent to that of x_i' and y_i'*, denoted by $[x_i, y_i]D_i[x_i', y_i']$, if for any x_j and y_j, $\mathbf{x} \sim \mathbf{y}$ if and only if $(x_i', x_j) \sim (y_i', y_j)$.

RTI is illustrated in Figure 2.9. Suppose for each i, the relative distance between x_i and y_i is equivalent to that of x_i' and y_i' (e.g., dashed intervals) and the relative distance between y_i and t_i is equivalent to that of y'_i and t'_i (e.g., dotted intervals). In other words, we can get $\mathbf{x}', \mathbf{y}', \mathbf{t}'$ from $\mathbf{x}, \mathbf{y}, \mathbf{t}$ by a distance-preserving shift. Then RTI requires that $\mathbf{x}$ is indifferent with $\mathbf{y}$ in the presence of $\mathbf{t}$ (solid curve $I_{\mathbf{t}}$) if and only if $\mathbf{x}'$ is indifferent with $\mathbf{y}'$ in the presence of $\mathbf{t}'$ (dashed curve $I_{\mathbf{t}'}$).

Axiom 9 (Reference Translation Invariance (RTI)) *For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}$, $\boldsymbol{x}'$, $\boldsymbol{y}'$, $\boldsymbol{t}' \in X$, if for each $i \in \{1, 2\}$, $[x_i, y_i]D_i[x'_i, y'_i]$, $[y_i, t_i]D_i[y'_i, t'_i]$, then*

$$\boldsymbol{x} \succsim_{\boldsymbol{t}} \boldsymbol{y} \text{ if and only if } \boldsymbol{x}' \succsim_{\boldsymbol{t}'} \boldsymbol{t}'.$$

RTI is much weaker than standard translation invariance. In fact, it is a weakening of WARP. First, note that by the definition of relative distances D_1 and D_2, $(x_1, x_2) \sim (y_1, y_2)$ iff $(x'_1, x_2) \sim (y'_1, y_2)$ iff $(x'_1, x'_2) \sim (x'_1, y'_2)$. Second, remember that WARP requires that the irrelevant third alternatives do not affect a comparison between a and b. Therefore, under WARP, $\mathbf{x} \sim_{\mathbf{t}} \mathbf{y}$ if and only if $\mathbf{x} \sim \mathbf{y}$ and $\mathbf{x}' \sim_{\mathbf{t}'} \mathbf{y}'$ if and only if $\mathbf{x}' \sim \mathbf{y}'$.

With the last axiom, called *cancellation*, in addition to transitivity, we can use existing methods to obtain additive representations. It is well known that two axioms are necessary and sufficient to have an additive representation: transitivity and cancellation (see Krantz et al. 1971, Fishburn and Rubinstein 1982, Wakker 1988, and Tversky and Kahneman 1991).[13] In particular, we use cancellation for $\succsim$ and $\succsim_{\mathbf{0}}$ where $\mathbf{0} = (0, 0)$. Although we need to have cancellation for each $\succsim_{\mathbf{t}}$, since the last axiom *Reference Translation Invariance* connects $\succsim_{\mathbf{t}}$ and $\succsim_{\mathbf{t}'}$ for any t and t', it turns out enough to have cancellation for only $\succsim_{\mathbf{0}}$.

Axiom 10 (Cancellation) *For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$,*

i) if $(x_1, x_2) \sim (y_1, z_2)$ and $(y_1, y_2) \sim (z_1, x_2)$, then $(x_1, y_2) \sim (z_1, z_2)$;

ii) if $(x_1, x_2) \sim_{\boldsymbol{0}} (y_1, z_2)$ and $(y_1, y_2) \sim_{\boldsymbol{0}} (z_1, x_2)$, then $(x_1, y_2) \sim_{\boldsymbol{0}} (z_1, z_2)$.

Figure 2.10 illustrates cancellation for $\succsim$. Solid and dashed curves represent indifferences. Intuitively, it requires that if the relative advantage of x_1 *over* y_1 is equivalent to that of z_2 *over* x_2 (dashed intervals) (i.e., $(x_1, x_2) \sim (y_2, z_2)$) and the relative advantage of y_1 *over* z_1 is equivalent to that of x_2 over y_2 (dotted intervals),

[13] Cancellation is sometimes called the Thomsen condition or double cancellation.

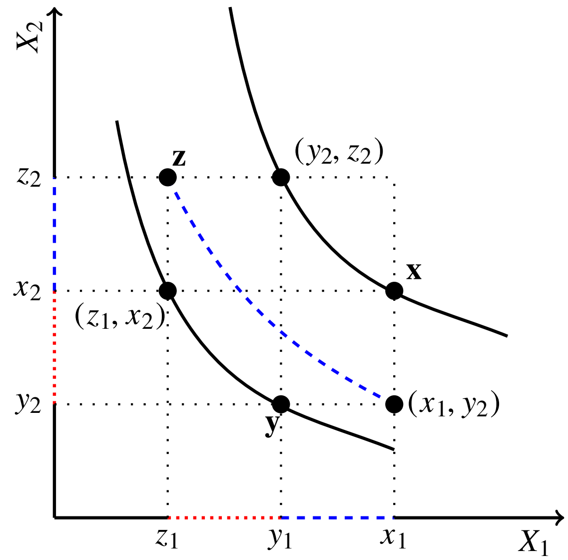

Figure 2.10: Cancellation

then the relative advantage of x_1 *over* z_1 is equivalent to that of z_2 *over* y_2 (combined intervals).

It turns out, RTI connects $\succeq_{\mathbf{t}}$ and $\succeq_{\mathbf{t'}}$ in the following way: $\succeq_{\mathbf{t}}$ satisfies cancellation if $\succeq_{\mathbf{0}}$ satisfies cancellation. Therefore, it is enough to impose cancellation on $\succeq_{\mathbf{0}}$ in cancellation (ii).

Finally, we can state our main theorem. Theorem 4 characterizes (2.2) and also provides a uniqueness result, which guarantees that utility functions have cardinal meaning.

Theorem 4 *C satisfies regularity, transitivity, cancellation, INEA, and RTI if and only if there exist strictly increasing and continuous functions $f, u_1, u_2 : \mathbb{R}_+ \to \mathbb{R}_+$ such that $f(\mathbb{R}_+) = u_1(\mathbb{R}_+) = u_2(\mathbb{R}_+) = \mathbb{R}_+$ and for any menu $A \in \mathscr{A}$,*

$$C(A) = \arg\max_{\boldsymbol{x}\in A}\{f(u_1(x_1) - u_1(m_1^A)) + f(u_2(x_2) - u_2(m_2^A))\}.$$

Moreover, for any two vectors of continuous functions (f, u_1, u_2) and (f', u_1', u_2') such that $f(1) = f'(1)$ and $u_1(1) = u_1'(1)$, if $C = C_{(f,u_1,u_2)} = C_{(f',u_1',u_2')}$, then $(f, u_1, u_2) = (f', u_1', u_2')$.

The uniqueness result can be stated in two steps. First, by Remark 1 in Section 2.2, we have $\mathbf{x} \succeq \mathbf{y}$ if and only if $u_1(x_1) + u_2(x_2) \geq u_1(y_1) + u_2(y_2)$. It is known that u_1 and u_2 are unique up to a linear transformation (e.g., see Krantz et al. 1971). Second, for given u_1 and u_2, it turns out f is also unique up to a linear transformation. In other words, after fixing $f(1)$ and $u_1(1)$, functions f, u_1, and u_2 are unique.

2.3.2 Characterizing Diminishing Sensitivity

In Section 2.2.2, we showed that diminishing sensitivity is sufficient to rationalize the compromise and attraction effects. In this subsection, we show that diminishing sensitivity is in fact necessary for the attraction effect by behaviorally characterizing two versions of diminishing sensitivity. Since Proposition 2 shows the equivalence between the attraction and compromise effects, we can say that diminishing sensitivity is also necessary for the compromise effect in our model. Now we will impose two novel axioms. The first axiom characterizes weak diminishing sensitivity while the second axiom characterizes diminishing sensitivity.

Axiom 11 (Weak Bound on the Attraction Effect (WBAE)) *For any $\boldsymbol{x}, \boldsymbol{y} \in X$ and $z_1 \in \mathbb{R}_+$ such that $x_1 > y_1 > z_1$, if $\boldsymbol{x} \sim_{(z_1, x_2)} \boldsymbol{y}$, then $\boldsymbol{x} \succ \boldsymbol{y}$.*

Roughly speaking, WBAE requires that (z_1, x_2) gives an advantage to $\mathbf{y}$ over $\mathbf{x}$. Evidently, WBAE is very similar to the attraction effect. In fact, we now show that WBAE is equivalent to observing the attraction effect in some neighborhood without assuming our model. Suppose $\mathbf{x} \sim \mathbf{y}$ and WBAE is satisfied. Then with continuity and monotonicity, there is $\epsilon > 0$ such that $(x_1 - \epsilon, x_2) \succ \mathbf{y}$ and $(x_1 - \epsilon, x_2) \prec_{(z_1, x_2)} \mathbf{y}$. Note that we obtained the attraction effect at alternatives $(x_1 - \epsilon, x_2)$, $\mathbf{y}$, and (z_1, x_2). Therefore, WBAE is equivalent to observing the attraction effect in a neighborhood of a triple $\mathbf{x}, \mathbf{y}$, and (z_1, x_2). In our model, since WBAE is equivalent to weak diminishing sensitivity, observing the attraction effect in some neighborhood is equivalent to weak diminishing sensitivity. Now we define the second axiom.

Axiom 12 (Bound on the Attraction Effect (BAE)) *For any $\boldsymbol{x}, \boldsymbol{y} \in X$ and $z_1, t_1 \in \mathbb{R}_+$ such that $x_1 > y_1 > z_1 > t_1$, if $\boldsymbol{x} \sim_{(z_1, x_2)} \boldsymbol{y}$, then $\boldsymbol{x} \prec_{(t_1, x_2)} \boldsymbol{y}$ and $\boldsymbol{x} \succ \boldsymbol{y}$.*

Roughly speaking, BAE requires that (z_1, x_2) helps $\mathbf{y}$ over $\mathbf{x}$, but (t_2, x_2) helps more when $t_1 < z_1$. Now we illustrate a connection between BAE and the attraction effect without assuming our model. Suppose $\mathbf{x} \sim_{(z_1, x_2)} \mathbf{y}$ and BAE is satisfied. With BAE, we have $\mathbf{x} \prec_{(t_1, x_2)} \mathbf{y}$ and $\mathbf{x} \succ \mathbf{y}$. Similar to the discussion of WBAE, by continuity and monotonicity, there is $\epsilon > 0$ such that $(x_1 + \epsilon, x_2) \succ_{(z_1, x_2)} \mathbf{y}$,

$(x_1 + \epsilon, x_2) \prec_{(t_1, x_2)} \mathbf{y}$, and $(x_1 + \epsilon, p) \succ \mathbf{y}$. Now we argue that BAE is summarzied by two properties. First, notice that we obtained the attraction effect at $(x_1 + \epsilon, x_2)$, $\mathbf{y}$, and (t_1, x_2). Second, $(x_1 + \epsilon, x_2) \succ_{(z_1, x_2)} \mathbf{y}$ implies that (t_1, x_2) is more likely to cause a preference reversal compared to (z_1, x_2) since $z_1 > t_1$. Therefore, BAE is equivalent to observing the attraction effect in the neighborhood of a triple $\mathbf{x}, \mathbf{y}$, and (t_1, x_2) while not in the neighborhood of a triple $\mathbf{x}, \mathbf{y}$, and (z_1, x_2). In other words, BAE dictates the degree of observing the attraction effect in some neighborhood (recall Proposition 3). In our model, since diminishing sensitivity is equivalent to BAE, diminishing sensitivity is equivalent to observing the attraction effect in the neighborhood of $\mathbf{x}, \mathbf{y}$, and (l, p) but not in that of $\mathbf{x}, \mathbf{y}$, and (z_1, x_2).

Finally, we show that (W)BAE characterizes (weak) diminishing sensitivity.

Proposition 5 *Suppose $C = C_{(f, u_1, u_2)}$ for strictly increasing continuous functions f, u_1, and u_2 such that $f(+\infty) = u_1(+\infty) = u_2(+\infty) = +\infty$. Then*

i) C satisfies WBAE if and only if it satisfies weak diminishing sensitivity.

ii) C satisfies BAE if and only if it satisfies diminishing sensitivity.

2.3.3 Sketch of the Proof of Theorem 4

Last, we briefly discuss the roles of INEA and RTI in the representation theorem. To do so, we need to know the role of the other three axioms: regularity, transitivity, and cancellation. First, with regularity, there are functions $\{V_A\}_{A\in\mathscr{A}}$, such that for any $A \in \mathscr{A}$,

$$C(A) = \arg\max_{\mathbf{x}\in A} V_A(\mathbf{x}).$$

Second, cancellation (i) (in addition to transitivity) on $\succeq$ is necessary and sufficient to have functions u_1 and u_2 such that $V_{\{\mathbf{x},\mathbf{y}\}}(\{\mathbf{x}\}) = u_1(x_1) + u_2(x_2)$; that is,

$$\mathbf{x} \succeq \mathbf{y} \text{ if and only if } u_1(x_1) + u_2(x_2) \geq u_1(y_1) + u_2(y_2).$$

Similarly, cancellation (ii) on $\succeq_{\mathbf{t}}$ (for a given $\mathbf{t}$) is necessary and sufficient to have functions $u^1_{\mathbf{t}}$ and $u^2_{\mathbf{t}}$ such that for any $A = \{\mathbf{x}, \mathbf{y}, \mathbf{t}\}$ with $\mathbf{x}, \mathbf{y} > \mathbf{t}$, $V_A(\mathbf{x}) = u^1_{\mathbf{t}}(x_1) + u^2_{\mathbf{t}}(x_2)$; that is,

$$\mathbf{x} \succeq_{\mathbf{t}} \mathbf{y} \text{ if and only if } u^1_{\mathbf{t}}(x_1) + u^2_{\mathbf{t}}(x_2) \geq u^1_{\mathbf{t}}(y_1) + u^2_{\mathbf{t}}(y_2).$$

Now note that there is no connection between utilities for different menus since the above conditions are separately imposed on $\succeq$ and $\succeq_{\mathbf{t}}$. In particular, there is no connection between the utilities for doubleton menus ($u_1 + u_2$), tripleton menus with dominated alternative $\mathbf{t}$ ($u^1_{\mathbf{t}} + u^2_{\mathbf{t}}$), and other menus with at least three alternatives A (V_A).

In other words, INEA and RTI connect $u_1 + u_2$, $u^1_{\mathbf{t}} + u^2_{\mathbf{t}}$, and V_A. We use two steps. In the first step, INEA connects $u^1_{\mathbf{t}} + u^2_{\mathbf{t}}$ and V_A. In particular, we show that the comparison between $\mathbf{x}$ and $\mathbf{y}$ in A is only affected by $\mathbf{m}^A$; that is, $V_A(\mathbf{x}) = V_{\{\mathbf{x},\mathbf{y},\mathbf{m}^A\}}(\mathbf{x})$ and $V_A(\mathbf{y}) = V_{\{\mathbf{x},\mathbf{y},\mathbf{m}^A\}}(\mathbf{y})$. In the second step, RTI connects $u_1 + u_2$ and $u^1_{\mathbf{t}} + u^2_{\mathbf{t}}$. In particular, we show that we can obtain $u^i_{\mathbf{t}}$ from u_i by some common distortion function f; that is, $u^i_{\mathbf{t}}(x_i) = f(u_i(x_i) - u_i(t_i))$. While the first step is naturally expected from INEA, the second step is a non-obvious implication of RTI because we need to prove that that $u^i_{\mathbf{t}}(x_i)$ is independent of t_j and is a function of the utility difference $u_i(x_i) - u_i(t_i)$.

Sketch of the proof of step 2: Most of the proof of Theorem 4 is devoted to proving the last step. Essentially we need to show that $V_{\{\mathbf{x},\mathbf{y},\mathbf{m}^A\}}(\mathbf{x}) = f(u_1(x_1) - u_1(m^A_1)) + f(u_2(x_2) - u_2(m^A_2))$ for some f. With cancellation (ii), we know that $V_{\{\mathbf{x},\mathbf{y},\mathbf{m}^A\}}(\mathbf{x}) = u^1_{\mathbf{m}^A}(x_1) + u^2_{\mathbf{m}^A}(x_2)$. Therefore, if we can prove that $V_{\{\mathbf{x},\mathbf{y},\mathbf{m}^A\}}(\mathbf{x}) = W(u_1(x_1) - u_1(m^A_1), u_2(x_2) - u_2(m^A_2))$ for some W (Lemma 5); then we can prove that $u^i_{\mathbf{m}^A}(x_i)$ is a function of $u_i(x_i) - u_i(m^A_i)$.

In order to prove Lemma 5, we will construct $W(u_1(x_1) - u_1(m^A_1), u_2(x_2) - u_2(m^A_2))$. In particular, for any $\mathbf{x}$ and $\mathbf{m}^A$, we find x^*_1 such that $(x^*_1, m^A_2) \sim_{\mathbf{m}^A} \mathbf{x}$ (Lemma 1) and set $W(u_1(x_1) - u_1(m^A_1), u_2(x_2) - u_2(m^A_2)) = u_1(x^*_1) - u_1(m^A_1)$. Finally, we need to make sure that W is well defined.

For example, we shall prove that for any x'_1 and $m^{A'}_1$, $u_1(x'_1) - u_1(m^{A'}_1) = u_1(x_1) - u_1(m^A_1)$ implies $W(u_1(x_1) - u_1(m^A_1), u_2(x_2) - u_2(m^A_2)) = W(u_1(x'_1) - u_1(m^{A'}_1), u_2(x_2) - u_2(m^A_2))$. In other words, we shall prove that if $u_1(x^*_1) - u_1(m^A_1) = u_1(x^{**}_1) - u_1(m^{A'}_1)$ for some x^{**}_1, then we have $(x^{**}_1, m^A_2) \sim_{(m^{A'}_1, m^A_2)} (x'_1, x_2)$. Now recall the definition of D_1. Note that $u_1(x'_1) - u_1(m^{A'}_1) = u_1(x_1) - u_1(m^A_1)$ is equivalent to $[x'_1, m^{A'}_1]D_1[x_1, m^A_1]$ and $u_1(x^{**}_1) - u_1(m^{A'}_1) = u_1(x^*_1) - u_1(m^A_1)$ is equivalent to $[x^{**}_1, m^{A'}_1]D_1[x^*_1, m^A_1]$. In other words, $(x^{**}_1, m^A_2), (x'_1, x_2), (m^{A'}_1, m^A_2)$ can be obtained from $(x^*_1, m^A_2), \mathbf{x}, \mathbf{m}^A$ by a distance-preserving shift. Therefore, RTI concludes the proof since $(x^*_1, m^A_2) \sim_{\mathbf{m}^A} \mathbf{x}$ if and only if $(x^{**}_1, m^A_2) \sim_{(m^{A'}_1, m^A_2)} (x'_1, x_2)$.

2.4 Applications

Explicit modeling of reference points allows us to apply our model to many different contexts. In order to demonstrate the usefulness of the model, we apply our model in two different contexts and diminishing sensitivity plays an important role in both cases. First, we discuss the standard intertemporal consumption choice. The main implication of the model is that if there is a borrowing constraint, then the consumer prefers to consume away from the constraint even if the constraint is not binding.

Second, we discuss risky choice. Our model can explain contradicting risk behavior (in the sense of the expected utility). More precisely, the model can allow behaviors such that an agent need to have both high and low degrees of risk aversion, i.e., the agent prefers riskier options in some cases and safer options in other cases.

2.4.1 Intertemporal Choice with Non-Binding Borrowing Constraint

Let us consider a two-period intertemporal choice model. A consumer lives two periods and her utility function is

$$U_{\mathbb{B}}(c_1, c_2) = (c_1 - c_1^{\mathbb{B}})^{\mu} + \beta(c_2 - c_2^{\mathbb{B}})^{\mu},$$

where β is a discount factor, c_i is consumption in period $i \in \{1, 2\}$, and $c_i^{\mathbb{B}}$ is the possible minimum consumption level in period i for the given budget set $\mathbb{B}$. Suppose $\mu < 1$, so we will have diminishing sensitivity.

The consumer earns income y_i in period i and she can borrow at most $\bar{b}$ in the first period at the interest rate z_2. We assume the consumer exhausts all her discounted total income $M \equiv y_1 + \frac{y_2}{1+r}$. Then the budget set is

$$\mathbb{B} = \{(c_1, c_2) \in \mathbb{R}_+^2 | \, c_1 + \frac{c_2}{1+r} = M \text{ and } c_1 \leq y_1 + \bar{b}\,\}.$$

Now we calculate the optimal consumption profile for two cases, without and with a borrowing constraint, and compare them.

Without a Borrowing Constraint: Suppose there is no borrowing constraint; that is, $\bar{b} \geq \frac{y_2}{1+r}$. Then the consumer's maximization problem is:

$$\max_{(c_1, c_2) \in \mathbb{B}} U_{\mathbb{B}}(c_1, c_2) = c_1^{\mu} + \beta c_2^{\mu},$$

where $\mathbb{B} = \{(c_1, c_2) \in \mathbb{R}_+^2 | \, c_1 + \frac{c_2}{1+r} = M\,\}$ and $(0, 0)$ is the reference point.

By a direct calculation, the optimal consumption levels are

$$c_1^* = \frac{M}{1 + \beta^{\frac{1}{1-\mu}}(1+r)^{\frac{\mu}{1-\mu}}} \text{ and } c_2^* = \frac{M \cdot \big(\beta(1+r)\big)^{\frac{1}{1-\mu}}}{1 + \beta^{\frac{1}{1-\mu}}(1+r)^{\frac{\mu}{1-\mu}}}.$$

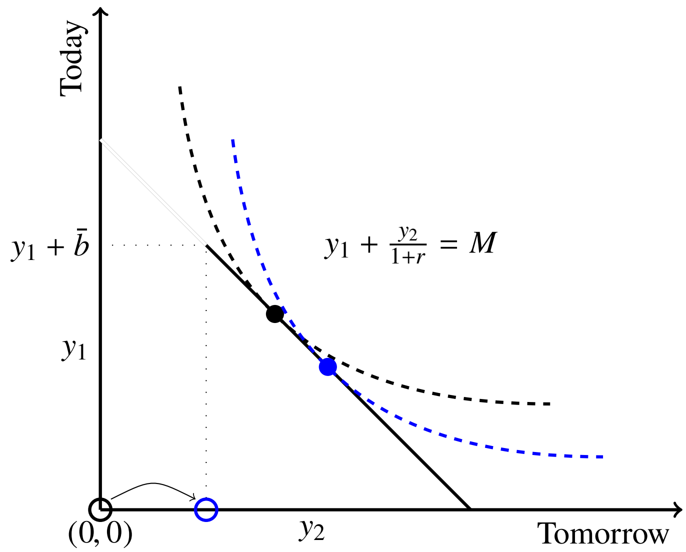

Figure 2.11: The Effect of a Non-Binding Borrowing Constraint

Note that, similar to the standard lifetime consumption model, the optimal consumption levels are proportional to M.

With a Borrowing Constraint: Now suppose there is a borrowing constraint; that is, $\bar{b} < \frac{y_2}{1+r}$. Then the consumer's maximization problem is:

$$\max_{(c_1,c_2)\in\mathbb{B}} U_{\mathbb{B}}(c_1, c_2) = c_1^{\mu} + \beta(c_2 - [y_2 - (1+r)\bar{b}])^{\mu}$$

where

$\mathbb{B} = \{(c_1, c_2) \in \mathbb{R}^2_+ | \, c_1 + \dfrac{c_2}{1+r} = M$ and $c_1 \leq y_1 + \bar{b}\,\}$ and $(0, y_2 - (1+r)\bar{b})$ is the reference point.

By a direct calculation, the new optimal consumption levels are

$$c_1^{**} = \frac{M - (\frac{y_2}{1+r} - \bar{b})}{1 + \beta^{\frac{1}{1-\mu}}(1+r)^{\frac{\mu}{1-\mu}}} = \frac{y_1 + \bar{b}}{1 + \beta^{\frac{1}{1-\mu}}(1+r)^{\frac{\mu}{1-\mu}}} \text{ and } c_2^{**} = \frac{M \cdot (\beta(1+r))^{\frac{1}{1-\mu}} + (y_2 - \bar{b}(1+r))}{1 + \beta^{\frac{1}{1-\mu}}(1+r)^{\frac{\mu}{1-\mu}}}.$$

In the standard model, the optimal consumption profiles are the same when the constraint is not binding; i.e., $y_1 + \bar{b} > c_1^*$. However, in our model, the two cases are different even if the constraint is not binding (see Figure 2.11). The first implication of the model and diminishing sensitivity is the effect of non-binding constraint on the optimal consumption profile.

Observation 2: The consumer decreases her first period consumption; that is, $c_1^* > c_1^{**}$, even if the borrowing constraint is not binding ($c_1^* < y_1 + \bar{b}$). Moreover, as the constraint gets tighter (as $\bar{b}$ decreases), she consumes less in the first period. In other words, c_1^{**} is increasing in $\bar{b}$.

The intuition of Observation 2 is as follows. The borrowing constraint decreases the consumer's the maximum consumption level for today and increases minimum consumption level for tomorrow. Then by diminishing sensitivity ($\mu < 1$), the relative value of tomorrow's consumption with respect to today's consumption will increase. Therefore, the consumer will decrease her consumption today and increase consumption tomorrow.

It is common to explain deviations from the standard model of consumer choice such as *excess sensitivity of consumption to income* and *a hump-shaped consumption profile* with liquidity or borrowing constraints (see Attanasio 1999 and Attanasio and Weber 2010). However, there is no strong empirical evidence that constraints were actually binding at the time of the decision. For example, Jappelli (1990) directly asked consumers whether they applied for and were denied credit and only 12.5 percent of 1982 respondents answered that they were denied credit. Moreover, Deaton (1991) shows that liquidity constraints are rarely binding by simulation. Our model gives a justification for those explanations since liquidity constraints can have an effect on choices even if they are not binding.[14]

Observation 3: c_i^{**} is positively correlated with y_i, fixing M constant. Moreover, $c_i^{**} = \alpha_i y_i + \beta_i$ for some α_i and β_i.

Contrast to permanent income hypothesis, it is well known that consumption depends on current income. For example, Figure 2.12, borrowed from Attanasio and Weber (2010), reports life-cycle profiles for two education groups in the UK. The left hand side of Figure 2.12 shows income and consumption paths for the group with compulsory education; the dotted curve is disposable income and the solid curve is nondurable consumption. Note that the consumption path closely follows the income path. Moreover, Campbell and Mankiw (1991) showed that in many different countries, a large number of consumers who follow a "rule of thumb" set

[14] A similar behavior (a non-binding constraint matters) is obtained by Hayashi (2008) in the context of choice under ambiguity (i.e., choice over Anscombe-Aumann acts). He also considers a menu-dependent choice, and in his model, an agent responds to non-binding constraints because of anticipated ex post regrets. It is not obvious how to relate our results with Hayashi (2008) since regret aversion is defined in the context of choice under uncertainty and naturally has a dynamic interpretation, while diminishing sensitivity is defined in a deterministic environment.

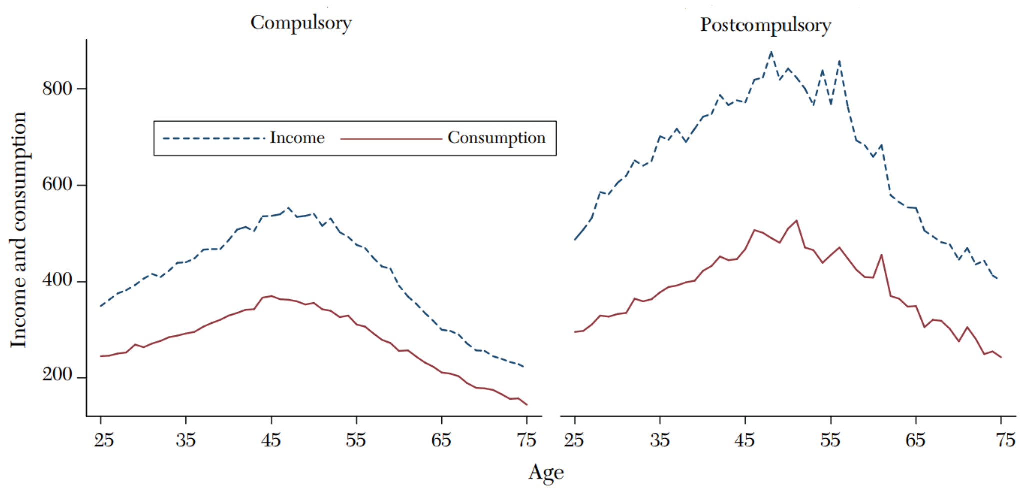

Figure 2.12: Average Income and Consumption by Education, Attanasio and Weber (2010)

their consumption proportional to their income. Observation 3 shows that our model is consistent with the above empirical regularities on consumption paths.

We could apply Observation 3 in a different context. For example, in a context of financial decision, it is known that younger people invest in risky asset less than older people even after controlling the income effect via risk preference. Observation 3 suggests an additional income effect to younger people's financial decisions. Therefore, the income effect via reference-dependent preference might be an additional cause for younger people to invest less.

We can also make another interesting observation: when $\bar{b}$ is small, c_1^{**} is more dependent on y_1. In fact, Zeldes (1989) observed that consumers with a low level of assets (low $\bar{b}$) more tightly link their consumption to their income.

2.4.2 Observing Different Degrees of Risk Aversion

Now we turn to risky choice. In particular, we focus on binary lotteries $\mathbf{x}$ where $\mathbf{x}$ denotes a binary lottery that gives x_1 dollars with probability x_2 and nothing with probability $1 - x_2$. Now recall that the utility of $\mathbf{x}$ for a given menu A:

$$U_A\mathbf{x} = f(u_1(x_1) - u_1(m_1^A)) + f(u_2(x_2) - u_2(m_2^A)).$$

The objective of this subsection is to obtain contradicting risk behavior: an agent makes a risky choice in case 1 and makes a safe choice in case 2 and the estimated degrees of risk aversion (in the sense of EUT) from two cases contradict each other.

Note that by diminishing sensitivity (the strict concavity of f), f introduces an additional concavity to the total value of alternatives since utilities are distorted by f. Moreover, the additional concavity is menu-dependent. Therefore, an agent will act as if her degree of risk aversion (in the standard sense) is high in some menus and as if it is low in other menus. Now let us show the above intuition more formally.

Suppose an agent has the following additive reference-dependent preference: For any menu A of binary lotteries and binary lottery $\mathbf{x} \in A$,

$$U(\mathbf{x}|A) = (\alpha \log(x_1) - \alpha \log(m_1^A))^\mu + (\log(x_2) - \log(m_2^A))^\mu.$$

Here α is the degree of risk aversion (in the standard sense) because when $\mu = 1$, the model reduces to the expected utility theory: the agent maximizes expected utility $x_2 \cdot x_1^\alpha$. Moreover, μ is independent from the agent's risk preference since the concavity of f is motivated by riskless choices.

Take four binary lotteries $\mathbf{x}$, $\mathbf{y}$, $\mathbf{x'}$, and $\mathbf{y'}$ with $x_1 > y_1$ and $x_2 < y_2$ and $x'_1 > y'_1$ and $x'_2 < y'_2$. Suppose the agent prefers $\mathbf{x}$ over $\mathbf{y}$ in the menu $\{\mathbf{x}, \mathbf{y}\}$ (choosing the riskier option), but prefers $\mathbf{y'}$ over $\mathbf{x'}$ in some menu A (choosing the safer option). In the standard model, the two choices contradict each other when $\frac{\log(y'_2/x'_2)}{\log(x'_1/y'_1)} < \frac{\log(y_2/x_2)}{\log(x_1/y_1)}$ because choosing $\mathbf{x}$ over $\mathbf{y}$ implies that $\alpha > \frac{\log(y_2/x_2)}{\log(x_1/y_1)}$ (the agent is risk loving), but choosing $\mathbf{y'}$ over $\mathbf{x'}$ implies that $\alpha < \frac{\log(y'_2/x'_2)}{\log(x'_1/y'_1)}$ (the agent is risk averse).

However, in our model the agent could choose $\mathbf{y'}$ over $\mathbf{x'}$ from A because $f(x_1) = x^\mu$ introduces an additional menu-dependent concavity to utilities of alternatives. To illustrate, note that choosing $\mathbf{x}$ over $\mathbf{y}$ implies $\alpha > \frac{\log(y_2/x_2)}{\log(x_1/y_1)}$, as in the standard model. However, choosing $\mathbf{y'}$ over $\mathbf{x'}$ from A implies

$$\alpha < \left(\frac{(\log(y'_2/m_2^A))^\mu - (\log(x'_2/m_2^A))^\mu}{(\log(x'_1/m_1^A))^\mu - (\log(y'_1/m_1^A))^\mu} \right)^{\frac{1}{\mu}} \equiv \alpha(\mu).$$

If the reference point of A satisfies $\frac{\log(x'_1/m_1^A)}{\log(y'_1/m_1^A)} < \frac{\log(y'_2/m_2^A)}{\log(x'_2/m_2^A)}$, then $\alpha(\mu)$ converges to infinity as $\mu \to 0$. Therefore, as long as $\frac{\log(x'_1/m_1^A)}{\log(y'_1/m_1^A)} < \frac{\log(y'_2/m_2^A)}{\log(x'_2/m_2^A)}$ and μ is small enough,[15] we can have $\frac{\log(y_2/x_2)}{\log(x_1/y_1)} < \alpha < \alpha(\mu)$. Therefore, when μ is small, the agent could act as if she is very risk-averse even if α is large.

[15]For example, let $A = \{\mathbf{x'}, \mathbf{y'}, \mathbf{z'}\}$ where $x'_1 > y'_1 > z'_1$ and $x'_2 < y'_2 < z'_2$. Since (z'_1, x'_2) is the reference point of A; i.e., $(m_1^A, m_2^A) = (z'_1, x'_2)$, the inequality $\frac{\log(x'_1/m_1^A)}{\log(y'_1/m_1^A)} = \frac{\log(x'_1/z'_1)}{\log(y'_1/z'_1)} < \frac{\log(y'_2/m_2^A)}{\log(x'_2/m_2^A)} = \frac{\log(y'_2/x'_2)}{\log(x'_2/x'_2)} = \infty$ always holds.

Last, we briefly discuss the estimation of risk preference when the agent has an additive reference-dependent preference. Because of diminishing sensitivity or the strict concavity of f, measuring risk preferences from observed choices could be misleading. However, if we estimate risk preferences only using binary comparisons, then it provides an unambiguous measure for the degree of risk because the effect of f is cancel out (recall Remark 1 on $\geq$ in Section 2.2).

2.5 Model with General Menu-Dependent References

As we mentioned in the introduction, (2.1) is too general to have testable implications on observed choice behavior since reference points are exogenously given. The objective of the current paper is to endogenize reference points and so we focused on menu-dependent reference points. In this section, we show that, in fact, we do not lose any generality by focusing on menu-dependent reference points. In other words, we will show that any observed choice data can be rationalized by a model with menu-dependent reference points. This result also suggests that we should have more specific models for reference points in order to obtain testable predictions.

For the sake of simplicity, suppose we observe choices from a finite number of finite menus $A_1, A_2, \ldots, A_N$. Suppose $C(A_n) \in A_n$ for each n. The next result shows that any such choices can be rationalized by a reference-dependent model with menu-dependent reference points as long as monotonicity is satisfied.

Proposition 6 *If C satisfies monotonicity, then there exist $\lambda > 0$ and reference-points $\boldsymbol{r}^1, \boldsymbol{r}^2, \ldots, \boldsymbol{r}^N$ such that for each n, $\boldsymbol{r}^n \in A_n$ and*

$$C(A_n) = \arg\max_{\boldsymbol{x} \in A_n} f(x_1 - r_1^n) + f(x_2 - r_2^n) \textit{ where } f(t) = \begin{cases} t & \textit{if } t \geq 0. \\ -\lambda t & \textit{if } t < 0. \end{cases}$$

In fact, we can rationalize observed choices by a model with a very specific functional form. Therefore, Proposition 6 also suggests that a specific functional form may not give us more predictive power without the further specification of menu-dependence.

In this paper, we took the specific approach of using minimums of menus as a reference point. Another factor to determine reference points might be the maximums of the menu. However, as mentioned in the introduction, models in which reference points only depend on the maximums cannot explain the attraction effect. Although there are many other possibilities, as long as reference points are increasing in the minimums, it is possible to rationalize the compromise and attraction effects. But behavioral predictions of the general models will be weaker compared to only using the minimums. In Appendix B.3, we will consider a model in which reference points depend on both the maximums and the minimums. The main implications stay the same, but behavioral predictions are weaker.

2.6 Related Literature

Tversky and Kahneman (1991) provided the first explicit model of reference-dependent preferences. Their model is an extension of the prospect theory of Kahneman and Tversky (1979) in the context of riskless choice. The model of Tversky and Kahneman (1991) is further developed by Munro and Sugden (2003) (to the n-dimensional commodity space), Sugden (2003) (to subjective expected utility), and Koszegi and Rabin (2006, 2007) (expectation-based reference points, to be discussed later).

The main objective of the current paper is to develop an axiomatic model of reference-dependent preferences with endogenous reference points. However, in most of the previous work, reference points are exogenously given. To the best of our knowledge, Ok et al. (2014) is the only paper that axiomatically develops a model with purely endogenous reference points. As we explained in the introduction, the authors weaken WARP to allow for choice behavior that exhibits the attraction effect, but not the compromise effect.

The literature related to reference-dependent behavior and prospect theory is too large to be discussed here (see Barberis 2013). We will now narrow our focus and discuss related literature in the following three ways. First, we discuss two main approaches to model reference points in the literature. Second, we discuss papers that accommodate behavioral phenomena we study in this paper. Finally, we attempt to place our model in the broad choice theory literature.

1. Two main approaches to model reference points:

In the first approach, an agent uses an exogenously given alternative as a reference point. For example, that alternative is the status quo, the default option (Masatlioglu and Ok 2005, Sagi 2006, and Apesteguia and Ballester 2009), or the initial endowment (Masatlioglu and Ok 2013). However, there are many choice situations where no alternative is exogenously given as the status quo or the initial endowment. For example, there is no sensible status quo or initial endowment in the compromise and attraction effects. Since in our model the reference point can be defined for any given menu, our approach complements their approach in situations where there is no sensible default option or initial endowment.

In the second approach, an agent has probabilistic beliefs over possible outcomes and she uses her expectations of the outcome as a reference point (Koszegi and Rabin 2006, 2007). This approach builds on Tversky and Kahneman (1991), as did we. Moreover, it is more natural in repeated or dynamic choice. However, there are two crucial differences. First, in their approach, the reference point is not purely endogenous because probabilistic beliefs are not directly observable, whereas in our model a reference point is completely determined by the menu. Second, their model reduces to the standard rational model when there is no underlying uncertainty. Therefore, their model cannot allow reference-dependent behavior such as the compromise and attraction effects since the effects are normally defined in a riskless environment. On the other hand, our model can be applied to both risky and riskless environments (Section 2.4.2).

2. Papers that discuss behavioral phenomena we study:

In the literature, the attraction and compromise effects are usually used as motivations for menu-dependent or context-dependent preferences. Therefore, it is already well known that these two effects can be explained by menu-dependent preferences or context-dependent preferences (e.g., Simonson and Tversky 1992, Wernerfelt 1995, Kamenica 2008, and Bordalo et al. 2013).[16] However, there is a common understanding among psychologists that the attraction and compromise effects are not separate phenomena, but rather two manifestations of the same behavior. To the best of our knowledge, this is the first paper that formally shows that, not only are the two effects implications of diminishing sensitivity, but they are also equivalent to diminishing sensitivity.[17]

It might be useful to compare our model with information-theoretic models of consumer choice (Wernerfelt 1995 and Kamenica 2008) that are consistent with the two effects. For example, Kamenica (2008) studies a model in which there is a market

[16]Moreover, there are a number of recent works on stochastic choice that are consistent with the two effects. For example, see Natenzon (2010), Echenique et al. (2013), and Fudenberg et al. (2013a).

[17]De Clippel and Eliaz (2012) rationalize the attraction and compromise effects as the result of a single bargaining protocol (to be discussed in part 3). However, behavioral postulates leading to the two effects are different.

with rational consumers who learn values of attributes of a good from a product line. Consumers obtain menu-dependent utility as a result of a market equilibrium that leads to preference reversals. However, there is evidence of the compromise and attraction effects that information-theoretic models cannot fully explain. First, in information-theoretic models, it is important that consumers face the binary menu first and the tripleton menu second in order to have different information in two choice situations (because consumers cannot unlearn information after facing the tripleton menu). However, the two effects are observed even if a decision maker sees the tripleton menu first and the binary menu second (e.g., see Sivakumar and Cherian 1995 and Wiebach and Hildebrandt 2012). Second, the two effects are robustly exhibited in non-market situations (e.g., choice over binary lotteries, Herne 1999, and choice over bundles of chewing gum and chocolate cookies, Herne 1998), and learning the values of attributes from a product line does not seem to be the main factor of preferences reversals.

Now, we narrow our focus to reference-dependent models which discuss either the attraction effect, the compromise effect, life-cycle consumption profile, or risky choice. Koszegi and Rabin (2009) and Pagel (2013) apply the expectation-based reference dependent model of Koszegi and Rabin (2006) to life-cycle consumption choice. They explain empirical observations about consumption profiles, including excess sensitivity of consumption to income and a hump-shaped consumption profile. The main difference is that we rely on a non-binding borrowing constraint while their models rely on uncertainty since the expectation-based reference dependent model reduces to the standard model when there is no uncertainty.[18]

Explaining two different risk attitudes and introducing reference-dependent behavior are two of the most important contributions of Kahneman and Tversky (1979). As mentioned in the introduction, loss aversion and diminishing sensitivity help to explain many different anomalies in risky choice (e.g., Benartzi and Thaler 1995, Koszegi and Rabin 2007 and Bordalo et al. 2012). One contribution of our paper is to obtain two different risk behaviors relying on diminishing sensitivity, but not relying on loss aversion.

Last, the salience theory of Bordalo et al. (2013) is important to mention in detail. Salience theory focuses on alternatives with two attributes, quality and price. It generalizes standard models in the way that an agent uses menu-dependent weight

[18]Moreover, in some of Pagel (2013)'s results, she also needs to have a hyperbolic discounting agent.

functions, w_q and w_p, that distort evaluations of attributes of alternatives: if the quality dimension of an alternative is more salient than the price dimension, then the agent overweights quality and underweights price ($w_q > 1 > w_p$) and vice versa ($w_q < 1 < w_p$).[19] In their model, the agent uses the average quality ($\bar{q}$) and the average price ($\bar{p}$) of the menu as references in order to decide the saliency of dimensions. In particular, for a given menu $A = \{(q_k, p_k)_{k=1}^n\}$,

$$C(A) = \arg\max_k \Big\{ \underbrace{w_q(\sigma(q_k, \bar{q}), \sigma(p_k, \bar{p})) \cdot q_k}_{=V(q_k, p_k, \bar{q}, \bar{p})} - w_p(\sigma(q_k, \bar{q}), \sigma(p_k, \bar{p})) \cdot p_k \Big\}$$

where the saliency of dimensions are determined by numbers $\sigma(q_k, \bar{q})$ and $\sigma(p_k, \bar{p})$ for some function σ.

Our model and salience theory have two important similarities. Let us consider the utilities of the first dimension in our model and salience theory: $f(u_1(x_1) - u_1(m_1^A))$ and $V(q_k, p_k, \bar{q}, \bar{p})$. First, both terms depend on some reference points that are purely determined by the menu, m_1^A and $(\bar{q}, \bar{p})$. Second, diminishing sensitivity properties on two terms (in particular, on f and σ) can explain the attraction and compromise effects.

However, salience theory cannot provide bounds on preference reversals. In particular, salience theory predicts that if (z_1, z_2') leads to the attraction effect, then for any $t_2 \in [0, z_2']$, (z_1, t_2) must cause a preference reversal (See Proposition 4.i of Bordalo et al. 2013). Therefore, salience theory is not consistent with the two decoy effect and symmetric dominance.

3. Choice theory literature:

In this paper, we weakened WARP in order to obtain a model that is consistent with observed behavioral phenomena. This approach is in the same spirit as a body of work that seeks to characterize models of non-standard choice in terms of direct axioms on choice behavior. This body of work includes, for example, rational shortlisting (Manzini and Mariotti 2007), choosing the uncovered set of some preference relation (Ehlers and Sprumont 2008), choosing two finalists (Eliaz et al. 2011), choosing an alternative that survived after a sequential elimination of alternatives by lexicographic semiorders (Manzini and Mariotti 2012b), multiself (De Clippel and Eliaz 2012), choosing from a subset of a menu because of limited attention (Masatlioglu et al. 2012), comparing all pairs by certain orders (Apesteguia

[19] The authors discuss salience theory in the context of choice under risk in Bordalo et al. (2012).

and Ballester 2013), and considering only alternatives that (belong to the best category (Manzini and Mariotti 2012a), optimal according to some rationalizing criteria (Cherepanov et al. 2013), and pass some menu-dependent threshold (Manzini et al. 2013)). The main difference between our paper and the above papers is that, except for the fact that we study a reference-dependent behavior, attributes of alternatives are known,[20] which allows us to obtain stronger predictions (e.g., bounds on preference reversals as in Section 2.2.3), while the above papers study alternatives with unknown attributes (as in abstract choice theory).[21]

The paper by De Clippel and Eliaz (2012) is important to mention; it axiomatically models choices as a solution to an intrapersonal bargaining problem among two selves of an individual. Each self is endowed with a preference relation, so having two preference relations for two selves is similar to having two-attribute alternatives. However, their model requires that $C(\{\mathbf{x}, \mathbf{y}\}) = \{\mathbf{x}, \mathbf{y}\}$ for any $x_1 > y_1$ and $x_2 < y_2$. Therefore, technically, they allow only “weak" forms of the attraction and compromise effects: the original two alternatives ($\mathbf{x}$ and $\mathbf{y}$) are indifferent and the third alternative breaks the tie (in favor of $\mathbf{y}$.)[22]

[20]This is a common assumption in economics. For example, in consumer theory, an alternative is a bundle of goods, and the i-th dimension represents the quantity of the i-th good. In social choice, an alternative is an allocation, and the i-th dimension represents the amount of money allocated to the i-th person. In finance, an alternative is a state-contingent prospect, and the i-th dimension represents the return at the i-th state.

[21]In fact, only Masatlioglu et al. (2012), De Clippel and Eliaz (2012), and Manzini et al. (2013) can be consistent with the attraction and two decoy effects. Although Manzini and Mariotti (2012a) and Cherepanov et al. (2013) are consistent with the attraction effect, the two decoy effect violates Weak WARP, an axiom that is necessary for the characterization of these two papers.

[22]More precisely, they only allow $C(\{\mathbf{x}, \mathbf{y}\}) = \{\mathbf{x}, \mathbf{y}\}$ and $C(\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = \{\mathbf{y}\}$. In other words, they do not allow the compromise (attraction) effect such that $C(\{\mathbf{x}, \mathbf{y}\}) = \{\mathbf{x}\}$ and $C(\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}) = \{\mathbf{y}\}$.

Chapter 3

THEORY OF DECISIONS BY INTRA-DIMENSIONAL COMPARISONS

3.1 Introduction

Making a choice between two multidimensional alternatives is a difficult task unless one dominates the other. It is not surprising, therefore, that a decision maker adopts some procedure, or heuristic, to make this choice. One such heuristic called the *Intra-Dimensional Comparison* (IDC) heuristic has been documented in the experimental work of Tversky (1969). The IDC heuristic is a procedure that compares multidimensional alternatives dimension-by-dimension and makes a decision based on those comparisons. We develop an axiomatic model of the IDC heuristic and provide a general framework that is applicable to many different contexts, such as risky choice and social choice.

To illustrate the IDC heuristic, consider the following two scenarios. In the first scenario, a decision maker (henceforth, DM) has to choose a lottery from two binary lotteries. Let us denote by (x, p) the binary lottery that returns \$$x$ with probability p, and the prize \$0 with probability $1-p$. Suppose the DM compares binary lotteries $(\$11, 0.43)$ and $(\$10, 0.45)$. There is no obvious answer to the question of how many dollars are equivalent to a 2% increase in the probability of winning. Instead, it is easier for the DM to compare the binary lotteries dimension-by-dimension (i.e., \$11 with \$10 and 43% with 45%) because the compared numbers represent the same attributes. In this scenario, the DM may prefer $(\$11, 0.43)$ over $(\$10, 0.45)$ because their winning probabilities are not very different and the former has a higher prize.

In the second scenario, a social planner (SP) needs to choose an allocation. Suppose she wants to allocate \$24 to three people and is considering two possible allocations: $(\$12, \$4, \$8)$ and $(\$14, \$3, \$7)$ (person 1 gets either \$12 or \$14, person 2 gets either \$4 or \$3, and so on). Although all dimensions are expressed in dollars, it may be easier for the SP to compare the allocations dimension-by-dimension (i.e., \$12 with \$14, \$4 with \$3, and \$8 with \$7). In this scenario, the SP may choose $(\$12, \$4, \$8)$ over $(\$14, \$3, \$7)$ because although person 1 is worse off, he still gets \$12, and the other two people are better off in the first allocation.

As we have seen in the above two scenarios, the IDC heuristic simplifies and guides choice in individual decision making as well as in social choice. The IDC heuristic was documented in the experimental work of Tversky (1969). Tversky run the following experiment in the context of the first scenario. Subjects were asked to choose one lottery from each of all possible pairs from $a = (5, \frac{7}{24})$, $b = (4.5, \frac{9}{24})$, and $c = (4, \frac{11}{24})$. Tversky (1969) obtained a systematic violation of transitivity: almost half of the subjects preferred the lottery with the higher payoff among adjacent pairs (a and b or b and c), while on the extreme pair (a and c), they preferred the lottery with the higher winning probability.[1]

In the words of one of Tversky's subjects in a post-experimental interview, "There is a small difference between lotteries a and b or b and c, so I would pick the one with higher payoff. However, there is a big difference between lotteries a and c, so I would pick the one with higher probability." In other words, the subject compared prizes and probabilities separately and made a decision based on those comparisons; i.e., used the IDC heuristic.[2]

The goal of this paper is to develop an axiomatic model of the IDC heuristic and discuss its implications. The main representation theorem of the paper shows that under standard axioms, in addition to an axiom called *Separability*, there are dimension-specific functions that represent dimension-by-dimension comparisons and a function that aggregates those comparisons in order to make a decision. Moreover, these functions are unique up to a certain normalization. Separability guarantees that comparisons within one dimension are independent from comparisons within other dimensions.

Now we introduce our model of the IDC heuristic in the aforementioned two scenarios and discuss some implications of them. First, let us describe our model in the context of the first scenario: choice over binary lotteries. Consider a DM who has to choose between lotteries (x, p) and (y, q) where $x > y$ and $q > p$. Our discussion above suggests that she compares x with y and q with p separately. She then makes a decision based on numbers $f(x, y)$ and $g(q, p)$, where $f(x, y)$ measures the advantage of x over y and $g(q, p)$ measures the advantage of q over p. We can

[1] Therefore, we obtain a violation of transitivity $a = (5, \frac{7}{24}) \succ b = (4.5, \frac{9}{24}) \succ c = (4, \frac{11}{24}) \succ a$ which is called the *Similarity Cycle*. More generally, the similarity cycle is a triple $(x, p), (y, q), (z, r)$ of binary lotteries such that $x > y > z$ and $(x, p) \succ (y, q) \succ (z, r) \succ (x, p)$. While Tversky's original experiment involved only a small number of subjects, this result was replicated by Lindman and Lyons (1978) and Budescu and Weiss (1987). Moreover, Day and Loomes (2010) produced a similar result by using different lotteries with real incentives.

[2] By using an eye-tracking experiment in the same setting with Tversky (1969), Arieli et al. (2011) found that when decision-making is difficult, subjects compare prize and probability separately.

imagine that if $f(x,y) > g(q,p)$, then the prize-dimension becomes more salient, and she prefers (x,p) over (y,q) since $x > y$. If by contrast, $f(x,y) < g(q,p)$, then the probability-dimension becomes more salient, and she prefers (y,q) over (x,p) since $q > p$. Finally, if $f(x,y) = g(q,p)$, then she is indifferent between (x,p) and (y,q). More formally, we say that a binary relation $\succeq$ on binary lotteries is an *IDC relation* if there are functions f and g such that for any binary lotteries (x,p) and (y,q),

$$(x,p) \succeq (y,q) \text{ if and only if } f(x,y) \geq g(q,p). \tag{3.1}$$

For example, when $f(x,y) = 1 - \frac{u(y)}{u(x)}$ and $g(q,p) = 1 - \frac{p}{q}$, (3.1) gives us Expected Utility Theory (EUT) preferences on binary lotteries. By allowing a more general form for g, the model can accommodate well-known deviations of EUT: the *Common Ratio Effect* (a version of the Allais Paradoxes) and the *Similarity Cycle* and *Regret Cycle* (both violations of transitivity that have opposite directions). We are not aware of any other axiomatic model that can accommodate all deviations mentioned.[3]

Next, we define our general model in the context of the second scenario, social choice. A social planner (SP) for the society $N = \{1, 2, \ldots, n\}$ compares two allocations $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, where x_i and y_i are the amounts of money that person i receives from the allocations. If the SP uses the IDC heuristic, then she asks each i to say how much he prefers x_i over y_i; in other words, she obtains a real number $f_i(x_i, y_i)$ that measures the advantage of x_i over y_i for person i, and aggregates these real numbers. More formally, we say that a binary relation $\succeq$ on n-vectors is an *IDC relation* if there exist functions $\{f_i\}_{i=1}^n$ and W such that for any $\mathbf{x}$ and $\mathbf{y}$,

$$\mathbf{x} \succeq \mathbf{y} \text{ if and only if } W\big(f_1(x_1,y_1), \ldots, f_n(x_n,y_n)\big) \geq 0. \tag{3.2}$$

Note that (3.2) generalizes (3.1). In contrast to the Bergson-Samuelson social welfare criterion[4] in which the SP aggregates preferences before comparing alternatives, in our model the SP aggregates preferences after she compares alternatives dimension-by-dimension. The IDC relations are (possibly) intransitive. For exam-

[3]The common ratio effect is a violation of the Independence Axiom such that $(x,p) \prec (y,q)$ and $(x,\alpha p) \succ (y,\alpha q)$ with $x > y$ and $1 > \alpha > 0$. See Allais (1953) and Kahneman and Tversky (1979). The Regret Cycle is a violation of transitivity such that $(x,p) \prec (y,q) \prec (z,r) \prec (x,p)$ with $x > y > z$, which is observed in the experiments of Loomes et al. (1991) and Day and Loomes (2010). We can generate the Similarity Cycle, the Regret Cycle, and the Common Ratio Effect with distance-based functions $g(q,p) = (1 - \frac{p}{q})[\frac{1-\beta p(q-p-\mu)}{1-\delta}]$ and $f(x,y) = 1 - \frac{u(y)}{u(x)}$ where $\beta, \mu, \delta \in [0,1)$.

[4] In the Bergson-Samuelson social welfare criterion, the SP compares $W(u_1(x_1), \ldots, u_n(x_n))$ and $W(u_1(y_1), \ldots, u_n(y_n))$ for given functions $u_1, \ldots, u_n$, and W (See Mas-Colell et al. (1995)).

ple, the model allows Condorcet Cycles in which the SP compares three allocations **x**, **y**, and **z**, and **y** is preferred to **x**, **z** to **y**, but **x** to **z**.[5]

The remainder of the paper is organized as follows. First, we discuss related literature in Section 3.1.1. In Section 3.2, we introduce the basic model and a behavioral foundation for (3.2). We also discuss some specifications and properties of f_i in Section 3.2.2 and the proof of the main theorem is in Appendix C.

3.1.1 Related Literature

Representations similar to (3.2) are characterized in Bouyssou and Pirlot (2002). They focus on binary aggregators, so they cannot have a uniqueness result. In other words, their representations have no cardinal meaning. Moreover, the implications of their representations for economic contexts and specifications are not worked out.

Now we discuss two axiomatic works on intransitive preferences related to our work. The first work is the *Similarity Relation Model* (SRM) of Rubinstein (1988), a model of choice under risk that used an idea of the IDC heuristic.

SRM consists of two similarity relations on prizes and probabilities, denoted by $\sim_x$ and $\sim_p$ respectively. Suppose a DM compares binary lotteries (x, p) and (y, q) where $x > y$ and $p < q$. In SRM, the DM uses the following procedure: if she considers p and q to be similar ($p \sim_p q$) and x and y not to be similar ($x \nsim_x y$), then she prefers (x, p) over (y, q); if instead she considers x and y to be similar ($x \sim_x y$) and p and q not similar ($p \nsim_p q$), then she prefers (y, q) over (x, p); finally, if she considers both x and y and p and q to be similar ($x \sim_x y$ and $p \sim_p q$) or not similar ($x \nsim_x y$ and $p \nsim_p q$), then the SRM procedure is not specified. As in the IDC heuristic, SRM compares prizes and probabilities separately and makes a decision based on these comparisons if possible. The main difference between our model and SRM is that the similarity relations $\sim_x$ and $\sim_p$ are exogenously given in SRM and are not unique, i.e., there could be multiple similarity relations that are compatible with one binary relation. Moreover, the procedure that is generated by the similarity relations is not complete.[6] On the other hand, we fully characterize the set of IDC relations (which are complete), f and g are endogenously derived from each IDC relation, and they are unique.

[5] For example, suppose the SP wants to allocate \$60 to three people and considers three possible allocations: $\mathbf{x} = (\$24, \$20, \$16)$, $\mathbf{y} = (\$16, \$24, \$20)$, and $\mathbf{z} = (\$20, \$16, \$24)$. Let $f_i(x_i, y_i) = -f_i(y_i, x_i) = \frac{x_i - y_i}{x_i}$ for any i and x_i, y_i with $x_i \geq y_i$. Suppose $W(f_1, f_2, f_3) = f_1 + f_2 + f_3$. Then we obtain a Condorcet Cycle because $\frac{24-20}{24} + \frac{20-16}{20} - \frac{24-16}{24} = \frac{1}{30} > 0$.

[6] There are cases in which the SRM procedure is almost never specified.

The second work is the Relative Discounting Model of Ok and Masatlioglu (2007) in intertemporal choice. They model a DM who compares intertemporal prospects (x, t) where (x, t) denotes an intertemporal prospect that gives \$$x$ at time t. Although their model does not build on the IDC heuristic, when there are only two dimensions, our model coincides with the Relative Discounting Model. More precisely, they have a representation similar to (3.1) in which the time-dimension corresponds to the probability-dimension in (3.1).

3.2 The Basic Model

Let $X \equiv \prod_{i=1}^{n} X_i \subseteq \mathbb{R}^n$ be a set of alternatives with n attributes, where $X_i = (a_i, b_i]$ for some $a_i, b_i \in \mathbb{R}$ with $a_i < b_i$ for each i.[7] Let $\succeq$ be a binary relation on X. Let $\succ$ (resp. $\sim$) denote the asymmetric (resp. symmetric) part of $\succeq$. For any $\mathbf{x}, \mathbf{y} \in X$, we say that $\mathbf{x}$ *dominates* $\mathbf{y}$, denoted by $\mathbf{x} > \mathbf{y}$, if $x_i \geq y_i$ for each i and $\mathbf{x} \neq \mathbf{y}$. For any vector $\mathbf{t} \in \mathbb{R}^n$, we denote by $\mathbf{t}_{-i}$, the $(n-1)$-dimensional vector that remains after t_i is excluded from $\mathbf{t}$. Similarly, we can define $\mathbf{t}_{-i,-j}$.

Now we define the model and the IDC relations. We say a function $f : \mathbb{R}^2 \rightarrow (-1, 1)$ is a *distance-based function* if it is strictly increasing in its first argument and $f(x, y) = -f(y, x)$ for all $x, y \in \mathbb{R}$. Note that $f(x, x) = 0$ and $f(x, y) > 0$ if $x > y$. Roughly speaking, $f(x, y)$ is a relative advantage of x over y. For example, $f(x, y)$ can be a function of the utility differences between x and y, i.e., $f(x, y) = h(u(x) - u(y))$ for some functions h and u. We say a function $W : (-1, 1)^n \rightarrow \mathbb{R}$ is an *aggregator* if it is strictly increasing in all its arguments and for any i and $t_i \in (-1, 1)$, $W(t_i, \mathbf{0}_{-i}) = t_i$. For example, $W(\mathbf{t}) = \sum_{i=1}^{n} t_i + \alpha \cdot \prod_{i=1}^{n} t_i$ is an aggregator for all $\alpha \in [-1, 1]$.

Definition 10 *(IDC Relation) A binary relation $\succeq$ on X is an* IDC relation *if there exist distance-based functions $\{f_i\}_{i=1}^{n}$ and an aggregator W such that for any $\boldsymbol{x}, \boldsymbol{y} \in X$,*

$$\boldsymbol{x} \succeq \boldsymbol{y} \text{ if and only if } W\big(f_1(x_1, y_1), \ldots, f_n(x_n, y_n)\big) \geq 0.$$

The functions $f_1, \ldots, f_n$ represent dimension-by-dimension comparisons as in the IDC heuristic. We call f_i the i^{th}-dimensional distance-based function. Distance-based functions have three natural interpretations. First, f_i measures similarity (or

[7] We focus on left-open intervals because comparisons between alternatives on the left boundaries (e.g., $(a_i, \mathbf{x}_{-i})$ and $(a_j, \mathbf{y}_{-j})$) may violate the representation. For example, let us consider a binary relation $\succeq$ on $[0, \overline{x}] \times [0, 1]$ the set of binary lotteries and the representation (3.1). Then a comparison between zero lotteries $(0, p)$ and $(0, q)$ (with $p \neq q$) is not consistent with our representation because $f(0, 0) \neq g(q, p)$ and $(0, p) \sim (0, q)$.

dissimilarity) of x_i and y_i. For example, in the first scenario, if prizes x and y are similar, then $|f(x, y)|$ is small, and if x and y are distinct, then $|f(x, y)|$ is large. But f also measures a degree of (dis)similarity of prizes. The similarity relations of Rubinstein (1988) have close features and in fact, IDC relations can be seen as smooth and complete versions of the procedure he analyzed.[8] Second, f_i measures how salient the i^{th}-dimension is compared to other dimensions. For example, in the first scenario, $f > g$ means that the prize-dimension is more salient than the probability-dimension. Finally, f_i can be seen as a function of utility difference. We will discuss this interpretation in Section 3.2.2.

3.2.1 Separability and Representation Theorem

In this subsection, we characterize our model. In particular, we can characterize (3.2) by an axiom called *Separability* in addition to standard postulates. We begin by imposing four standard properties: completeness, continuity, strong monotonicity, and richness. We concentrate all of these properties in one axiom, called *Regularity*.

Axiom 13 *(Regularity) Let $\succeq$ be a binary relation on X.*

1. *(Completeness) For any $\boldsymbol{x}, \boldsymbol{y} \in X$, either $\boldsymbol{x} \succeq \boldsymbol{y}$ or $\boldsymbol{y} \succeq \boldsymbol{x}$;*
2. *(Continuity) For any $\boldsymbol{x} \in X$, $\{\boldsymbol{y} \in X | \boldsymbol{y} \succeq \boldsymbol{x}\}$ and $\{\boldsymbol{y} \in X | \boldsymbol{x} \succeq \boldsymbol{y}\} \cup X_0$ are closed sets where $X_0 \equiv \{t \in \prod_{i=1}^{n}[a_i, b_i] | \exists i \text{ such that } t_i = a_i\}$;*
3. *(Strong Monotonicity) For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$, if $\boldsymbol{x} \sim \boldsymbol{y}$ and $\boldsymbol{y} \succ \boldsymbol{z}$, then $\boldsymbol{x} \succ \boldsymbol{z}$;*
4. *(Richness) For any $\boldsymbol{x} \in X$, i, and $\boldsymbol{y}_{-i} \in X_{-i}$, there exists $y_i \in X_i$ such that $\boldsymbol{x} \succ \boldsymbol{y} = (y_i, \boldsymbol{y}_{-i})$.[9]*

Completeness states that any two alternatives are comparable. Continuity requires, loosely speaking, that an upper contour set and a lower contour set of any

[8] Note that any IDC relation $\succeq$ with distance-based functions f and g is consistent with the similarity relation model of Rubinstein (1988) (SRM) with the following similarity relations $\sim_x$ and $\sim_p$: $x \sim_x y$ iff $|f(x, y)| \leq \lambda$ and $p \sim_p q$ iff $|g(p, q)| \leq \lambda$ for some fixed $\lambda \in (0, 1)$. Here, we say that a binary relation $\succeq$ is consistent with SRM if there are similarity relations $\sim_x$ and $\sim_p$ such that for any $x > y$ and $p < q$, if $p \sim_p q$ and $x \nsim_x y$, then $(x, p) \succ (y, q)$, and if $x \sim_x y$ and $p \nsim_p q$, then $(y, q) \succ (x, p)$.

[9] In the context of choice over binary lotteries, if one wants to use the domain $[0, \overline{x}] \times [0, 1]$ instead of $(0, \overline{x}] \times (0, 1]$, then Richness is equivalent to the following simpler condition: $(x, p) \succ (y, 0)$ and $(x, p) \succ (0, q)$ for any $(x, p), (y, q) \gg (0, 0)$. It requires that any zero lottery is worse than non zero lotteries. For more details, see the discussion in Section 3.3.

alternative are closed. Technically, a lower contour set is closed after we take into account that each $X_i = (a_i, b_i]$ is left-open. Hence the presence of X_0 in the axiom. Richness states that for any vector $\mathbf{y}_{-i}$, there exists y_i such that $\mathbf{y} = (y_i, \mathbf{y}_{-i})$ is worse than a fixed lottery $\mathbf{x}$. Strong monotonicity implies the standard monotonicity axiom: $\mathbf{x} > \mathbf{y}$ implies $\mathbf{x} \succ \mathbf{y}$. Strong monotonicity also requires transitivity where one of three alternatives dominates one of the other two alternatives ($\mathbf{y} > \mathbf{z}$). We do not impose transitivity beyond this minimal case. Therefore, we allow binary relations that violate transitivity.

The key axiom in this paper is called *Separability* and is closely related to the IDC heuristic.[10] In the IDC heuristic, the DM compares alternatives dimension-by-dimension and Separability requires that these dimension-by-dimension comparisons are independent. For example, in the context of choice over binary lotteries, Separability requires that a comparison between prizes x and y is independent of a comparison between probabilities q and p. Figure 3.1 illustrates the intuition behind Separability for $n = 2$. In Figure 3.1, solid and dashed curves represent indifference curves. Consider two indifferent pairs of binary lotteries (x, p) and (y, q) and (x', p) and (y', q); i.e., $(x, p) \sim (y, q)$ and $(x', p) \sim (y', q)$ (two solid indifference curves). Since these lotteries use the same probabilities, the indifferences suggest that the relative advantage of x with respect to y is equal to that of x' with respect to y'. If this is the case, then the same should hold even if we change the probabilities: we should have $(x, p') \sim (y, q')$ if and only if $(x', p') \sim (y', q')$ for any probabilities p' and q' (two dashed indifference curves). This is the content of Separability on choice over binary lotteries. It is a simple exercise to show that any EUT preference on binary lotteries satisfies Separability.

Axiom 14 *(Separability) For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x'}, \boldsymbol{y'} \in X$ and i, if*

$$(x_i, \boldsymbol{x}_{-i}) \sim (y_i, \boldsymbol{y}_{-i}), \quad (x'_i, \boldsymbol{x}_{-i}) \sim (y'_i, \boldsymbol{y}_{-i}), \text{ and}$$

$$(x_i, \boldsymbol{x'}_{-i}) \sim (y_i, \boldsymbol{y'}_{-i}), \text{ then } (x'_i, \boldsymbol{x'}_{-i}) \sim (y'_i, \boldsymbol{y'}_{-i}).$$

[10]This axiom is used in the theories of additive utility representation. For example, it is called the *corresponding tradeoff condition* in Keeney and Raiffa (1976) and *triple cancellation* in Wakker (1988). When there are only two dimensions and transitivity is satisfied, Separability is implied by *Cancellation* (sometimes it is called Double Cancellation or Thomsen condition). There are a large literature on separability (See Blackorby et al. (2008)). However, they focus on horizontal separability in line with additive utility models in which transitivity is assumed while we use vertical separability in line with IDC heuristic.

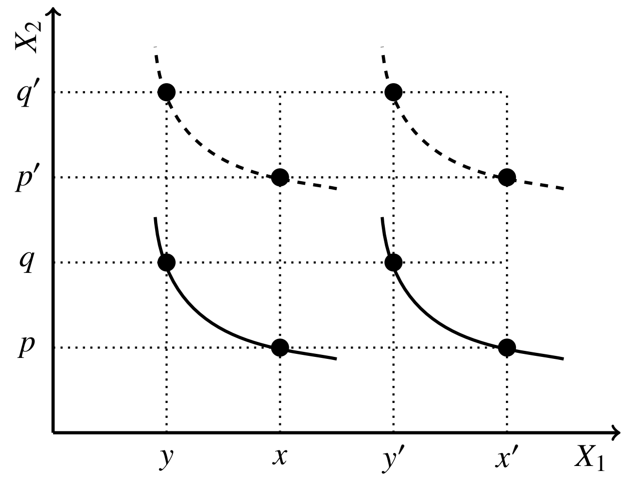

Figure 3.1: Separability when $n = 2$

Separability states that for each i, a comparison between x_i and y_i is independent from the other dimensions (Lemma 7). In the IDC heuristic, the DM compares alternatives dimension-by-dimension and Separability guarantees that these comparisons can be represented by well-defined functions. Now, we are ready to state the main representation theorem.

Theorem 5 *A binary relation $\succeq$ on X satisfies Regularity and Separability if and only if it is an IDC relation with some continuous aggregator W and continuous distance-based functions $\{f_i\}_{i=1}^n$ such that for any $\boldsymbol{t} \in (-1, 1)^n$, $W(\boldsymbol{t}) = W(\boldsymbol{t}_{-n}, 0_n) + t_n$ and $f_n(b_n, x_n) = \frac{b_n - x_n}{b_n - a_n}$ for any $x_n \in X_n = (a_n, b_n]$. Moreover, $\{f_i\}_{i=1}^n$ and W are unique.*

Separability captures the idea of the IDC heuristic and guarantees dimension-specific distance-based functions $f_1, \ldots, f_n$ and an aggregator W. Regularity guarantees that these functions are well-behaved. Note that when $n = 2$, Theorem 5 also characterizes (3.1) since $W(t_1, t_2) = t_1 + t_2$.

We can derive unique distance-based functions and an aggregator function up to a certain normalization. It is obvious that we cannot get any uniqueness without normalizing distance-based functions and an aggregator. For example, any common monotonic transformation of f and g, $h(f)$ and $h(g)$, satisfies (3.1). Therefore, we normalize one of the distance-based functions, e.g., f_n. However, it is not necessary to specify f_n on all points of X_n^2. When $n = 2$, it is enough to normalize f_n as

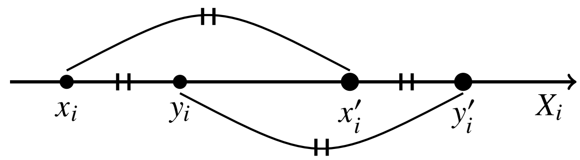

Figure 3.2: Difference-i

$f_n(b_n, x_n) = \frac{b_n - x_n}{b_n - a_n}$ for any $x_n \in X_n$. When $n \geq 3$, we also need a normalization on W such that $W(\mathbf{t}) = W(\mathbf{t}_{-n}, 0_n) + t_n$ for each $\mathbf{t} \in (-1, 1)^n$. We discuss some specifications and properties of f_i in the next subsection.

To gain an intuition of the proof of Theorem 5, we construct f and g for (3.1) on choice over binary lotteries. Let $X_1 = (0, \overline{x}]$ and $X_2 = (0, 1]$. For any $x, y \in X_1$ with $x \geq y$, set $f(x, y) = 1 - p$ whenever $(x, p) \sim (y, 1)$. By Regularity, f is a well-defined function (Lemma 6). For any $p, q \in X_2$ with $q \geq p$, set $g(q, p) = f(x, y)$ whenever $(x, p) \sim (y, q)$. Obviously, we can find two different pairs x and y and x' and y' such that $(x, p) \sim (y, q)$ and $(x', p) \sim (y', q)$. Therefore, we need to show that $f(x, y) = f(x', y')$. Take any two pairs x and y and x' and y' such that $(x, p) \sim (y, q)$ and $(x', p) \sim (y', q)$. Let $r \in X_2$ such that $(x, r) \sim (y, 1)$. By Separability, we have $(x', r) \sim (y', 1)$. Then, by the construction of f, we obtain $f(x, y) = f(x', y') = 1 - r$ which implies that g is well-defined.

3.2.2 Some Specifications and Properties of Distance-Based Functions f_i

Finally, we discuss some specifications and properties of f_i. A natural specification of f_i is $f_i(x_i, y_i) = h_i(u_i(x_i) - u_i(y_i))$ for some functions h_i and u_i. In other words, f_i is a function of utility difference. In fact, many models (e.g., additive utility models, EUT, the Additive Difference Model of Tversky (1969), and the Regret Theory of Loomes and Sugden (1982)) follow this specification. It can be characterized by the following axiom: for given i, we define an axiom called *Difference-i*.

Axiom 15 (Difference-i) *For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}', \boldsymbol{y}' \in X$, if*

$$\begin{aligned} (x_i, \boldsymbol{x}_{-i}) &\sim (y_i, \boldsymbol{y}_{-i}) \text{ and } (x'_i, \boldsymbol{x}_{-i}) \sim (y'_i, \boldsymbol{y}_{-i}), \text{ then} \\ (x_i, \boldsymbol{x}'_{-i}) &\sim (x'_i, \boldsymbol{y}'_{-i}) \text{ iff } (y_i, \boldsymbol{x}'_{-i}) \sim (y'_i, \boldsymbol{y}'_{-i}). \end{aligned}$$

Figure 3.2 illustrates Difference-i. It requires that if the relative advantage of x_i over y_i is equal to that of x'_i over y'_i (represented by the segments $\overline{x_i y_i}$ and $\overline{x'_i y'_i}$), then the relative advantage of x_i over x_i' is equal to that of y_i over y_i' (represented by the arcs $\overline{x_i x_i'}$ and $\overline{y_i y_i'}$). Now we state the characterization for f_i.

Proposition 7 *For given i, if an IDC relation $\succeq$ with distance-based functions $\{f_j\}_{j=1}^n$ is continuous and satisfies Difference-i, then there are continuous strictly increasing functions h_i and u_i such that $f_i(x_i, y_i) = h_i(u_i(x_i) - u_i(y_i))$.*

Since this characterization directly follows from the result of Suppes and Winet (1955), we omit the proof. Next, we discuss a property on distance-based functions, called *Diminishing Sensitivity*. Formally,

Diminishing Sensitivity: A distance-based function f satisfies *Diminishing Sensitivity* if $f(x+\epsilon, y+\epsilon) < f(x, y)$ for any $x > y$ and $\epsilon > 0$.

Diminishing sensitivity requires that the relative advantage of x over y decreases as the average $\frac{x+y}{2}$ goes up. In other words, the DM is less sensitive when x and y are large. The idea of diminishing sensitivity is commonly used in the literature (e.g., Tversky and Kahneman (1991) and Bordalo et al. (2012)). In fact, the same condition is called diminishing sensitivity in Bordalo et al. (2012). It is also related to the Weber-Fechner law of human perception, which states that the perceived intensity is proportional to the logarithm of the stimulus.[11] We think diminishing sensitivity is likely to be satisfied in two contexts we discussed in the introduction. Diminishing sensitivity is also natural when f_i measures (dis)similarity or saliency. Moreover, when f is a function of utility difference; i.e., $f(x, y) = h(u(x) - u(y))$, diminishing sensitivity is equivalent to the strict concavity of u. In particular, when $f(x, y) = 1 - \exp(u(y) - u(x)) = 1 - \frac{\exp(u(y))}{\exp(u(x))} = 1 - \frac{\tilde{u}(y)}{\tilde{u}(x)}$, diminishing sensitivity is equivalent to the strict log-concavity of $\tilde{u}$.

In general, characterizing diminishing sensitivity is difficult. However, in our model, the following simple axiom characterizes diminishing sensitivity.

Axiom 16 (Diminishing Sensitivity) *For any $\boldsymbol{x}, \boldsymbol{y} \in X$ and i, if $(x_i, \boldsymbol{x}_{-i}) \sim (y_i, \boldsymbol{y}_{-i})$ and $x_i > y_i$, then $(x_i + \epsilon, \boldsymbol{x}_{-i}) \prec (y_i + \epsilon, \boldsymbol{y}_{-i})$ for any $\epsilon > 0$ such that $x_i + \epsilon \in X_i$.*

[11] To illustrate, suppose x and y are the magnitudes of some stimulus with $x > y$. Then the difference between the perceived intensities of x and y is proportional to $\log(\frac{x}{y})$. Now note that the perceived intensity difference decreases as the average $\frac{x+y}{2}$ goes up since $\log(\frac{x}{y}) > \log(\frac{x+\epsilon}{y+\epsilon})$ for any $\epsilon > 0$.

Further specifications may depend on the context of decision-making. So we will not discuss further specifications of f_i.[12]

3.3 Discussion on Changing the Domain

Here let us briefly discuss how to change left-open intervals $(a_i, b_i]$ to closed intervals $[a_i, b_i]$ and obtain the same representation result. We discuss it in the context of choice over binary lotteries. Suppose we want to obtain (3.1) for the domain $[0, \overline{x}] \times [0, 1]$. In this case, Richness can replaced by the following simpler condition.

Richness*: $(x, p) \succ (y, 0)$ and $(x, p) \succ (0, q)$ for any $(x, p), (y, q) \gg (0, 0)$.

Under Richness*, we can obtain Theorem 5 for the domain $[0, \overline{x}] \times [0, 1]$ with two exceptions. First of all, the main technical difficulty of using the domain $[0, \overline{x}] \times [0, 1]$ is that zero lotteries are indifferent with each other; that is, $(0, p) \sim (0, q) \sim (x, 0) \sim (y, 0)$ for any $(x, p), (y, q)$. This violates our representation (3.1) because $f(0, 0) = 0 < g(q, p)$ and $f(x, y) > g(0, 0) = 0$ when $x > y$ and $q > p$. Therefore, we cannot compare zero lotteries in our representation. However, comparisons between zero lotteries and non zero lotteries are consistent with (3.1). Second, because of Richness*, we need to have $f(x, 0) = 1$ and $g(p, 0) = 1$ when $x, p > 0$. Therefore, f and g are continuous and strictly increasing only at $(0, \overline{x}]^2$ and $(0, 1]^2$, respectively.

[12] A natural specification of W is that $W(\mathbf{t}) = \sum_{t=1}^{n} t_i$. This specification generalizes the theories of additive utilities (Debreu (1960b) and Wakker (1988)), the Additive Difference Model of Tversky (1969), the Regret Theory of Loomes and Sugden (1982), and a version of the Salience Theory of Bordalo et al. (2012).

Chapter 4

THE PERCEPTION-ADJUSTED LUCE MODEL

4.1 Introduction

We study the role of perception in individual stochastic choice. Perception is captured through *priority orders,* which determine whether an alternative, or object of choice, is perceived sooner or later than other alternatives. The perception priority order could represent differences in familiarity, or salience, of the objects of choice.

Our main contribution is to identify a perception priority order from an agent's violations of *independence from irrelevant alternatives* (IIA), the rationality axiom behind Luce's (1959) model of choice. We attribute any violation of Luce's model to the role of perception, and use these violations to back out a perception order. Our model, a *perception-adjusted Luce model (PALM)*, reduces to Luce's when perception plays no role, and uses perception to capture violations of Luce's model.

In PALM, an agent makes choices as if she were following a sequential procedure. In the procedure, the agent considers different alternatives in sequence, following a perception priority order. The probability of choosing an alternative depends on the probability that no alternative that was considered, or perceived, before was chosen. The choice probability also depends on relative utility, just as in Luce's model. The sequential nature of PALM allows us to explain well-known behavioral phenomena, such as the attraction and compromise effects, and the consequences of forced choice and choice overload (see Sections 4.4.1 and 4.5).

We use stochastic choice data to construct a perception priority order. We start from a primitive stochastic choice, and when the choice satisfies certain axioms, we can construct a PALM model. The perception priority order comes from the observed violations of Luce's IIA. Luce's IIA says that the relative choice probabilities of alternative a over b should not be affected by adding a third alternative c. So suppose that we have a violation of IIA, and that adding c changes the probability of choosing a relative to that of choosing b. What can we conclude about perception? We claim that a *decrease* in the relative probability of choosing a over b is an indication that a has *higher* perception priority than b.

The reasoning is as follows. Adding c should in principle decrease the probability of choosing a and the probability of choosing b. The reason is that c competes with a and b. But when a has a higher perception priority than b, then the very fact that a is chosen with lower probability means that b has a higher chance of being perceived. So there is a second effect of adding c, and it favors choosing b. The final consequence is that adding c provokes a larger decrease in the probability of choosing a than in the probability of choosing b. This means that the resulting violation of Luce's IIA takes the form of a decrease in the relative probability of choosing a over b.

The second idea in our construction is to use the perception priority order to define a *hazard rate*. The hazard rate is the probability of choosing an object, conditional on not choosing any of the objects with higher perception priority. So hazard rates incorporate the effects of perception. We impose two axioms. The first requires that the perception priority be complete and transitive. The second axiom is imposed on hazard rates, and says that hazard rates must satisfy the IIA. Since hazard rates are obtained from choices by accounting for priority, and hazard rates equal the primitive choice probability where priority does not matter, our axiom means that perception explains all the deviations from IIA.

The resulting model of choice is what we call PALM, the perception-adjusted Luce model. In PALM, an agent who is faced with a choice problem considers the different alternatives in order of their priority. Each time one alternative is considered, it is chosen with probability dictated by an underlying Luce model. So the probability that a given alternative is chosen depends both on its utility (as in Luce) and on its priority in perception.

Despite having a tight axiomatic characterization, PALM is quite flexible and can accomodate many behavioral phenomena, including some of the best known violations of Luce's model in experiments. It can explain the attraction and compromise effects: Section 4.5 has the details. Some of these effects stem from violations of the regularity axiom; PALM can violate the regularity axiom. In Section 4.5, we also use PALM to explain recent experimental findings on how forcing agents to make a choice affects their choices. Another application of PALM is to choice overload (see Section 4.4). An increase in the number of objects can lead to an increased probability of not making a choice, when the objects are similar to each other.

It is instructive to see how PALM can accommodate the attraction effect. Doyle et al. (1999) is a representative experiment with evidence in favor of the attraction effect: Doyle et al. present customers with a choice of baked beans. The first choice is between two types of baked beans: a and b; a is Heinz baked beans, while b is a local cheap brand called Spar. In the experiment, b was chosen 19% of the time. The authors then introduced a third option, c, a more expensive version of the local brand Spar. After c was introduced, b was chosen 33% of the time. This pattern of choices cannot be explained by Luce's model; indeed it cannot be explained by any model of random utility. It can, however, be explained by PALM.

Suppose that perception is related to the familiarity of the brand of beans. Since a is the well-known Heinz brand, it is likely to be the highest priority alternative. Also, b is at least as familiar as c because b and c are the same brands, and c is introduced later. Given this perception priority, if the utility of a is large enough, PALM produces the attraction effect in Doyle et al.'s experiment. As we explained above, the addition of c in principle hurts the choice probabilities of both a and b, but, while a does not benefit from b's potential decrease, b does benefit from the decrease in the probability of choosing a because b has lower priority than a. The magnitude of this positive effect depends on the utility of a; if the utility of a is large enough, then the indirect positive effect overcomes the direct negative effect, and that is how PALM produces an increase in the probability of choosing a. This increase in the probability of choosing a is the attraction effect.

There are models within the axiomatic literature that explain the attraction and compromise effects. Section 4.7 discusses the related literature. Most of proofs are collected in Appendix D.

4.2 Primitives and Luce's model

Let X be a countable and nonempty set of *alternatives*, and $\mathscr{A}$ be a set of finite and nonempty subsets of X. Suppose that $\mathscr{A}$ includes all sets with two and three elements. We model an agent who makes a probabilistic choice from $A_0 \equiv A \cup \{x_0\}$, with $A \in \mathscr{A}$. The element $x_0 \notin X$ represents an outside option that is always available to the agent. Choosing the outside option can be interpreted as the agent not making a choice. Let $X_0 \equiv X \cup \{x_0\}$.

Definition 11 *A function* $\rho : X_0 \times \mathscr{A} \to [0, 1]$ *is called* a stochastic choice function *if*

$$\sum_{a \in A_0} \rho(a, A) = 1$$

for all $A \in \mathscr{A}$. A stochastic choice function ρ is nondegenerate *if $\rho(a, A) \in (0, 1)$ for all $A \in \mathscr{A}$ with $|A| \geq 2$ and $a \in A$.*

We write $\rho(B, A)$ for $\sum_{b \in B} \rho(b, A)$, and say that $\rho(\emptyset, A) = 0$. Note that we allow for $\rho(x_0, A) = 0$. So it is possible that the outside option is never chosen with positive probability, even when ρ is nondegenerate.

Definition 12 (IIA) *A stochastic choice function ρ satisfies* Luce's independence of irrelevant alternatives (IIA) axiom at *$a, b \in X_0$ if, for any $A \in \mathscr{A}$,*

$$\frac{\rho(a, \{a, b\})}{\rho(b, \{a, b\})} = \frac{\rho(a, A)}{\rho(b, A)}.$$

Moreover, ρ satisfies IIA *if ρ satisfies IIA at a, b for all $a, b \in X$.*

Luce (1959a) proves that, if a non-degenerate stochastic choice function satisfies IIA, then it can be represented by the following model (also referred to as multinomial logit):

Definition 13 (Luce's Model) *ρ satisfies the (extended) Luce's model if there exists a real-valued function u on X_0 such that*

$$\rho(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a') + u(x_0)}. \tag{4.1}$$

Luce presented his model with no outside option. Here we allow for an outside option, and use the version of Luce's model in which not choosing in A is possible. Note that $\rho(x_0, A) = 0$ iff $u(x_0) = 0$.

Luce's model satisfies a monotonicity property: $\rho(x, A) \geq \rho(x, B)$, if $A \subset B$. This property is called *regularity*.

4.2.1 PALM

Perception priority. We capture the role of perception through a weak order $\succsim$. The idea is that when $a \succ b$, then a tends to be perceived sooner than b, and when $a \sim b$, then a and b are perceived simultaneously.

A PALM decision maker is described by two parameters: a weak order $\succsim$ and a utility function u. She perceives each element of a set A sequentially according to the perception priority $\succsim$. Each perceived alternative is chosen with probability described by μ, a function that depends on utility u according to Luce's formula (4.1). Formally, the representation is as follows.

Definition 14 (PALM) *A perception-adjusted Luce model (PALM) is a pair* $(u, \succsim)$ *of a weak order* $\succsim$ *on* X*, and a function* $u : X_0 \to \mathbf{R}$ *such that*

$$\rho(a, A) = \mu(a, A) \prod_{\alpha \in A/\succsim:\, \alpha \succ a} (1 - \mu(\alpha, A)), \tag{4.2}$$

where

$$\mu(a, A) = \frac{u(a)}{\sum_{b \in A} u(b) + u(x_0)}.$$

The notation $A/\succsim$ is standard: $A/\succsim$ is the set of equivalence classes in which $\succsim$ partitions A. That is, (i) if $A/\succsim = \{\alpha_i\}_{i \in I}$, then $\cup_{i \in I} \alpha_i = A$; and (ii) $x \sim y$ if and only if $x, y \in \alpha_i$ for some $i \in I$. The notation $\alpha \succ a$ means that $x \succ a$ for all $x \in \alpha$. Luce's model is a special case of PALM, in which $a \sim b$ for all $a, b \in X$.

For any PALM $(u, \succsim)$, we denote by $\rho_{(u,\succsim)}$ the stochastic choice defined through (4.2). (When there is no risk of confusion, we write ρ instead of $\rho_{(u,\succsim)}$.)

The PALM has a procedural interpretation. Consider the following procedure. First consider the highest-priority alternatives in A, and choose each of them with probability given by $\mu(\cdot, A)$; these probabilities obey a Luce formula. This means that if α is the set of highest-priority elements of A, then each $a \in \alpha$ is chosen with probability $\mu(a, A)$. With probability $1 - \mu(\alpha, A)$ none of the elements in α is chosen. If none of the elements of α are chosen, then move on to the second-highest priority alternatives, and choose each of them with the Luce probability specified by μ. And so on and so forth.

For example, consider the menu $A = \{x, y, z\}$ with $x \succ y \succ z$. In the PALM, the agent first looks at x and chooses x with "Luce probability" $\mu(x, A)$. With probability $1 - \mu(x, A)$, x is not chosen, and the agents moves on to consider y, the second-highest priority element. She chooses y with probability $\mu(y, A)$. This means that the probability of choosing y is $\mu(y, A)(1 - \mu(x, A))$. Finally, the probability of choosing z is equal to $\mu(z, A)(1 - \mu(x, A))(1 - \mu(y, A))$. If, instead of having $x \succ y \succ z$, we have that $x \sim y \succ z$ then the probability of choosing z is equal to $\mu(z, A)(1 - \mu(x, A) - \mu(y, A))$. The idea captured by $x \sim y$ is that x and y are perceived, and considered, simultaneously. So the probability of choosing an option that has higher priority than z is $\mu(x, A) + \mu(y, A)$.

4.3 Axioms

We introduce the revealed perception priority order derived from ρ, and the resulting hazard rate function. The hazard rate function will be a "perception adjusted" random choice function. It coincides with the random choice function except where violations of Luce's IIA are present. When there are violations of Luce's IIA, they will be attributed to the role of perception. So in our model the hazard rate will satisfy IIA, even when the primitive stochastic choice violates IIA.

<u>Revealed perception priority.</u> We denote by $\succsim^*$ the revealed priority relation that we obtain from the data in ρ. To define $\succsim^*$, first we identify the direct revealed priority relation $\succsim^0$ from ρ. The revealed priority relation $\succsim^*$ is defined as the transitive closure of $\succsim^0$.

We shall attribute all violations of IIA to the role of perception. That is, we require that $a \sim^0 b$ when IIA holds at a and b. In other words, when two alternatives a and b do not exhibit a violation of IIA then we impose that they are equivalent from the view point of perception: they have the same perception priority.

In contrast, if a and b are such that IIA fails at a and b, meaning that there is some third alternative whose presence affects the relative probability of choosing a over b, then we shall require that a and b are strictly ordered by $\succ^0$. We shall require that either $a \succ^0 b$ or that $b \succ^0 a$. Which of the two orderings, $a \succ^0 b$ or $b \succ^0 a$, is determined by the nature of the violation of IIA.

Suppose that IIA fails at a and b because there is some c such that

$$\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}. \tag{4.3}$$

In words, the presence of c lowers the probability of choosing a relative to the probability of choosing b. When does adding an option hurt one alternative relatively more than another? We claim that this happens when a has *higher priority* than b. The reason is that by adding c we are "muddying the waters." We are making the choice between a and b less clear than before, and thus diluting the advantage held by the high priority a over the low priority b.

As we explained in the introduction, we seek to model perception through an order in which alternatives are considered. Adding c to $\{a, b\}$ would in principle decrease the probability of choosing both a and b because c competes with a and b; but when a has higher priority than b, then the sole fact that a's choice probability decreases implies that choosing b becomes more likely. The reason is that b is only chosen when a is not chosen, so the decrease in the probability of choosing a increases the probability of choosing b. Of course, by adding c we may also be decreasing the probability of choosing b because c and b are in competition, so the *net* effect on the probability of choosing b is not determined. However, we do know that $\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}$. And thus the direction of violation of Luce's IIA is dictated by perception priority.

Definition 15 *Let a and b be arbitrary elements in X.*
(i)

$$a \sim^0 b \quad \text{if} \quad \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})},$$

for all $c \in X$;
(ii)

$$a \succ^0 b \quad \text{if} \quad \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})},$$

for all $c \in X$ such that $c \nsim^0 a$ and $c \nsim^0 b$, and if there is at least one such c. We write $a \succsim^0 b$ if $a \sim^0 b$ or $a \succ^0 b$.
(iii) Define $\succsim^*$ as the transitive closure *of $\succsim^0$: that is, $a \succsim^* b$ if there exist $c_1, \ldots, c_k \in X$ such that*

$$a \succsim^0 c_1 \succsim^0 \cdots c_k \succsim^0 b.$$

The binary relation $\succsim^*$ is called the revealed perception priority *derived from ρ.*

It is important to note that

$$\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}$$

does not always imply that $a \succ b$. It will imply that $a \succ b$ only when c has either more or less priority than both a and b. When c is inbetween, then its presence may also disproportionally hurt b, as it has higher priority than b.[1]

We shall impose the following condition on ρ:

[1] To illustrate, consider the case $a \succeq c \succ b$. As we explained before, adding c to $\{a,b\}$ has negative effects on the choice probabilities of both a and b because c competes with a and b. It also has a positive effect on b because b will be chosen only after a is not chosen, and a is not chosen with higher probability after we add c. However, when $a \succeq c \succ b$, then c also directly hurts b (but not a) because b will be chosen only after c is not chosen. Therefore, when $a \succeq c \succ b$, we *can* have

$$\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} < \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}.$$

Axiom 17 (Weak Order) *The relation $\succsim^*$ derived from ρ is a weak order.*

<u>Hazard rate.</u> The second important component of our analysis is the hazard rate function. The hazard rate is the probability of choosing an object, conditional on not choosing any of the objects with higher perception priority.

Definition 16 (Hazard Rate) *For all $a \in X$ and $A \in \mathscr{A}$, define*

$$q(a, A) = \frac{\rho(a, A)}{1 - \rho(A_a, A)},$$

where $A_a = \{b \in A | b \succ^* a\}$, $A \in \mathscr{A}$ and $a \in A$. For the outside option, we also define $q(x_0, A) = 1 - \sum_{a \in A} q(a, A)$. Here q is called ρ's hazard rate function.

We ascribe all violations of IIA to the role of perception, and the hazard rate is the tool that we use to that purpose.

Axiom 18 (Hazard Rate IIA) *The hazard rate function q satisfies Luce's IIA; that is, for any $a, b \in X_0$, and $A \in \mathscr{A}$,*

$$\frac{q(a, \{a, b\})}{q(b, \{a, b\})} = \frac{q(a, A)}{q(b, A)}.$$

The idea behind Hazard Rate IIA is that all violations of Luce's IIA are explained by the perception priority order. The definition of q implies that

$$\frac{q(a, A)}{q(b, A)} = \frac{\rho(a, A)}{\rho(b, A)} \frac{\rho(\overline{A}_b, A)}{\rho(\overline{A}_a, A)}. \tag{4.4}$$

(Where $\overline{A}_a = A_0 \setminus A_a$ and $\overline{A}_b = A_0 \setminus A_b$.) If Luce's IIA is violated, we must have a change in the "relative probability" of choosing a over b: $\frac{\rho(a,A)}{\rho(b,A)} \neq \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})}$. Hazard Rate IIA implies that the "relative hazard rate" stays the same, $\frac{q(a,\{a,b\})}{q(b,\{a,b\})} = \frac{q(a,A)}{q(b,A)}$. This means that the far-right term of (4.4), $\frac{\rho(\overline{A}_b,A)}{\rho(\overline{A}_a,A)}$, must change as well.

In particular, when either $a \sim c \succ b$ or $a \succ c \succ b$ and the utility of c is large enough, we will have

$$\frac{\rho(a, \{a, b\})}{\rho(b, \{a, b\})} < \frac{\rho(a, \{a, b, c\})}{\rho(b, \{a, b, c\})}.$$

Observe that the definition of $\succ^0$ involves $c \in X$ such that $c \nsim^0 a$ and $c \nsim^0 b$. The subtlety in the definition of $\succ^0$ is to rule out the case $a \sim c \succ b$.

Now, if $a \sim^* b$ then $A_a = A_b$, and Hazard Rate IIA implies the Luce IIA formula for a and b. *Therefore Hazard Rate IIA only differs from Luce's IIA for alternatives that are strictly ordered by perception priority.*

So suppose that a has higher priority than b, and that the relative probability of choosing a over b is smaller when the choice set is $A \cup \{c\}$ than when the choice set is A. Hazard Rate IIA means that the perception priority explains the change in relative probabilities: we must have a compensating decrease in the probability of choosing an element that is perceived before b, relative to the probability of choosing an element that is perceived before a. The explanation is that a was "hurt" relative to b because the choice of a or b depends in part on the probability of choosing an element with higher perception priority, and the addition of c decreased the relative probability of choosing an element with higher priority than b.

In other words, the relative probability of choosing a over b decreased, and therefore Luce's IIA was violated, because the probability of choosing an element that is perceived before b increased relative to the probability of choosing an element that is perceived before a. Hazard Rate IIA means that the only permissible violations of Luce's IIA are those that can be explained in this fashion by the perception priority order.

4.4 Theorem

Before stating the theorem, we also define an additional technical condition called "richness". Richness requires that X has infinitely many alternatives. We do not need this condition to prove the sufficiency of the axioms: that the axioms imply a PALM representation. We need it to prove the necessity of the axioms, in particular, the result that $\succsim = \succsim^*$.

Richness: For any pair $a, b \in X$ with $a \succ b$, there is $c \in X$ with $c \succ a$ or $b \succ c$.[2]

Theorem 6 *If a nondegenerate stochastic choice function ρ satisfies Weak Order and Hazard Rate IIA, then there is a PALM $(u, \succsim)$ such that $\succsim^* = \succsim$ and $\rho = \rho_{(u,\succsim)}$.*

Conversely, for a given PALM $(u, \succsim)$, if $\succsim$ satisfies Richness, then $\rho_{(u,\succsim)}$ satisfies Weak Order and Hazard Rate IIA, and $\succsim = \succsim^*$.

[2] We can prove Theorem 6 when X is finite by slightly modifying the revealed perception priority order $\geq^*$. See Appendix D.7.

The proof of the theorem is in Appendix D.1. The sufficiency of the axioms for the representation is straightforward. The converse of Theorem 6 states, not only that PALM satisfies the axioms, but that $\succsim$ must coincide with $\succsim^*$. The perception priority is thus identified from data on stochastic choice. Therefore, u is unique up to multiplication by a positive scalar. The bulk of the proof is devoted to establishing that $\succsim = \succsim^*$.

Despite the tight behavioral characterization in Theorem 6, PALM is very flexible and can account for some well known behavioral phenomena. We discuss two of them in the next section:

1. The compromise effect; violations of Luce's IIA (Section 4.5.1).

2. The attraction effect; violations of regularity (Section 4.5.2).[3]

4.4.1 Discussion of Outside Option

It is useful to compare how Luce and PALM treat the outside option, the probability of not making a choice from a set A.

For PALM, the utility of the outside option is:

$$u(x_0) = \sum_{a \in A} u(a) \left(\frac{1}{\sum_{a \in A} q(a, A)} - 1 \right). \tag{4.5}$$

In (extended) Luce's model, the utility of the outside option has a similar expression. Indeed,

$$\hat{u}(x_0) = \sum_{a \in A} u(a) \left(\frac{1}{\sum_{a \in A} \rho(a, A)} - 1 \right), \tag{4.6}$$

with ρ in place of the hazard rates q .

It is interesting to contrast the value of $u(x_0)$ according to Equation (4.5) with what one would obtain from Equation (4.6). Given a PALM model $(u, \succsim)$, we can calculate $\hat{u}(A)$ from $\rho_{(u,\succsim)}$ by application of Equation (4.6). If we do that, we obtain

1. $\hat{u}(x_0) \geq u(x_0)$,

2. and $\hat{u}(x_0) = u(x_0)$ when $a \sim b$ for all $a, b \in A$.

[3] PALM can also account for violations of stochastic transitivity.

The inequality $\hat{u}(x_0) \geq u(x_0)$ reflects that there are two sources behind choosing the outside option in PALM. One source is the utility $u(x_0)$ of not making a choice; this is the same as in Luce's model with an outside option. The second source is due to the sequential nature of choice in PALM. When we consider an agent that chooses sequentially, following the priority order $\succsim$, then it is possible that we exhaust the elements in A without making a choice. When that happens, it would seem to inflate (or bias) the value of the outside option; as a result we get that $\hat{u}(x_0) \geq u(x_0)$. For example, when the utility of the outside option is zero, the outside option will not be chosen in Luce's model. However, in PALM, the outside option will be chosen with positive probability because of the second source behind choosing the outside option. Therefore, the utility of the outside option must be negative in order to choosing the outside option with zero probability.[4]

Choice Overload

The outside option in PALM allows us to capture various behavioral phenomena. One example is "choice overload:" the idea that a subject may be inclined to make *no choice* when presented with many alternatives. The paper by Iyengar and Lepper (2000) is a well known study of choice overload. Iyengar and Lepper run an experiment where subjects had to choose among a large set of nearly identical alternatives. They find that a large fraction of subjects make no choice whatsoever, and that the fraction of subjects who make no choice increases from 26% to 40% as the number of alternatives increases. These results are easily captured by PALM.

Let $A = \{a_1, \ldots, a_n\}$ be a menu with n elements, each of which provide the same Luce utility; so $u(x_1) = u(x_2) = \ldots = u(x_n) > 0$. Suppose that the n elements in A are strictly ordered by the perception priority $\succsim$, and that $u(x_0) = 0$. Then the probability of choosing the outside option is

$$\rho(x_0, A) = (1 - 1/n)^n,$$

which is monotone increasing in n. In other words, the probability of not making a choice in A increases as the cardinality of A increases. Moreover, $\rho(x_0, A)$ goes from about 25% to $\frac{1}{e} \approx 37\%$ as n increases. So PALM approximately matches the numbers in the Iyengar and Lepper experiment.

[4] For example, when $a \succ b$ and $u(x_0) = -u(a)$, we have $\rho(x_0, \{a, b\}) = 0$.

Forced Choice

Another advantage of PALM's treatment of the outside option is that it allows us to understand forced choice. In particular, the presence of the outside option allows us to compare environments in which agents are forced to make a choice with environments in which they are not forced to make a choice. Our model is consistent with the experimental results of Dhar and Simonson (2003) on the effects of forced choice on choice (Section 4.5.3). Moreover, we show that if an agent chooses not to make choice with high probability, then utility and perception are positively correlated (Section 4.6).

4.5 Compromise and Attraction Effects

The compromise and attraction effects are well-known deviations from Luce's model. See Rieskamp et al. (2006) for a survey. In this section, we demonstrate how PALM can capture each of these phenomena.

The compromise and attraction effects are defined in the same kind of experimental setup. An agent makes choices from the sets $\{x, y\}$ and $\{x, y, z\}$. The "effects" relate to the consequences of adding the alternative z. PALM can explain the attraction effect when one uses *familiarity* to infer perception priority, so familiar objects are perceived before unfamiliar ones. The role of familiarity in the compromise and attraction effects is documented in Sheng et al. (2005) and Ratneshwar et al. (1987), respectively.

In the following, we review experiments on the compromise effects and attraction effects. To apply PALM to such experiments, we need to discuss what is the outside of option in the experiments. We claim that one natural interpretation of the outside option is opting out from the experiments. In the next two sections we show how PALM captures the compromise and attraction effects.[5] Then we investigate how the option to opt out of the experiment affects the results.

4.5.1 Compromise Effect– Violation of IIA

Consider three alternatives, x, y and z. Suppose that x and z are "extreme" alternatives, while y represents a moderate middle ground, a compromise. In the experiment studied by Simonson and Tversky (1992), x is X-370, a very basic model of Minolta camera; y is MAXXUM 3000i, a more advanced model of the

[5] Appendix D.8 discusses a modification of PALM which avoids the outside option. Using the modification, we illustrate that the outside option does not really play a role in explaining the two effects, but the sequential procedure does.

same brand; and z is MAXXUM 7000i, the top of the line offered by Minolta in this class of cameras.

Model	Price ($)	Choices Exp. 1	Choices Exp. 2
x (X-370)	169.99	50%	22%
y (MAXXUM 3000i)	239.99	50 %	57%
z (MAXXUM 7000i)	469.99	N/A	21%

Figure 4.1: Compromise effect in Simonson and Tversky (1992)

The agent's choice set is $\{x, y\}$ in Experiment 1 and $\{x, y, z\}$ in Experiment 2. The experimental data show that the probability of choosing y increases when moving from Experiment 1 to 2 (see Figure 4.1). Simonson and Tversky (1992) call this phenomenon the *compromise effect.* As in Rieskamp et al. (2006), the compromise effect can be written as follows:

$$\frac{\rho(x, \{x, y, z\})}{\rho(y, \{x, y, z\})} < 1 \leq \frac{\rho(x, \{x, y\})}{\rho(y, \{x, y\})}. \tag{4.7}$$

Proposition 8 *When $x \succ y \succsim z$, $\rho_{(u,\succsim)}$ exhibits the* compromise effect *(i.e., (4.7)) if and only if $u(y) > u(x)$ and*

$$u(z) + u(x_0) > \frac{u^2(x) - u^2(y) + u(x)u(y)}{u(y) - u(x)} \geq u(x_0). \tag{4.8}$$

Proposition 8 results from a straightforward calculation so the proof is omitted.

Simonson and Tversky (1992)'s explanation for the compromise effect is that subjects are averse to extremes, which helps the "compromise" option y when facing the problem $\{x, y, z\}$. PALM can capture the compromise effect when we assume that y is "in between" x and z with respect to priority. One rationale for $x \succ y \succsim z$ is familiarity. The basic camera model may be more familiar, while the top of the line is the least familiar.

4.5.2 Attraction Effect– A Violation of Regularity

PALM can accommodate violations of regularity. We focus on the attraction effect, a well-known violation of regularity. A famous example of the attraction effect is documented by Simonson and Tversky (1992) using the following experiment. Consider our three alternatives again, x, y and z. Suppose now that y and z are different variants of the same good: y is a Panasonic microwave oven (meaning a higher quality and expensive good[6]), while z is a more expensive version of y: z is dominated by y. The alternative x is an Emerson microwave oven (meaning a lower quality and cheap good). A more recent example, which we discussed in the introduction, is due to Doyle et al. (1999). As we mentioned in the introduction, the findings in Doyle et al.'s experiments fit the story in PALM particularly well.

Option	Choices Exp. 1	Choices Exp. 2
x (Emerson)	57 %	27 %
y (Panasonic I)	43 %	60 %
z (Panasonic II)	N/A	13 %

Figure 4.2: Attraction effect in Simonson and Tversky (1992)

Simonson and Tversky (1992) (p. 287) asked subjects to choose between x and y in Experiment 1, and to choose among x, y, and z in Experiment 2 (see Figure 4.2). They found that the share of subjects who chose y in Experiment 2 is higher than in Experiment 1. This finding is called the *attraction effect*. As in Rieskamp et al. (2006), the effect can be described as follows:

$$\rho(y, \{x, y, z\}) > \rho(y, \{x, y\}). \tag{4.9}$$

Proposition 9 *If $x \succ y \succsim z$ and $u(x)$ is large enough, then $\rho_{(u,\succsim)}$ exhibits the* attraction effect *(i.e., (4.9)).*

Proof of Proposition 9: We have

$$\rho(y, \{x, y, z\}) > \rho(y, \{x, y\}) \Leftrightarrow$$

$$\Leftrightarrow q(y, \{x, y, z\})(1 - q(x, \{x, y, z\})) > q(y, \{x, y\})(1 - q(x, \{x, y\}))$$

$$\Leftrightarrow u(x) > \sqrt{(u(y) + u(z) + u(x_0))(u(y) + u(x_0))}$$

■

The assumption $x \succ y \succsim z$ means that the Emerson microwave x is more salient than the Panasonic microwaves, perhaps because of its price. The first Panasonic

[6]Microwave ovens were at one point expensive; see "Money for nothing" by Dire Straits.

microwave y is at least as salient as z since there are the same brands. It is also possible to tell a story of familiarity for the microwaves experiment. The Emerson microwave x is likely to be the most familiar alternative since it is the cheapest and simplest model. In Doyle et al.'s experiments (as discussed in the introduction), perception is related to the familiarity of the brand of beans.

A different, symmetric, experiment would be to add an alternative t to enhance the choice of x. So t could be a more expensive version of x. Heath and Chatterjee (1995) found that one is *less* likely to observe the attraction effect when the third alternative is dominated by the low-quality alternative (x), compared to the high-quality alternative (y). More precisely, one is more likely to have $\rho(y, \{x, y, z\}) > \rho(y, \{x, y\})$ compared to $\rho(x, \{x, y, t\}) > \rho(x, \{x, y\})$. PALM is consistent with this finding: we cannot have $\rho(x, \{x, y, t\}) > \rho(x, \{x, y\})$ when $x > y$.[7]

4.5.3 The Effects of Forced Choice

Dhar and Simonson (2003) run choice experiments in which agents may not have to make a choice. In their design, "no-choice" and "forced choice" are two experimental treatments. Under the no-choice option, subjects can opt not to make a choice. Under the forced-choice treatment, subjects must make a choice, as in the experiments described in the two previous sections. Dhar and Simonson show that the introduction of the no choice option weakens the compromise effect and decrease the relative share of an option that is "average" on all dimensions. In our model, not making a choice corresponds to choosing the outside option x_0. We proceed to illustrate how PALM can capture the evidence presented by Dhar and Simonson.

Consider two PALM models, $\rho = \rho_{(u,\succsim)}$ and $\rho^f = \rho_{(u^f,\succsim^f)}$. Suppose that these two models only differ in $u(x_0)$. Thus, $u(x) = u^f(x)$ for any $x \in X$ and $\succsim=\succsim^f$. We assume that $\rho(x_0, A) > \rho^f(x_0, A)$ for all $A \in \mathscr{A}$. Roughly speaking, in the PALM ρ^f, a decision maker chooses the outside option less often.

In PALM, choosing the outside option more frequently is tied to a larger utility of the outside option. In particular:

Condition ♠: $\rho(x_0, \{x, y\}) > \rho^f(x_0, \{x, y\})$ iff $\rho(x_0, \{x, y, z\}) > \rho^f(x_0, \{x, y, z\})$ iff $u(x_0) > u^f(x_0)$.

[7]However, it is still possible that the relative probability of choosing x increases; that is, $\frac{\rho(x,\{x,y,t\})}{\rho(y,\{x,y,t\})} > \frac{\rho(x,\{x,y\})}{\rho(x,\{x,y\})}$, when t is added where $x \succsim t > y$.

Proposition 10 *Any PALM model satisfies Condition ♠.*

Recall our discussion of the outside option in PALM. There are two sources behind the choice of the outside option. One is the sequential nature of choice, and the other is the utility of the outside option. Proposition 10 says that the utility of the outside option is the source behind "global" increases in the probability of choosing the outside option.

In light of Proposition 10, we can trace the probability of opting out of an experiment, and not making a choice, to the incentives provided for participation in the experiment. In particular, consider the findings of Dhar and Simonson. Fix three alternatives $x, y, z \in X$. Suppose that $x \succ y \succ z$. So y can be interpreted as an "average" option. Given our assumption on ρ and ρ^f, and under Condition ♠, we assume that $u(x_0) > u^f(x_0)$.

In first place, PALM can capture Dhar and Simonson's finding that the no-choice option decreases the relative share of an average alternative (recall that ρ represents the case where subjects exercise the outside "no-choice" option more):

Proposition 11 *If* $u(x_0) > u^f(x_0)$*, then*

$$\frac{\rho(y,\{x,y\})}{\rho(x,\{x,y\})} > \frac{\rho^f(y,\{x,y\})}{\rho^f(x,\{x,y\})} \text{ and } \frac{\rho(y,\{x,y,z\})}{\rho(x,\{x,y,z\})} > \frac{\rho^f(y,\{x,y,z\})}{\rho^f(x,\{x,y,z\})}.$$

Proof of Proposition 11: By a direct calculation, $\frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})} = \frac{u(x)}{u(y)}(1 + \frac{u(x)}{u(y)+u(x_0)})$ and $\frac{\rho^f(x,\{x,y\})}{\rho^f(y,\{x,y\})} = \frac{u(x)}{u(y)}(1 + \frac{u(x)}{u(y)+u^f(x_0)})$. Since $f(t) = \frac{u(x)}{u(y)}(1 + \frac{u(x)}{u(y)+t})$ is decreasing in t, we obtain $u(x_0) > u^f(x_0)$ if and only if $\frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})} < \frac{\rho^f(x,\{x,y\})}{\rho^f(y,\{x,y\})}$. Similarly, $\frac{\rho(x,\{x,y,z\})}{\rho(y,\{x,y,z\})} = \frac{u(x)}{u(y)}(1 + \frac{u(x)}{u(y)+u(z)+u(x_0)})$ and $\frac{\rho^f(x,\{x,y,z\})}{\rho^f(y,\{x,y,z\})} = \frac{u(x)}{u(y)}(1 + \frac{u(x)}{u(y)+u(z)+u^f(x_0)})$. Since $g(t) = \frac{u(x)}{u(y)}(1 + \frac{u(x)}{u(y)+u(z)+t})$ is decreasing in t, we obtain $u(x_0) > u^f(x_0)$ if and only if $\frac{\rho(x,\{x,y,z\})}{\rho(y,\{x,y,z\})} < \frac{\rho^f(x,\{x,y,z\})}{\rho^f(y,\{x,y,z\})}$. ■

In second place, PALM can capture Dhar and Simonson's finding that the no-choice option weakens the compromise effect as follows:

Proposition 12 *If* $u(x_0) > u^f(x_0)$*, then*

$$\frac{\rho^f(x,\{x,y\})}{\rho^f(y,\{x,y\})} \Big/ \frac{\rho^f(x,\{x,y,z\})}{\rho^f(y,\{x,y,z\})} > \frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})} \Big/ \frac{\rho(x,\{x,y,z\})}{\rho(y,\{x,y,z\})}.$$

4.6 Correlation between Utility u and Perception Priority $\succsim$

Perception and utility are two independent parameters in PALM. Therefore, PALM allows us to model scenarios where perception is positively correlated with utility, negatively correlated, or simply unrelated.

In experimental settings, Reutskaja et al. (2011) find no intrinsic correlation between utility and perception (a similar finding is reported in Krajbich and Rangel 2011). High-utility items are not *per se* more likely to be perceived more prominently than others. It is therefore important that PALM not force a particular relation between perception and utility.

However, we argue that when an agent chooses the outside option with high probability, it is likely that utility and perception are positively correlated. Now we give some conditions under which $u(a) > u(b)$ if and only if $a \succ b$.

Proposition 13 *Suppose $a \nsim b$ and $u(a) \neq u(b)$. If $\rho(x_0, \{a, b\}) \geq \min\{\rho(a, \{a, b\}), \rho(b, \{a, b\})\}$ and $u(x_0) \leq 0$, then $u(a) > u(b)$ if and only if $a \succ b$.*

The condition that $\rho(x_0, \{a, b\}) \geq \min\{\rho(a, \{a, b\}), \rho(b, \{a, b\})\}$ means that the probability of choosing the outside option must be large enough. This property is necessary to achieve positive correlation, as evidenced in the following result.

Proposition 14 *If $a \succ b$, $u(a) > u(b)$, and $u(b) - u(a) \leq u(x_0)$, then*

$$\rho(a, \{a, b\}) > \rho(x_0, \{a, b\}) \geq \rho(b, \{a, b\}) = \min\{\rho(a, \{a, b\}), \rho(b, \{a, b\})\}.$$

4.7 Related Literature

Section 4.5 explains how PALM relates to the relevant empirical findings, including the compromise, and attraction effects. We now proceed to discuss the relation between PALM and some of the most important theoretical models of stochastic choice.

There is a non-axiomatic literature proposing models that can explain the compromise and attraction effects. Rieskamp et al. (2006) is an excellent survey. Examples are Tversky (1972b), Roe et al. (2001) and Usher and McClelland (2004). The latter two papers propose *decision field theory*, which allows for violations of Luce's regularity axiom. The recent paper by Natenzon (2010) presents a learning model, in which an agent learns about the utility of the different alternatives, and makes a choice with imperfect knowledge of these utilities. Learning is random, hence choice is stochastic. Natenzon's model can explain these effects (as well as the similarity effect of Tversky 1972b).

We shall not discuss these papers here, and focus instead on the more narrowly related axiomatic literature in economics.

1) The benchmark economic model of rational behavior for stochastic choice is the random utility model. Luce's model is a special case of both PALM and random utility. So PALM and random utility are not mutually exclusive; PALM is, however, not always a random utility model.

The random utility model is described by a probability measure over preferences over X; $\rho(x, A)$ is the probability of drawing a utility that ranks x above any other alternative in A. The random utility model is famously difficult to characterize behaviorally: see the papers by Falmagne (1978), McFadden and Richter (1990), and Barberá and Pattanaik (1986).

As we have seen in Section 4.5, there are instances of PALM that violate the regularity axiom. A random utility model must always satisfy regularity. Thus PALM is not a special case of random utility. Moreover, Luce's is a random utility model, and a special case of PALM. So the class of PALM and random utility models intersect, but they are distinct.

2) The recent paper by Gul et al. (2014) presents a model of random choice in which object attributes play a key role. Object attributes are obtained endogenously from observed stochastic choices. Their model has the Luce form, but it applies sequentially; first for choosing an attribute and then for choosing an object. In terms of its empirical motivation, the model seeks to address the similarity effect.

Gul, Natenzon and Pesendorfer's model is a random utility model (in fact they show that any random utility model can be approximated by their model). There are therefore instances of PALM that cannot coincide with the model in Gul et al. (2014). (Importantly, PALM can explain violations of the regularity axiom.) On the other hand, Luce's model is a special case of their model and of PALM. So the two models obviously intersect.

3) Manzini and Mariotti (2014) study a stochastic choice model where *attention* is the source of randomness in choice. In their model, preferences are deterministic, but choice is random because attention is random. Manzini and Mariotti's model

takes as parameters a probability measure g on X, and a linear order $>_M$. Their representation is then

$$\rho(a, A) = g(a) \prod_{a' >_M a} (1 - g(a')).$$

In PALM, perception is described by the (non-stochastic) perception priority relation $\succsim$. Choice is stochastic because it is dictated by utility intensities, similarly to Luce's model. In Manzini and Mariotti, in contrast, attention is stochastic, but preference is deterministic.

Manzini and Mariotti's representation looks superficially similar to ours, but the models are in fact different to the point of not being compatible, and seek to capture totally different phenomena. Manzini and Mariotti's model implies that IIA is violated for any pair x and y, so their model is incompatible with Luce's model. PALM, in contrast, has Luce as a special case. Appendix D.2 shows that the two models are disjoint. Any instance of their model must violate the PALM axioms, and no instance of PALM can be represented using their model. So their model and ours seek to capture completely different phenomena.

4) A closely related paper is Tserenjigmid (2013). In this paper, an order on alternative also matters for random choice, and the model can explain the attraction and compromise effects (as well as the similarity effect). The source of violations of IIA is not perception, but instead a sort of menu-dependent utility.

5) The paper by Fudenberg et al. (2013b) considers a decision maker who chooses a probability distribution over alternatives so as to maximize expected utility, with a cost function that ensures that probabilities are non-degenerate. One version of their model can accommodate the attraction effect, and one can accommodate the compromise effect.

6) Some related studies use the model of non-stochastic choice to explain some of the experimental results we describe in Section 4.5. This makes them quite different, as the primitives are different. The paper by De Clippel and Eliaz (2012) is important to mention; it gives an axiomatic foundation for models of non-stochastic choice that can capture the compromise effect. PALM gives a different explanation for the compromise effect, in the context of stochastic choice.

Another related paper is Lleras et al. (2010). (See also Masatlioglu et al. (2012) for a different model of attention and choice.) They attribute violations of IIA to the role of attention. They elicit revealed preference (not perception priority, but preference) in a similar way to ours. When the choice from $\{x, y, z\}$ is x and from $\{x, z\}$ is z, then they conclude that x is revealed preferred to z (this is in some sense, the opposite of the inference we make).

7) Some papers study deliberate stochastic choice due to non-expected utility or uncertainty aversion. Machina (1985) proposes a model of stochastic choice of lotteries. In Machina's paper, an agent deliberately randomizes his choices due to his non-expected utility preferences. Machina does not provide an axiomatization. Saito (2015) axiomatizes a model of stochastic choice of act. In Saito's model, an agent deliberately randomizes his choices because of non-unique priors over the set of states. Saito's primitives is preferences over sets of acts (i.e., payoff-profiles over the set of states).

8) Ravid (2015) studies a random choice procedure that is similar to PALM. First, an agent picks an option at random from the choice set; the option becomes "focal." Second, she compares the focal option to each other alternative in the set. Third, the agent chooses the focal option if it passes all comparisons favorably. Otherwise, the agent draws a new focal option with replacement. Ravid (2015) characterizes the procedures by an relaxation of IIA termed Independence of Shared Alternatives (ISA).

9) Marley (1991) axiomatizes a more general model than the model proposed by Ravid. Marley (1991) called the model, binary advantage models. In the models, choice probabilities depend on a measure of binary advantage and an aggregation function which maps choice set to a strictly positive number. The model by Ravid is a special case of binary advantage models, in which the aggregation function is constant and equal to one.

Chapter 5

THE ORDER-DEPENDENT LUCE MODEL

5.1 Introduction

Luce (1959b)'s model of random choice (or multinomial logit model) has been widely used and studied in econometrics and economic theory (see McFadden 2001). This model relies on the axiom of *independence from irrelevant alternatives* (IIA) which says that the probability of choosing alternative x relative to that of choosing alternative y is not affected by the presence of alternatives other than x or y. This kind of menu-independence assumption is also common in standard choice theory models. However, many consistent violations of IIA are documented both in the real world and in lab experiments.[1]

In this paper, we develop a menu-dependent version of Luce's Model in which menu-dependence is caused by an underlying linear order on alternatives. To illustrate the menu-dependence of interest, consider a consumer choice in a grocery store. It is well-understood that the location of a product in the shop can make a crucial difference to its sales. Traditionally, "eye-level shelving" is best, followed by "waist-level", "knee-level", and "ankle-level" (Dreze et al. 1995). Simple product reorganization can produce changes in sales of 5-6% (Dreze et al. 1995). Since space is limited, the location of one product affects the location of all other products which in turn affects the sales of those products. Therefore, the locations of products potentially cause menu-dependent choice.

Another example of menu-dependent behavior is the effect of ballot order on vote shares. Meredith and Salant (2013) estimated that, in California city council and school board elections, candidates listed first win office between 4-5% more often than expected, absent order effects. Candidates listed first in primary or non-partisan elections for U.S. state or federal offices gain about 2% points (Koppel and Steen 2004, Ho and Imai 2008).

In this paper, we focus on an environment in which alternatives are ordered by some underlying (fixed) linear order and an agent's behavior is influenced by the

[1]For example, there are many experimental results that suggest a violation of IIA. See Tversky (1972a), Huber et al. (1982), Simonson and Tversky (1992), and Tversky and Simonson (1993). Also see Rieskamp et al. (2006) for survey.

ordering over alternatives in the menu she faces. We axiomatically model an agent who makes a random choice in such environment. We characterize our model by two weakenings of IIA.

In many cases, the underlying linear order is subjective, unobservable, or too costly to obtain for a researcher. For example, the attractiveness or salience of alternatives affect the agent's behavior, but they are purely subjective. Also, advertisement expenditures and Google search ranking give an objective ranking on alternatives, but it is costly to obtain for the researcher. One of the main contributions of the paper is that we can uniquely identify the underlying linear order from observed choice probabilities (Propositions 16-17).

Despite having a tight axiomatic characterization, our model can explain many well-known behavioral phenomena. In Section 5.5, we will show that our model can accommodate the *similarity effect* of Tversky (1972a), the *compromise effect* of Tversky and Simonson (1993), the *attraction effect* of Huber et al. (1982), a violation of stochastic transitivity (Tversky 1969), and *the choice overload* (Iyengar and Lepper 2000). Since the attraction effect is a violation of regularity and random utility models always satisfy regularity (see Section 5.5.2), our model is not a special case of random utility models.

The paper proceeds as follows. In Section 5.2, we introduce basic notations and definitions. In Section 5.3, we focus on a case in which the underlying linear order on alternatives is observed and provide a characterization theorem. In Section 5.4, we focus on a case in which the underlying linear order on alternatives is not observed. Section 5.5 discusses aforementioned behavioral phenomena and Section 5.6 concludes with related literature. Proofs are collected in Appendix E.

5.2 Basic

Let X be a finite set of alternatives with at least three elements, and $\mathscr{A}$ be a set of nonempty subsets (menus) of X. We assume that $\mathscr{A}$ includes all binary and tripleton menus. We model an agent who makes a probabilistic choice from A with $A \in \mathscr{A}$.

Definition 17 (Random Choice Rule) *A function $p : X \times \mathscr{A} \to [0,1]$ is called a* random choice rule *if*

$$\sum_{a \in A} p(a, A) = 1$$

for all $A \in \mathscr{A}$. A random choice rule p is nondegenerate *if $p(a, A) > 0$ for any A and $a \in A$.*

Throughout the paper we focus on nondegenerate random choice rules since a zero probability is empirically indistinguishable from a positive but small probability. Now we formally define IIA and Luce's model.

Definition 18 (IIA) *A random choice rule p satisfies* Luce's independence of irrelevant alternatives (IIA) axiom *if for any $a, b \in X$ and $A, B \in \mathscr{A}$ with $a, b \in A \cap B$,*

$$\frac{p(a, A)}{p(b, A)} = \frac{p(a, B)}{p(b, B)}.$$

Luce (1959b) proves that, if a nondegenerate random choice rule satisfies IIA, then it can be represented by the following model: p satisfies Luce's model if there exists a positive-valued function u (menu-independent utility function) on X such that

$$p(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a')}.$$

In this paper, we focus on a menu-dependent version of Luce's model. However, without a specific kind of menu-dependence, we cannot obtain any testable predictions on observed data. In other words, any nondegenerate random choice rule can be represented by some menu-dependent version of Luce's Model; that is, there exists a positive-valued function u (menu-dependent utility function) on $X \times \mathscr{A}$ such that

$$p(a, A) = \frac{u(a, A)}{\sum_{a' \in A} u(a', A)}.$$

In this paper, we consider a menu-dependent Luce's model in which menu-dependence is determined by some underlying linear order on alternatives. Let R be the underlying linear order on X of interest. Let $R(a, A) \equiv |\{b \in A : bRa\}| + 1$ which gives a ranking of an alternative a in a menu A by R. Now we formally define our model.

Definition 19 *An* order-dependent Luce Model *(ODLM) is a pair (u, R) of a linear order R on X and a function $u : X \times \mathbb{N} \to \mathbb{R}_+$ such that for any $A \in \mathscr{A}$,*

$$p(a, A) = \frac{u(a, R(a, A))}{\sum_{a' \in A} u(a', R(a', A))}. \tag{5.1}$$

For any ODLM (u, R), we denote by $p_{(u,R)}$ the random choice rule defined through (5.1). The ODLM includes the following three different choices:

- **Luce's model:** if $u(a, i) = u(a, j)$ for any $a \in X$ and $i, j \in \mathbb{N}$, then an OLDM produces Luce's model.
- **Top-3:** when $u(\cdot, i) > 0$ only if $i \leq 3$, an ODLM (u, R) produces Top-3 choice.
- **Advertisement Expenditures:** suppose R is a ranking of products in a market by their advertisement expenditures. Suppose agent's evaluations of products are weighted depending on that ranking, then the agent's choice generates an ODLM (u, R) such that $u(a, R(a, A)) = v(a) \cdot w(R(a, A))$ where v is a utility function and w is a weight function.

Before we proceed to the characterization theorems, let us discuss two simple predictions of the model. For sake of clarity, we say that *IIA is satisfied at alternatives $a, b,$ and c*, if the probability of choosing a from menu $\{a, b, c\}$ relative to that of choosing b does not change if c is eliminated from the menu; that is,

$$\frac{p(a, \{a, b, c\})}{p(b, \{a, b, c\})} = \frac{p(a, \{a, b\})}{p(b, \{a, b\})}.$$

Now take any ODLM (u, R) and three alternatives a, b, and c with $a\, R\, b\, R\, c$. Suppose $u(c, 2) > u(c, 3)$. Then the model predicts, the following two observations.

Observation 4: IIA is satisfied at a, b, and c since

$$\frac{p(a, \{a, b, c\})}{p(b, \{a, b, c\})} = \frac{u(a, 1)}{u(b, 2)} = \frac{p(a, \{a, b\})}{p(b, \{a, b\})}.$$

In other words, an alternative with a lower ranking (c) has no effect on alternatives with higher ranking (a and b).

Observation 5: IIA is violated at a, c, and b since

$$\frac{p(a, \{a, b, c\})}{p(c, \{a, b, c\})} = \frac{u(a, 1)}{u(c, 3)} > \frac{p(a, \{a, c\})}{p(a, \{a, c\})} = \frac{u(a, 1)}{u(c, 2)}.$$

In other words, eliminating an alternative (b) hurts an alternative with higher ranking (a) more than an alternative with a lower ranking (c).

5.3 Axioms and Representation Theorem for given R

In some cases the ordering of alternatives is observable. For example, in elections the ordering of candidates on ballots is observable to both voters and researchers. In this section, we suppose that the linear order R is given. For the given order R, two weakenings of IIA are enough to characterize the ODLM.

To illustrate the first axiom, let us consider consumer choices in a grocery store. Suppose a grocery store has three shelves, top, middle, and bottom, and goods a, b, and c are on each one of them. Suppose being on the top shelf gives 5% advantage over being on the middle shelf, and being on the middle shelf gives 3% advantage over being on the bottom shelf in terms of sales. Consider two scenarios.

First, consider a case in which a is on the top, b is in the middle, and c is at the bottom. Then the elimination of c does not effect the probability of choosing a relative to that of choosing b because the rankings of a and b in $\{a, b, c\}$ are the same that of in $\{a, b\}$. Therefore, a still should have 5% advantage over b, so IIA should be satisfied consistent with Observation 4.

Second, consider a case in which a is in the middle, b is at the bottom, and c is on the top. Now let us replace c with a new alternative c'. Since the rankings of a and b in $\{a, b, c\}$ are the same that of in $\{a, b, c'\}$, a should still have 3% advantage over b.

We formalize the idea of the above two scenarios and obtain a weaker version of IIA: the probability of choosing a relative to that of choosing b is constant across menus in which their rankings are the same. Formally,

Axiom 19 (R-IIA) *For any $A, B \in \mathscr{A}$ and $a, b \in A \cap B$, if*

$$R(a, A) = R(a, B) \text{ and } R(b, A) = R(b, B), \text{ then } \frac{p(a, A)}{p(b, A)} = \frac{p(a, B)}{p(b, B)}.$$

Indeed, R-IIA is weaker than IIA. Now we turn to the second axiom. To illustrate the second axiom, let us discuss the following indirect implication of IIA:

$$\frac{p(a, \{a, c\})}{p(c, \{a, c\})} \cdot \frac{p(c, \{b, c\})}{p(b, \{b, c\})} = \frac{p(a, \{a, b\})}{p(b, \{a, b\})}. \tag{5.2}$$

Luce interpreted (5.2) as follows: let $p(a, A)$ be the probability of a being the best element in A for some (random) preference $\succ^*$ on X; that is,

$$p(a, A) = Pr\left(a \succ^* b \text{ for all } b \in A \setminus \{a\}\right).$$

Then the product $p(a, \{a, c\}) \cdot p(c, \{b, c\})$ is equal to the probability of the event $\{a \succ^* c \succ^* b\}$ and the product $p(c, \{a, c\}) \cdot p(b, \{b, c\})$ is equal to the probability of the event $\{b \succ^* c \succ^* a\}$. Then we rewrite (5.2) in the following way:

$$\text{LHS of (5.2)} = \frac{Pr(a \succ^* c \succ^* b)}{Pr(b \succ^* c \succ^* a)} = \frac{Pr(a \succ^* b)}{Pr(b \succ^* a)} = \text{RHS of (5.2)}.$$

In other words, equation (5.2) requires that the product $\frac{p(a,\{a,c\})}{p(c,\{a,c\})} \cdot \frac{p(c,\{b,c\})}{p(b,\{b,c\})}$ ($= \frac{p(a,\{a,b\})}{p(b,\{a,b\})}$) does not depend on c. We weaken this property in the following way: the above product does not depend on c when c has a lower ranking than both a and b by R. Formally,

Axiom 20 (R-Constant Impact (R-CI)) *For any $a, b \in X$, the product $\frac{p(a,\{a,c\})}{p(c,\{a,c\})} \cdot \frac{p(c,\{b,c\})}{p(b,\{b,c\})}$ is constant across all $c \in X$ such that aRc and bRc; that is, for any $c, c' \in X$ such that aRc, bRc, aRc', and bRc',*

$$\frac{p(a, \{a, c\})}{p(c, \{a, c\})} \cdot \frac{p(c, \{b, c\})}{p(b, \{b, c\})} = \frac{p(a, \{a, c'\})}{p(c', \{a, c'\})} \cdot \frac{p(c', \{b, c'\})}{p(b, \{b, c'\})}.$$

Now we state the first main result of our paper.

Theorem 7 *A nondegenerate random choice rule p satisfies R-IIA and R-CI if and only if there is an ODLM (u, R) such that $p = p_{(u,R)}$. Moreover, u is unique up to multiplication by a positive scalar.*

5.3.1 Increasing Order-Dependent Luce Model

Let us consider the grocery store example again. In that example, evaluations of alternatives are increasing in order of shelves: being on the middle shelf is better than being on the bottom and being on the top shelf is better than being in the middle. In this subsection, we discuss a case in which the utilities of alternatives are increasing in rankings of alternatives (decreasing in R). Formally, we say an ODLM (u, R) is *increasing* if for any i, j with $i < j$, $u(a, i) \geq u(a, j)$.

Consider the grocery store example in which a is on the top, b is in the middle, and c is on the bottom. Similar to Observation 5 in Section 5.2, the elimination of b from $\{a, b, c\}$ helps c more than a because c moves to the middle shelf while a is still on the top shelf. Formally,

Axiom 21 (R-Increasing) $\frac{p(c,A)}{p(a,A)} \leq \frac{p(c,A\setminus\{b\})}{p(a,A\setminus\{b\})}$ *for all A and $a, b, c \in A$ with $a\,R\,b\,R\,c$.*

We also require the following condition: when $a\,R\,b\,R\,c$,

$$\frac{p(a,\{a,b\})}{p(b,\{a,b\})}\cdot\frac{p(b,\{b,c\})}{p(c,\{b,c\})}\geq\frac{p(a,\{a,c\})}{p(c,\{a,c\})}. \tag{5.3}$$

In the grocery story example, if the consumer prefers all alternatives equally, since the top shelf has 5% advantage over the middle one, then LHS of (5.3) is $(1+5\%)(1+5\%)$ and RHS of (5.3) is $(1+5\%)$. Now we state the second result of the paper. Above two conditions are enough to characterize increasing ODLMs.

Proposition 15 *For any ODLM (u, R), a random choice rule $p = p_{(u,R)}$ satisfies R-Increasing and (5.3) for any $a, b, c \in X$ with $aR\,bR\,c$ if and only if it is increasing.*

5.4 Identifying Unknown R and Revealed Order

In many cases R is not observable. For example, suppose that R is a ranking over products in a market by their levels of advertisement spending. Indeed, R affects the choices of consumers, but it is not observable to outside researchers. We identify R from observable choice data by using Observation 4.

Let us restate the main message of Observation 4: an alternative with lower ranking has no effect on alternatives with higher ranking. Therefore, in order to identify R, we simply take menus with three alternatives and check whether IIA is satisfied.

More formally, we say that *c has lower rank than a and b* if IIA is satisfied at a, b, and c; that is, $\frac{p(a,\{a,b,c\})}{p(b,\{a,b,c\})}=\frac{p(a,\{a,b\})}{p(b,\{a,b\})}$. Now for any two alternatives a and b, let us define the set of elements that have lower rank than both a and b:

$$L(a,b)=\{c\in X\setminus\{a,b\}|\ \frac{p(a,\{a,b,c\})}{p(b,\{a,b,c\})}=\frac{p(a,\{a,b\})}{p(b,\{a,b\})}\}.$$

Now we define a revealed order R^0 on X from the sets L.

Definition 20 (Revealed Order) *for any $a, b \in X$, we say a* has revealed higher ranking than *b, denoted by aR^0b, if there exists $a' \in X$ such that $b \in L(a, a')$.*

Two postulates on L are enough to obtain a well-behaved revealed order R^0 which almost uniquely identifies the underlying order R.

The first axiom is called *asymmetry*, which requires that if a has revealed higher ranking than b, then b cannot have revealed higher ranking than a. Formally,

Axiom 22 (Asymmetry) *For any $a, b \in X$, if there exists $c \in X \setminus \{a, b\}$ such that $b \in L(a, c)$, then $a \notin L(b, c')$ for all $c' \in X \setminus \{a, b\}$.*

The second axiom is called *transitivity*, which requires that if c does not have revealed higher ranking that b, and c does not have revealed higher ranking that a, then both a and b have revealed higher ranking than c. Formally,

Axiom 23 (Transitivity) *For any $a, b, c \in X$, if $b \notin L(a, c)$ and $a \notin L(b, c)$, then $c \in L(a, b)$.*

The following result proves that R^0 is well-behaved and almost complete under asymmetry and transitivity.

Proposition 16 (Sufficiency) *Suppose a nondegenerate random choice rule p satisfies Asymmetry and Transitivity. Then R^0 is transitive; that is, aR^0b and bR^0c imply aR^0c, and asymmetric; that is, aR^0b implies $\neg bR^0a$. Moreover, if there is a pair (a^*, b^*) such that neither $a^*R^0b^*$ nor $b^*R^0a^*$, then for any $(a, b) \neq (a^*, b^*)$, either aR^0b or bR^0a, and for any $c \in X \setminus \{a^*, b^*\}$, a^*R^0c and b^*R^0c.*

Proposition 16 shows that Asymmetry and Transitivity are sufficient for R^0 to be well behaved. The next result shows that Asymmetry and Transitivity are necessary under conditions on the utility function u give us enough variety in choice probabilities. Proposition 16 also shows that R^0 is almost complete.[2] Then the next result also shows that R^0 is in fact consistent with R.

Proposition 17 (Necessity) *For any ODLM (u, R), if $u(b, 2) \neq u(b, 3)$ and $\frac{u(a,1)}{u(b,2)} \neq \frac{u(a,2)}{u(b,3)}$ for all $a, b \in X$ with $a\,R\,b$, then $p_{(u,R)}$ satisfies Asymmetry and Transitivity. Moreover, for any $a, b \in X$, if aR^0b, then aRb.*

5.5 Behavioral Phenomena

5.5.1 Consistency with Violations of IIA–the Similarity Effect and Compromise Effects

The similarity and compromise effects are well-known deviations from Luce's model (see Rieskamp et al. (2006) for a survey). In this section, we demonstrate how the ODLM can capture each of these phenomena.

[2] In Appendix E.5, we discuss how to complete R^0.

The similarity and compromise effects are defined in the same kind of experimental setup. An agent makes choices from the menus $\{x, y\}$ and $\{x, y, z\}$. The "effects" relate to the consequences of adding the alternative z.

The Similarity Effect and Debreu's Red Bus/Blue Bus Example

Suppose that our three alternatives are such that x and z are somehow very similar to each other, and clearly distinct from y. This setup is discussed by Tversky (1972a), building on a well-known example of Debreu (1960a) in which the agent makes a transportation choice and x and z are a red bus and a blue bus while y is a train. Since z is more similar to x, adding z hurts x more than y. This effect is called the *similarity effect*, and can be formalized as follows:

$$\frac{p(x, \{x, y, z\})}{p(y, \{x, y, z\})} < \frac{p(x, \{x, y\})}{p(y, \{x, y\})}.$$

In Debreu's example, the agent is assumed to like bus and train equally; that is, $p(x, \{x, y\}) = p(y, \{x, y\}) = \frac{1}{2}$, but when there are two buses the probability of choosing blue bus halves; that is, $p(x, \{x, y, z\}) = \frac{1}{4}$ and $p(y, \{x, y, z\}) = \frac{1}{2}$.

Observation 6: When $y \, R \, z \, R \, x$, the ODLM is consistent with the similarity effect. In particular, when $u(x, 2) = 2$, $u(x, 3) = 1$, $u(y, 1) = 2$, then we obtain Derbeu's example:

$$\frac{p(x, \{x, y, z\})}{p(y, \{x, y, z\})} = \frac{u(x, 3)}{u(y, 1)} = \frac{1}{2} < \frac{p(x, \{x, y\})}{p(y, \{x, y\})} = \frac{u(x, 2)}{u(y, 1)} = 1.$$

The Compromise Effect

Consider again three alternatives x, y, and z. Suppose that x and z are "extreme" alternatives, while y represents a moderate middle ground, a compromise. In the experiment studied by Simonson and Tversky (1992), x is X-370, a very basic model of Minolta camera; y is MAXXUM 3000i, a more advanced model of the same brand; and z is MAXXUM 7000i, the top of the line offered by Minolta in this class of cameras.

In Experiment 1, the menu is $\{x, y\}$ and x is chosen at least as frequently as y. However, in Experiment 2, the menu is $\{x, y, z\}$ and y is chosen more often than x.[3]

[3] Essentially, it is a stochastic version of preference reversal.

Model	Price ($)	Choices Exp. 1	Choices Exp. 2
x (X-370)	169.99	50%	22%
y (MAXXUM 3000i)	239.99	50 %	57%
z (MAXXUM 7000i)	469.99	N/A	21%

Figure 5.1: The Compromise effect in Simonson and Tversky (1992)

Simonson and Tversky (1992) call this phenomenon the *compromise effect*. As in Rieskamp et al. (2006), the compromise effect can be written as follows:

$$\frac{p(x,\{x,y,z\})}{p(y,\{x,y,z\})} < 1 \leq \frac{p(x,\{x,y\})}{p(y,\{x,y\})}.$$

Simonson and Tversky (1992)'s explanation for the compromise effect is that subjects are averse to extremes, which helps the "compromise" option y when facing the problem $\{x,y,z\}$.

Observation 7: ODLM can capture the compromise effect when $yRzRx$. Moreover, we can replicate Figure 5.1 with the following numbers: $u(x,2) = u(y,1) = 1$, $u(x,3) = \frac{22}{57}$, and $u(z,2) = \frac{21}{57}$.[4]

5.5.2 Consistency with Violation of Regularity–Attraction Effect

The ODLM can also accommodate violations of regularity, another property that Luce's Model satisfy. In fact, regularity is the property that all Random Utility Models satisfy which requires that adding alternative to a menu weakly decreases the probability of choosing alternatives of the original menu. Formally,

Regularity: $p(a,A) \geq p(a,A \cup \{b\})$ for any $A \in \mathscr{A}$, and $a,b \in X$.

We focus on the well-known attraction effect (documented by Simonson and Tversky 1992) using the following experiment. Consider our three alternatives again, x, y, and z. Suppose now that y and z are different variants of the same good: y is a Cross pen (meaning a higher quality pen), while z is a pen of regular quality: y clearly dominates z. We give the alternative x a monetary value ($6). Then

Simonson and Tversky (1992) (p.287) asked subjects to choose between x and y in Experiment 1 and to choose among x, y, and z in Experiment 2. They found

[4]Then $p(x,xy) = \frac{u(x,2)}{u(x,2)+u(y,1)} = \frac{1}{1+1} = 0.5$, $p(x,xyz) = \frac{u(x,3)}{u(x,3)+u(y,1)+u(z,2)} = \frac{22/57}{22/57+1+21/57} = 0.22$, and $p(y,xyz) = \frac{u(y,1)}{u(x,3)+u(y,1)+u(z,2)} = \frac{1}{22/57+1+21/57} = 0.57$.

Option	Choices Exp. 1	Choices Exp. 2
x (\$6)	64 %	52 %
y (Cross pen)	36 %	46 %
z (Other pen)	N/A	2 %

Figure 5.2: Attraction effect in Simonson and Tversky (1992)

that in Experiment 2, the share of subjects who chose y becomes higher than that in Experiment 1. This effect is called the *a*ttraction effect. As in Rieskamp et al. (2006), the effect can be described as follows:

$$p(y, xyz) > p(y, xy).$$

Observation 8: The ODLM can capture the attraction effect when $yRzRx$. Moreover, we can replicate Figure 5.2 with numbers $u(x,2) = 64, u(y,1) = 36$, $u(x,3) = 40\frac{16}{23}$, and $u(z,2) = \frac{36}{23}$.[5]

5.5.3 Violation of Stochastic Transitivity

The ODLM also allows for violations of weak stochastic transitivity. Formally, weak stochastic transitivity is defined as follows.

Weak Stochastic Transitivity: For any $x, y, z \in X$, if $p(x, xy) \geq \frac{1}{2}$ and $p(y, yz) \geq \frac{1}{2}$, then $p(x, xz) \geq \frac{1}{2}$.

Violations of transitivity are consistently observed in lab experiments. For example, Figure 5.3 shows observed choice probabilities in the experiment of Tversky (1969). In the experiment, subjects were asked to choose between binary lotteries $x = (\$5, \frac{7}{24})$, $y = (\$4.5, \frac{9}{24})$, and $z = (\$4, \frac{11}{24})$. Here $(\$5, \frac{7}{24})$ denotes a binary lottery that gives \$5 with probability $\frac{7}{24}$ and gives nothing with probability $\frac{17}{24}$ and so on.[6]

Figure 5.3: Choice Probabilities in Tversky (1969)

Gambles	x and y	y and z	x and z
$x = (\$5, 7/24)$	67.5%		36%
$y = (\$4.5, 9/24)$	32.5%	65%	
$z = (\$4, 11/24)$		35%	64%

[5] Then $p(y, xy) = \frac{u(y,1)}{u(x,2)+u(y,1)} = 0.36$ and $p(y, xyz) = \frac{u(y,1)}{u(x,3)+u(y,1)+u(z,2)} = \frac{36}{40\frac{16}{23}+36+\frac{36}{23}} = 0.46$.

[6] Figure 5.3 is directly calculated from Tversky (1969)'s results. Tversky's result was replicated by Lindman and Lyons (1978), Budescu and Weiss (1987), and Day and Loomes (2010).

Observation 9: ODLM allows for violations for weak stochastic transitivity when $yRzRx$. Moreover, we can replicate Figure 5.3 with numbers $u(x,2)=1$, $u(y,1)=\frac{13}{27}$, $u(z,1)=\frac{16}{9}$.

5.5.4 Choice Overload

The *choice overload* is a scenario documented in both lab and field experiments, where an increase in the number of alternatives in menu leads to adverse consequences such as a decrease in the motivation to choose or the satisfaction with the finally chosen option (e.g., Chernev 2003 and Iyengar and Lepper 2000). One of usual explanations for the choice overload is that having too many alternatives makes it hard to choose (or find) the good alternative. Here we demonstrate that adding a new alternative into a menu may lead to a decrease in the agent's satisfaction with the his chosen option even if added alternative does not decrease the average level of utility of the menu.

We can convey the main intuition by just considering three alternatives: x, y, and z. Take an ODLM (u, R) such that $u(a,i) = w(i) \cdot u(a)$ where $w(i)$ is strictly decreasing in i. The following observation shows that adding alternative z into menu the $\{x, y\}$ decreases the expected utility of menu even if the utility of z is high enough.

Observation 10: *Suppose $xRzRy$, $u(y) > u(x)$, and the utility of z is equal to the expected utility of the menu $\{x, y\}$; that is, $u(z) = p(x,\{x,y\}) \cdot u(x) + p(y,\{x,y\}) \cdot u(y)$. If $w(3)$ is small enough, then*

$$p(x,\{x,y\}) \cdot u(x) + p(y,\{x,y\}) \cdot u(y) >$$

$$p(x,\{x,y,z\}) \cdot u(x) + p(y,\{x,y,z\}) \cdot u(y) + p(z,\{x,y,z\}) \cdot u(z).$$

Intuitively, adding z makes it harder to choose (or find) the best alternative y because y is the last alternative under the ordering R. If R is related to the search process that agents use to make consumption choices, then the intuition of Observation 10 is consistent with the usual explantation for the choice overload.

5.6 Related Literature

Section 5.5 explains how the ODLM relates to the relevant empirical findings, including the similarity, compromise, and attraction effects. We now proceed to discuss the relation between the ODLM and some of the most important theoretical models of stochastic choice.

There is a non-axiomatic literature that proposes several models which can explain the similarity, compromise, and attraction effects. Rieskamp et al. (2006) is an excellent survey. Examples are Tversky (1972a), Roe et al. (2001) and Usher and McClelland (2004). The latter two papers propose *d*ecision field theory, which allows for violations of Luce's regularity axiom. The recent work by Natenzon (2010) presents a learning model, in which an agent learns about the utility of the different alternatives randomly and makes a choice with imperfect knowledge of these utilities. Natenzon's model can explain all three effects. We shall not discuss these papers here, and focus instead on the more narrowly related axiomatic literature in economics. We separate literature in three categories.

1. Random Utility Models: The benchmark economic model of rational behavior for stochastic choice is the random utility model. Since Luce's model is a special case of both the ODLM and random utility, the ODLM and random utility intersect. However, the ODLM allows for the attraction effect while random utility models always satisfy regularity, so the ODLM is not a special case of random utility.

The recent paper by Gul et al. (2010) presents a random utility model in which object attributes play a key role as in Tversky (1972a). Their model has Luce's form, but it applies sequentially, and in terms of its empirical motivation, it seeks to address the similarity effect.

2. Models with Bounded Rational Agents: A closely related paper is Echenique et al. (2013). In this paper, an order on alternatives also matters for random choice, and the model can explain the attraction and compromise effects, as well as violations of stochastic transitivity. In their paper, the source of violations of IIA is limited perception while the utilities are menu-independent.

Manzini and Mariotti (2014) study a stochastic choice model where *attention* is the source of randomness in choice while preferences are deterministic. Their model can explain the similarity and compromise effects as well as violations of stochastic transitivity.

The paper by Fudenberg et al. (2015) considers a decision maker who chooses a probability distribution over alternatives so as to maximize expected utility, with a cost function that ensures that probabilities are non-degenerate. One version of their model can accommodate the attraction effect, and one can accommodate the compromise effect.

3. Non-Stochastic Choices: Our model is more closely related to two lines of research on choice theory. Before we discuss them, note that we can easily obtain a choice theoretic version of the ODLM in the following way: for any $A \in \mathscr{A}$,

$$a \text{ is chosen from menu } A \text{ iff } u(a, R(a, A)) > u(b, R(b, A)) \text{ for each } b \in A \setminus \{a\}.$$

Similar to Section 5.3.1, we can also define choice-theoretic versions of increasing ODLMs and decreasing ODLMs.

The first line of research is on limited attention and consideration set. Masatlioglu et al. (2012) attribute violations of WARP (the counterpart of IIA in deterministic choice models) to the role of attention. They elicit revealed preference in the following way: when the choice from $\{x, y, z\}$ is x and from $\{x, z\}$ is z, then they conclude that x is revealed preferred to z. In contrast, we conclude that x has a higher ranking than z in decreasing ODLMs (opposite of Observation 5). In fact, a choice-theoretic version of decreasing ODLMs is a special case of Masatlioglu et al. (2012). However, choice-theoretic versions of increasing ODLMs are not special cases of Masatlioglu et al. (2012).

The second line of research is on framing effects. In particular, Rubinstein et al. (2006) and Salant and Rubinstein (2008) discuss the effect of different frames (e.g., different rankings over alternatives) while in our paper the ordering is fixed. A more closely related paper is Yildiz (2012) which also discusses fixed ordering on alternatives. Since Yildiz focuses on choices in which the choice procedure is also influenced by the ordering on alternatives and an agent engage in some kind of sequential search, he obtains a very different model from ours. In particular, a random choice version of his model cannot have Luce's Model as a special case because of the sequentiality.

Besides these two lines of research, there are several recent papers on choice theory that explain behavioral phenomena in Section 5.5. For example, Kamenica (2008) discusses model of context-dependent preferences and explains the attraction and compromise effects as well as the choice overload; Ok et al. (2014) discusses model of (endogenous) reference-dependent preferences and explains the attraction effect; De Clippel and Eliaz (2012)'s model produces the compromise and attraction effects as solutions of some bargaining problems.

Appendix A

APPENDIX TO CHAPTER 1

A.1 Proofs

A.1.1 Proofs of Theorem 1 and Corollary 1

Proof of Theorem 1: Take any lottery X and its two outcomes x_i, x_j with $x_i > x_j$. I shall prove that $V_{(x_i, X)}(Z) \geq V_{(x_j, X)}(Z)$ for any $Z \in \Delta(\mathbb{R}_+)$. Let me consider the following special case of dynamic monotonicity.

First define the following new lotteries: for any $\epsilon \in (0, p_j)$, $X_\epsilon \equiv (p_i + \epsilon, x_i, p_j - \epsilon, x_j, X_{-i,-j})$. Since X_ϵ first-order stochastically dominates X, I must have:

Weak Dynamic Monotonicity. $W(X_\epsilon; Z) \geq W(X; Z)$ for any $Z \in \Delta(\mathbb{R}_+)$.

In terms of (1.3), the above is equivalent to the inequality:

$$V_0(X_\epsilon) + \beta\left((p_i + \epsilon)\, V_{(x_i, X_\epsilon)}(Z) + (p_j - \epsilon)\, V_{(x_j, X_\epsilon)}(Z) + \sum_{s \neq i,j} p_s\, V_{(x_s, X_\epsilon)}(Z)\right) \geq$$

$$V_0(X) + \beta\left(p_i\, V_{(x_i, X)}(Z) + p_j\, V_{(x_j, X)}(Z) + \sum_{s \neq i, j} p_s\, V_{(x_s, X)}(Z)\right);$$

equivalently,

$$\frac{V_0(X_\epsilon) - V_0(X)}{\beta} + \epsilon\left(V_{(x_i, X)}(Z) - V_{(x_j, X)}(Z)\right) + (p_i + \epsilon)\left(V_{(x_i, X_\epsilon)}(Z) - V_{(x_i, X)}(Z)\right) +$$

$$(p_j - \epsilon)\left(V_{(x_j, X_\epsilon)}(Z) - V_{(x_j, X)}(Z)\right) + \sum_{s \neq i, j} p_s\left(V_{(x_s, X_\epsilon)}(Z) - V_{(x_s, X)}(Z)\right) \geq 0.$$

By right-continuity, $V_{(x_s, X_\epsilon)} \xrightarrow{\text{U}} V_{(x_s, X)}$ for each s. That is, for any $\delta > 0$, there exists $\epsilon^* > 0$ such that for any $\epsilon \in (0, \epsilon^*)$,

$$|V_{(x_s, X_\epsilon)}(Z) - V_{(x_s, X)}(Z)| < \delta \text{ for any } Z \in \Delta(\mathbb{R}_+).$$

Fix any $\delta > 0$. Therefore, for any s, there exists $\epsilon_s(\delta) > 0$ such that for any $\epsilon < \epsilon_s(\delta)$,

$$|V_{(x_s, X_\epsilon)}(Z) - V_{(x_s, X)}(Z)| < \delta \text{ for any } Z \in \Delta(\mathbb{R}_+).$$

Let $\epsilon^*(\delta) = \min\{\epsilon_1(\delta), \ldots, \epsilon_n(\delta)\}$. Then for any $\epsilon < \epsilon^*(\delta)$ and s,

$$|V_{(x_s, X_\epsilon)}(Z) - V_{(x_s, X)}(Z)| < \delta \text{ for any } Z \in \Delta(\mathbb{R}_+).$$

Fix any $\epsilon \in (0, \epsilon^*(\delta))$. Then the inequality for weak dynamic monotonicity implies that

$$0 \leq \frac{V_0(X_\epsilon) - V_0(X)}{\beta} + \epsilon \left(V_{(x_i, X)}(Z) - V_{(x_j, X)}(Z)\right) + (p_i + \epsilon)\left(V_{(x_i, X_\epsilon)}(Z) - V_{(x_i, X)}(Z)\right) +$$

$$(p_j - \epsilon)\left(V_{(x_j, X_\epsilon)}(Z) - V_{(x_j, X)}(Z)\right) + \sum_{s \neq i, j} p_s \left(V_{(x_s, X_\epsilon)}(Z) - V_{(x_s, X)}(Z)\right) <$$

$$< \frac{V_0(X_\epsilon) - V_0(X)}{\beta} + \epsilon \left(V_{(x_i, X)}(Z) - V_{(x_j, X)}(Z)\right) + \delta.$$

Now I shall prove the Theorem 1. By way of contradiction, suppose there exists $Z \in \Delta(\mathbb{R}_+)$ such that $V_{(x_i, X)}(Z) < V_{(x_j, X)}(Z)$.

By nontriviality, there exists $Z^* \in \Delta(\mathbb{R}_+)$ such that

$$V_{(x_j, X)}(Z^*) - V_{(x_i, X)}(Z^*) > \left|\frac{V_0(X_\epsilon) - V_0(X) + \beta\,\delta}{\epsilon\,\beta}\right|.$$

Therefore, I obtain a violation of weak dynamic monotonicity since

$$0 < \frac{V_0(X_\epsilon) - V_0(X)}{\beta} + \epsilon \left(V_{(x_i, X)}(Z^*) - V_{(x_j, X)}(Z^*)\right) + \delta <$$

$$< \frac{V_0(X_\epsilon) - V_0(X)}{\beta} + \delta - \left|\frac{V_0(X_\epsilon) - V_0(X) + \beta\,\delta}{\beta}\right| \leq 0.$$

■

Proof of Corollary 1: For (1.6), Theorem 1 proves that when $x > x'$,

$$u_x^{-1}(\mathbb{E}[u_x(Z)]) \geq u_{x'}^{-1}(\mathbb{E}[u_{x'}(Z)]) \text{ any } Z \in \Delta(\mathbb{R}_+).$$

Now with the Jensen's Inequality, the above inequality implies that $u_{x'}$ is more concave than u_x. To illustrate, let $f_{(x,x')} \equiv u_{x'} \circ u_x^{-1}$ and $t_k = u_x(z_k)$ for each k. Then the above inequality is equivalent to

$$f_{x, x'}(\mathbb{E}[T]) \geq \mathbb{E}[f_{x, x'}(T)] \text{ any } T \in \Delta(\mathbb{R}_+).$$

By the Jensen's inequality, in order to satisfy the above inequality for any $T \in \Delta(\mathbb{R}_+)$, $f_{x, x'}$ must be concave.

■

A.1.2 Proofs of Theorems 2-3

Proof of Theorem 2 Suppose a continuous preference relation $\succeq$ on $\mathscr{L}$ is represented by a history-dependent model $\{V_{x,X}\}$ and satisfies time consistency, the axiom for history independence, and strong independence. I prove Theorem 2 with two steps.

Step 1. $V_0(Z) = V_{(x,X)}(Z)$ for any $X, Z \in \Delta(\mathbb{R}_+)$.

First, let $u(z) \equiv V_0(z)$ for any $z \in Z$. Take any $X, Z \in \Delta(\mathbb{R}_+)$ and let $\mu \equiv u^{-1}(V_0(Z))$. By (1.3), $(Z;0) \sim (\mu;0)$. By the axiom for history independence, $(Z;0) \sim (\mu;0)$ implies

$$\big(p_i, (x_i; Z), (X_{-i}; 0)\big) \sim \big(p_i, (x_i; \mu), (X_{-i}; 0)\big);$$

equivalently, by (1.3), $V_{(x_i,X)}(Z) = V_{(x_i,X)}(\mu)$. By time consistency and the definition of μ, $V_{(x_i,X)}(Z) = V_{(x_i,X)}(\mu) = V_0(\mu) = V_0(Z)$.

Step 2. $V_0(Z) = \mathbb{E}u(Z)$.

Take any $X, Y \in \Delta(\mathbb{R}_+)$. Let $z \equiv u^{-1}(\frac{V_0(X)}{\beta})$ and $t \equiv u^{-1}(\frac{V_0(Y)}{\beta})$. By (1.3), I have $(X;0) \sim (0;z)$ and $(Y;0) \sim (0;t)$. By strong independence,

$$(\alpha X + (1-\alpha)Y; 0) \sim (\alpha, (0;z), 1-\alpha, (0;t)) \text{ for any } \alpha \in [0,1];$$

equivalently, by (1.3) and time consistency,

$$V_0(\alpha X + (1-\alpha)Y) = \beta\left(\alpha\, u(z) + (1-\alpha)\, u(t)\right) \text{ for any } \alpha \in [0,1].$$

By the definitions of z and t, I have

$$V_0(\alpha X + (1-\alpha)Y) = \alpha\, V_0(X) + (1-\alpha)\, V_0(Y) \text{ for any } \alpha \in [0,1].$$

Therefore, for any $Z = (r_1, z_1, \ldots, r_m, z_m)$,

$$V_0(Z) = V_0(\sum_{k=1}^{m} r_k\, z_k) = \sum_{k=1}^{m} r_k V_0(z_k) = \mathbb{E}u(Z).$$

■

Proof of Theorem 3 Suppose $\succeq$ on $\mathscr{L}$ satisfies all three axioms. I prove Theorem 3 with the following four steps.

Step 1: First, let me construct V_0 and $V_{(x_i,\,X)}$.

First, let me define a function $\mu_0 : \Delta(\mathbb{R}_+) \to \mathbb{R}_+$. Take any $X \in \Delta(\mathbb{R}_+)$. Then let $\mu_0(X) \equiv \overline{x}$ whenever $(\overline{x}; 0) \sim (X; 0)$. Indeed, μ_0 is well-defined and $\mu_0(t) = t$ for any $t \in \mathbb{R}_+$. Now let $V_0(X) \equiv u(\mu_0(X))$. Second, let me define a function $\mu_{(x_i,\,X)}$ for any history (x_i, X). Take any history (x_i, X) and $Z \in \Delta(\mathbb{R}_+)$. Then let $\mu_{(x_i,\,X)}(Z) \equiv \overline{z}$ whenever

$$(p_i, (x_i; Z), (X_{-i}; 0)) \sim (p_i, (x_i; \overline{z}), (X_{-i}; 0)).$$

Indeed, $\mu_{(x_i,\,X)}$ is well-defined and $\mu_{(x_i,\,X)}(t) = t$ for any $t \in \mathbb{R}_+$. Let $V_{(x_i,\,X)}(Z) \equiv u(\mu_{(x_i,\,X)}(Z))$. By the above construction, I have $V_0(t) = V_{(x_i,\,X)}(t) = u(t)$ for any $t \in \mathbb{R}_+$.

Step 2: For any $(p_i, (x_i; Y_i))_{i=1}^n \in \mathscr{L}$,

$$(p_i, (x_i; Y_i))_{i=1}^n \sim \Big(X; U_2^{-1}\big(\sum_{i=1}^n p_i\, U_2(\mu_{(x_i,\,X)}(Y_i))\big)\Big).$$

Take any $L = (p_i, (x_i; Y_i))_{i=1}^n \in \mathscr{L}$. By Separability Between Parallel Histories,

since $(p_i, (x_i; Y_i), (X_{-i}; 0)) \sim (p_i, (x_i; \mu_{(x_i,\,X)}(Y_i)), (X_{-i}; 0))$ for each i,

I have

$$(p_i, (x_i; Y_i))_{i=1}^n \sim (p_i, (x_i; \mu_{(x_i,\,X)}(Y_i))_{i=1}^n).$$

Let $z_i \equiv \mu_{(x_i,\,X)}(Y_i)$. Then $(p_i, (x_i; Y_i))_{i=1}^n \sim (p_i, (x_i; z_i))_{i=1}^n$.

Let me find $\overline{z} \in \mathbb{R}_+$ such that $(p_i, (0; z_i))_{i=1}^n \sim (0; \overline{z})$. Then, by Regularity 4, $\overline{z} = U_2^{-1}\big(\sum_{i=1}^n p_i\, U_2(z_i)\big)$. By Weak Separability Between Today and Tomorrow ii),

$$(p_i, (x_i; z_i))_{i=1}^n \sim (X; \overline{z}) = \Big(X; U_2^{-1}\big(\sum_{i=1}^n p_i\, U_2(z_i)\big)\Big) = \Big(X; U_2^{-1}\big(\sum_{i=1}^n p_i\, U_2(\mu_{(x_i,\,X)}(Y_i))\big)\Big).$$

Step 3: For any $L = (p_i, (x_i; Y_i))_{i=1}^n, L' = (p'_k, (x'_k; Y'_k))_{k=1}^m \in \mathscr{L}$,

$$L \succeq L' \text{ iff } u(\mu_0(X)) + \beta\, u\Big(U_2^{-1}\big(\sum_{i=1}^n p_i\, U_2(\mu_{(x_i,\,X)}(Y_i))\big)\Big) \geq$$

$$\geq u(\mu_0(X')) + \beta\, u\Big(U_2^{-1}\big(\sum_{k=1}^m p'_k\, U_2(\mu_{(x'_k,\,X')}(Y'_k))\big)\Big).$$

By Weak Separability Between Today and Tomorrow i),

$$(X;\overline{z}) \sim (\mu_0(X);\overline{z}) \text{ since } (X;0) \sim (\mu_0(X);0).$$

By transitivity,

$$(p_i,(x_i;Y_i))_{i=1}^n \sim (X;\overline{z}) \sim (\mu_0(X);\overline{z}) = \big(\mu_0(X);U_2^{-1}\big(\sum_{i=1}^n p_i\,U_2(z_i)\big)\big) =$$

$$= \big(\mu_0(X);U_2^{-1}\big(\sum_{i=1}^n p_i\,U_2(\mu_{(x_i,X)}(Y_i))\big)\big).$$

Then by Regularity 3 and transitivity,

$$(p_i,(x_i;Y_i))_{i=1}^n \succsim (p'_k,(x'_k;Y'_k))_{k=1}^m \text{ iff}$$

$$\Big(\mu_0(X);U_2^{-1}\big(\sum_{i=1}^n p_i\,U_2(\mu_{(x_i,X)}(Y_i))\big)\Big) \succsim \Big(\mu_0(X');U_2^{-1}\big(\sum_{k=1}^m p'_k\,U_2(\mu_{(x'_k,X')}(Y'_k))\big)\Big) \text{ iff}$$

$$u(\mu_0(X))+\beta\,u\Big(U_2^{-1}\big(\sum_{i=1}^n p_i\,U_2(\mu_{(x_i,X)}(Y_i))\big)\Big) \geq u(\mu_0(X'))+\beta\,u\Big(U_2^{-1}\big(\sum_{k=1}^m p'_k\,U_2(\mu_{(x'_k,X')}(Y'_k))\big)\Big).$$

Step 4: Without loss of generality, suppose $u(0) = U_2(0) = 0$ and $u(1) = U_2(1) = 1$. Then $U_2 = u$.

Take any $(p,(0;z),1-p,(0;z')) \in \mathscr{L}$. Let me find $\lambda,\mu \in \mathbb{R}_+$ such that

$$(p,(0;z),1-p,(0;0)) \sim (\lambda;0) \text{ and } (p,(0;0),1-p,(0;z')) \sim (0;\mu).$$

By Step 3, I have

$$u(\lambda) = \beta\,u\Big(U_2^{-1}(p\,U_2(z))\Big) \text{ and } \beta\,u(\mu) = \beta\,u\Big(U_2^{-1}((1-p)\,U_2(z'))\Big).$$

By Additivity, I have $(p,(0;z),1-p,(0;z')) \sim (\lambda;\mu)$; equivalently,

$$\beta\,u\Big(U_2^{-1}(p\,U_2(z)+(1-p)\,U_2(z'))\Big) = u(\lambda)+\beta\,u(\mu) =$$

$$= \beta\,u\Big(U_2^{-1}(p\,U_2(z))\Big)+\beta\,u\Big(U_2^{-1}((1-p)\,U_2(z'))\Big).$$

Let $A \equiv p\,U_2(z)$ and $B \equiv (1-p)\,U_2(z')$. Therefore, I obtain

$$u(U_2^{-1}(A+B)) = u(U_2^{-1}(A)) + u(U_2^{-1}(B)) \text{ for any } A,B \in \mathbb{R}_+.$$

Since the above is a typical Cauchy functional equation for $u \circ U_2^{-1}$, I know that $u \circ U_2^{-1}$ is linear. Therefore, $u = U_2$.

Uniqueness: Since $u = U_2$, I have

$$(r_k, (0; z_k))_{k=1}^{m} \geq (r'_k, (0; z'_k))_{k=1}^{m'} \text{ iff } \mathbb{E}\, u(Z) \geq \mathbb{E}\, u(Z').$$

It is well known that, u that satisfies the above is unique up to the normalization $u(0) = 0$ and $u(1) = 1$. Now it is sufficient to prove that functions μ_0 and $\mu_{(x_i, X)}$ are unique. Recall Step 1. By deterministic monotonicity, μ_0 is unique since for any X, there exists unique $\overline{x}$ such that $(\overline{x}; 0) \sim (X; 0)$. Moreover, by deterministic monotonicity, $\mu_{(x_i, X)}$ is unique since for any Z, there exists unique $\overline{z}$ such that

$$\left(p_i, (x_i; \overline{z}), (X_{-i}; 0)\right) \sim \left(p_i, (x_i; Z), (X_{-i}; 0)\right).$$

■

A.2 Behavioral Foundations of HDEU and HDDA

A.2.1 Characterizing HDEU

I characterize HDEU (1.13) with additional three axioms.

Axiom 24 (Expected Utility) *A preference relation $\succeq$ on $\mathscr{L}$ satisfies the following two conditions.*

1. *(EU at Period 1) There exists a utility function U_1 such that for any $X, X' \in \Delta(\mathbb{R}_+)$,*
$$(X; 0) \succeq (X'; 0) \Leftrightarrow \mathbb{E}U_1(X) \geq \mathbb{E}U_1(X), \tag{A.1}$$
2. *(EU at Period 2) For any history (x_i, X), there exists a utility function $U_{(x_i, X)}$ such that for any $Z, Z' \in \Delta(\mathbb{R}_+)$,*
$$\left(p_i, (x_i; Z), (X_{-i}; 0)\right) \succeq \left(p_i, (x_i; Z'), (X_{-i}; 0)\right) \text{ iff } \mathbb{E}U_{(x_i, X)}(Z) \geq \mathbb{E}U_{(x_i, X)}(Z'). \tag{A.2}$$

This axiom assumes that $\succeq$ has an expected utility representation in static case. The first part of Axiom 24 requires that $\succeq$ has an expected utility representation when it compares lotteries in the first period. The second part of Axiom 24 requires that

$\succeq$ has an expected utility representation when it compares lotteries in the second period. I allow that utilities for difference periods and different histories can be different.

The second axiom is called *Additivity Today* which exploits additive structure of EU in the first period. It states that if receiving z today with probability r is equivalent to receiving $\overline{y}$ tomorrow and receiving z' today with probability r' is equivalent to receiving $\overline{x}$ today, then receiving $(r, z, r', z', 1-r-r', 0)$ is equivalent to receiving $(\overline{x}; \overline{y})$. More formally,

Axiom 25 (Additivity Today) *For any* $(r, z, r', z', 1-r-r', 0) \in \Delta(\mathbb{R}_+)$ *and* $\overline{x}, \overline{y} \in \mathbb{R}_+$,

$$\begin{aligned} \textit{if } ((r, z, 1-r, 0); 0) &\sim (0; \overline{y}) \textit{ and} \\ ((1-r', 0, r', z'); 0) &\sim (\overline{x}; 0), \textit{ then} \\ ((r, z, r', z', 1-r-r', 0); 0) &\sim (\overline{x}; \overline{y}). \end{aligned}$$

The third axiom is called Linearity. It states that if receiving an outcome x after a history (x_i, X) is equivalent to receiving a lottery $(p, x, 1-p, 0)$ after a history (y_j, Y), then replacing x with y does not change the equivalence; that is, if receiving an outcome y after the history (x_i, X) is equivalent to receiving a lottery $(p, y, 1-p, 0)$ after the history (y_j, Y). More formally,

Axiom 26 (Linearity) *For any* $X, Y \in \Delta(\mathbb{R}_+)$ *with* $(X; 0) \sim (Y; 0)$, $z, z' > 0$ *and* $r \in (0, 1]$,

$$\textit{if } (p_i, (x_i; z), (X_{-i}; 0)) \sim (q_j, (y_j; (r, z, 1-r, 0)), (Y_{-j}; 0)),$$

$$\textit{then } (p_i, (x_i; z'), (X_{-i}; 0)) \sim (q_j, (y_j; (r, z', 1-r, 0)), (Y_{-j}; 0)).$$

Under Axioms 3 and 24, there are functions u, U_1, and $U_{(x_i, X)}$ for each history (x_i, X). Without loss of generality, I assume $u(0) = U_1(0) = U_{(x_i, X)}(0) = 0$ and $u(1) = U_1(1) = U_{(x_i, X)}(1) = 1$. Now I can state the characterization theorem for HDEU (1.13).

Theorem 8 (HDEU) *Take any preference* $\succeq$ *on* $\mathscr{L}$ *that satisfies (1.12). If it satisfies Axiom 24 (Expected Utility), Additivity Today, and Linearity, then* $u = U_1$ *and for any history* (x, X), $U_{(x, X)} = u^{\mu(x, X)}$ *for some* $\mu(x, X) > 0$.

A.2.2 Characterizing HDDA

I also characterize HDDA (1.15) with additional two axioms.

Axiom 27 (Axioms of Gul's Disappointment Aversion) *A preference relation* $\succeq$ *on* $\mathscr{L}$ *satisfies the following two conditions.*

1. *(Disappointment Aversion at Period 1) There exists a pair* (U_0, δ_0) *such that for any* $X, X' \in \Delta(\mathbb{R}_+)$,

$$(X; 0) \succeq (X'; 0) \Leftrightarrow \mu_0(X|\delta_1) \geq \mu_0(X'|\delta_1), \tag{A.3}$$

2. *(Disappointment Aversion at Period 2) For any history* (x_i, X)*, there exists a pair* $(U_{(x_i, X)}, \delta(x_i, X))$ *such that for any* $Z, Z' \in \Delta(\mathbb{R}_+)$,

$$(p_i, (x_i; Z), (X_{-i}; 0)) \succeq (p_i, (x_i; Z'), (X_{-i}; 0)) \textit{ iff}$$

$$\mu_{(x_i, X)}(Z|\delta(x_i, X)) \geq \mu_{(x_i, X)}(Z'|\delta(x_i, X)). \tag{A.4}$$

This axiom assumes that $\succeq$ has a Gul's Disappointment Aversion representation in the static case. The first part of Axiom 27 requires that $\succeq$ has a Gul's Disappointment Aversion representation when it compares lotteries in the first period. The second part of Axiom 27 requires that $\succeq$ has a Gul's Disappointment Aversion representation when it compares lotteries in the second period. I allow that utilities and disappointment parameters for different periods and different histories can be different.

The second axiom is a weakening of Additivity Today, I call it *Weak Additivity*. First, let me rewrite Additivity Today in the following way: when $r + r' \leq 1$ and $s + s' \leq 1$,

$$\begin{aligned} \text{if } (r, z, 1-r, 0) &\succeq (s, t, 1-s, 0) \text{ and} \\ (1-r', 0, r', z') &\succeq (1-s', 0, s', t'), \text{ then} \\ (r, z, r', z', 1-r-r', 0) &\succeq (s, t, s', t', 1-s-s', 0). \end{aligned}$$

However, consistent violations of Additivity Today are documented in lab experiments. One well-known violation of Additivity Today is the Common Consequence Effect, a version of the Allais Paradox. For example, Kahneman and Tversky (1979) found the following instance of the Common Consequence Effect:

$$\begin{aligned}(0.33, \$27, 0.67, \$0) &\succ (0.34, \$24, 0.66, \$0) \text{ and} \\ (0.34, \$0, 0.66, \$24) &= (0.34, \$0, 0.66, \$24), \text{ but} \\ (0.33, \$27, 0.66, \$24, 0.01, \$0) &\prec (1, \$24).\end{aligned}$$

The main reason of the above violation is that the lottery $(0.34, \$0, 0.66, \$24)$ completely complements $(0.34, \$24, 0.66, \$0)$. As a result, the agent has to compare a risky lottery $(0.33, \$27, 0.66, \$24, 0.01, \&0)$ with a deterministic lottery $(1, \$24)$, but her behavior will be different from the case when she compares two risky lotteries $(0.33, \$27, 0.67, \$0)$ and $(0.34, \$24, 0.66, \$0)$. Therefore, I modify Additivity Today in the following way: the additive property should hold when i) there is no mixture between risky and deterministic lotteries (e.g., $r + r' = 1$ and $s + s' = 1$); and ii) lotteries do not complement each other (e.g., z' is small enough compared to z and t' is small enough compared to t). More formally, when $(r, z, 1-r, 0) \succeq z'$ and $(s, t, 1-s, 0) \succeq t'$,

$$\begin{aligned}\text{if } (r, z, 1-r, 0) &\succeq (s, t, 1-s, 0) \text{ and} \\ (r, 0, 1-r, z') &\succeq (s, 0, 1-s, t'), \text{ then} \\ (r, z, 1-r, z') &\succeq (s, t, 1-s, t').\end{aligned}$$

I require that the new additivity property holds for lotteries in the first period and also for lotteries in the second period. Formally,

Axiom 28 (Weak Additivity) *Take any* $(r, z, 1-r, z') \in \Delta(\mathbb{R}_+)$ *with* $(r, z, 1-r, 0) \succeq z'$.

1. *(Weak Additivity Today) For any* $\overline{x}, \overline{y} \in \mathbb{R}_+$,

$$\begin{aligned}\textit{if } ((r, z, 1-r, 0); 0) &\sim (0; \overline{y}) \textit{ and} \\ ((r, 0, 1-r, z'); 0) &\sim (\overline{x}; 0), \textit{ then} \\ ((r, z, 1-r, z'); 0) &\sim (\overline{x}; \overline{y}).\end{aligned}$$

2. *(Weak Additivity Tomorrow) For any* $\overline{x}, \overline{y} \in \mathbb{R}_+$ *and history* (x_i, X),

$$\begin{aligned}\textit{if } (p_i, (x_i; (r, z, 1-r, 0)), (X_{-i}; 0)) &\sim (X; \overline{y}) \textit{ and} \\ (p_i, (x_i; (r, 0, 1-r, z')), (X_{-i}; 0)) &\sim (\overline{x}; 0), \textit{ then} \\ (p_i, (x_i; (r, z, 1-r, z')), (X_{-i}; 0)) &\sim (\overline{x}; \overline{y}).\end{aligned}$$

Under Axioms 3 and 27, there are (u, β), (U_0, δ_0), and $(U_{(x_i, X)}, \delta(x_i, X))$ for any history (x_i, X). Without loss of generality, I assume $u(0) = U_1(0) = U_{(x_i, X)}(0) = 0$ and $u(1) = U_1(1) = U_H(1) = 1$. Now I can state the characterization theorem for HDDA (1.15).

Theorem 9 *Take any preference $\succeq$ on $\mathscr{L}$ that satisfies (1.12). If it satisfies Axioms of Gul's Disappointment Aversion, and Weak Additivity, then $u = U_0 = U_{(x,X)}$ for any history (x, X); that is, for any $X, Z \in \Delta(\mathbb{R}_+)$,*

$$\mu_0(X|\delta_0) \text{ is solution to } u(\mu) = \frac{\sum_{i=1}^n p_i\, u(x_i)(1 + \delta_0\ \mathbb{1}\{x_i \le \mu\})}{\sum_{i=1}^n p_i\ (1 + \delta_0\ \mathbb{1}\{x_i \le \mu\})} \text{ and}$$

$$\mu_{(x,X)}(Z|\delta(x, X)) \text{ is solution to } u(\mu) = \frac{\sum_{k=1}^m r_k\, u(z_k)(1 + \delta(x, X)\ \mathbb{1}\{z_k \le \mu\})}{\sum_{i=1}^n r_k\ (1 + \delta(x, X)\ \mathbb{1}\{z_k \le \mu\})}.$$

Proofs of Theorems 8-9

Proof of Theorem 8 I prove Theorem 8 with three steps. Recall Step 1 of the proof of Theorem 3. There are functions μ_0 and $\{\mu_{(x,X)}\}$ such that for any $L = (p_i, (x_i; Y_i))_{i=1}^n, L' = (p'_k, (x'_k; Y'_k))_{k=1}^m \in \mathscr{L}$,

$$L \succeq L' \text{ iff } u(\mu_0(X)) + \beta \sum_{i=1}^n p_i\, u(\mu_{(x_i, X)}(Y_i)) \ge u(\mu_0(X')) + \beta \sum_{k=1}^m p'_k\, u(\mu_{(x'_k, X')}(Y'_k)). \tag{A.5}$$

Step 1. By Axiom 24, for any X and Z,

$$\mu_0(X) = U_1^{-1}(\mathbb{E}\, U_1(X)) \text{ and } \mu_H(Z) = U_H^{-1}(\mathbb{E}\, U_H(Z)) \text{ where } H = (x_i, X).$$

Take any X and $\overline{x}$ such that $(X; 0) \sim (\overline{x}; 0)$. By (A.5), I have $\mu_0(X) = \overline{x}$. By Axiom 24, I then obtain the first equation of Step 1. Moreover, take any $H = (x_i, X)$, Z, and and $\overline{z}$ such that $(p_i, (x_i; Z), (X_{-i}; 0)) \sim 0)) \sim (p_i, (x_i; \overline{z}), (X_{-i}; 0))$. By (A.5), I have $\mu_H(Z) = \overline{z}$. By Axiom 24, I then obtain the second equation of Step 1.

Step 2. By Additivity Today, $u = U_1$

Take any $(r, z, (1 - r), z') \in \Delta(\mathbb{R}_+)$ and $\overline{x}, \overline{y} \in \mathbb{R}_+$. Suppose

$$((r, z, 1 - r, 0)); 0) \sim (0; \overline{y}) \text{ and } ((r, 0, (1 - r), -z'); 0) \sim (\overline{x}; 0);$$

equivalently, by (A.5),

$$u(\mu_0(r, z, 1 - r, 0)) = \beta\, u(\overline{y}) \text{ and } u(\mu_0(r, 0, 1 - r, z')) = u(\overline{x}).$$

By Additivity Today,

$$((r, z, 1-r, -z'); 0) \sim (\overline{x}; \overline{y});$$

equivalently, by (A.5),

$$u(\mu_0(r, z, 1-r, z')) = u(\overline{x}) + \beta\, u(\overline{y}).$$

Therefore, from the above three equalities, I obtain

$$u(\mu_0(r, z, (1-r), 0)) + u(\mu_0(r, 0, 1-r, z')) = u(\mu_0(r, z, 1-r, z')); \tag{A.6}$$

equivalently

$$u(U_1^{-1}(rU_1(z))) + u(U_1^{-1}((1-r)U_1(z'))) = u(U_1^{-1}(rU_1(z) + (1-r)U_1(z'))). \tag{A.7}$$

Let $A = rU_1(z)$ and $B = (1-r)U_1(z')$. Therefore, I have

$$u(U_1^{-1}(A)) + u(U_1^{-1}(B)) = u(U_1^{-1}(A+B)) \text{ for any } A, B \geq 0.$$

Since the above is a typical Cauchy functional equation, $u \circ U_1^{-1}$ is a linear function. Since $u(1) = U_1(1) = 1$, I have $u = U_1$.

Step 3: For any history H, by Linearity, $U_H = u^{\mu(H)}$ for some $\mu(H) > 0$.

Take any $X, Y \in \Delta(\mathbb{R}_+)$ with $(X; 0) \sim (Y; 0)$ and $(p_i, (x_i; x), (X_{-i}; 0)) \sim (q_j, (y_j; (p, x, 1-p, 0)), (Y_{-j}; 0))$ for some $x > 0$ and $p > 0$. By (A.5),

$$\mathbb{E}\, u(X) = \mathbb{E}\, u(Y) \text{ and}$$

$$\mathbb{E}\, u(X) + \beta\, p_i\, u(x) = \mathbb{E}\, u(Y) + \beta\, q_j\, u(U_H^{-1}(p\, U_H(x))) \text{ where } H = (y_j, Y).$$

From the above two equalities, I obtain

$$p_i\, u(x) = q_j\, u(U_H^{-1}(p\, U_H(x))).$$

By Linearity, I have

$$(p_i, (x_i; y), (X_{-i}; 0)) \sim (q_j, (y_j; (p, y, 1-p, 0)), (Y_{-j}; 0)) \text{ for any } y > 0.$$

Therefore,

$$p_i\, u(x) = q_j\, u(U_H^{-1}(p\, U_H(x))) \text{ iff } p_i\, u(y) = q_j\, u(U_H^{-1}(p\, U_H(y))) \text{ for any } y > 0.$$

Let $y = 1$ and $U_H(x) = t$ and $G \equiv u \circ U_H^{-1}$. Then I have

$$G(p)\, G(t) = G(p\, t) \text{ for any } t > 0 \text{ and } p \in (0, 1).$$

Since $G(1) = 1$, $G(p) = \frac{1}{G(\frac{1}{p})}$ when $p \cdot t = 1$. Therefore, I have

$$G(p)\, G(t) = G(p\, t) \text{ for any } t > 0 \text{ and } p > 0.$$

Since the above is a typical Cauchy functional equation, there exists $\alpha > 0$ such that $G(t) = t^{\alpha}$; that is, $U_H = u^{\frac{1}{\alpha}}$. Since α depends on H, let $\mu(H) \equiv \frac{1}{\alpha}$; that is, $U_H = u^{\mu(H)}$.

■

Proof of Theorem 9 I prove Theorem 9 with three steps.

Step 1. Let H be a history. By Axiom 27, for any X and Z,

$$\mu_0(X|\delta_0) \text{ is a unique solution to } U_1(\mu) = \frac{\sum_{i=1}^{n} p_i\, U_0(x_i)(1 + \delta_0\ \mathbb{1}\{x_i \leq \mu\})}{\sum_{i=1}^{n} p_i\, (1 + \delta_0\ \mathbb{1}\{x_i \leq \mu\})} \text{ and}$$

$$\mu_H(Z|\delta(H)) \text{ is a unique solution to } U_H(\mu) = \frac{\sum_{k=1}^{m} r_k\, U_H(z_k)(1 + \delta(H)\ \mathbb{1}\{z_k \leq \mu\})}{\sum_{i=1}^{n} p_i\, (1 + \delta(H)\ \mathbb{1}\{z_k \leq \mu\})}.$$

Take any X and $\overline{x}$ such that $(X; 0) \sim (\overline{x}; 0)$. By (A.5), $\mu_0(X) = \overline{x}$. By Axiom 27, I then obtain the first first equation of Step 1. Moreover, take any $H = (x_i, X)$, Z, and and $\overline{z}$ such that $(p_i, (x_i; Z), (X_{-i}; 0)) \sim (p_i, (x_i; \overline{z}), (X_{-i}; 0))$. By (A.5), $\mu_H(Z) = z$. By Axiom 27, I then obtain the second equation of Step 1.

Step 2. By Weak Additivity Today, $u = U_0$

Take any $(r, z, (1 - r), z') \in \Delta(\mathbb{R}_+)$ with $(r, z, (1 - r), 0) \geq z'$, and $\overline{x}, \overline{y} \in \mathbb{R}_+$. Suppose

$$((r, z, 1 - r, 0)); 0) \sim (0; \overline{y}) \text{ and } ((r, 0, (1 - r), -z'); 0) \sim (\overline{x}; 0);$$

equivalently, by (A.5),

$$u\big(\mu_1(r, z, 1 - r, 0)\big) = \beta\, u(\overline{y}) \text{ and } u\big(\mu_1(r, 0, 1 - r, z')\big) = u(\overline{x}).$$

By Weak Additivity Today,

$$((r, z, 1 - r, -z'); 0) \sim (\overline{x}; \overline{y});$$

equivalently, by (A.5),

$$u(\mu_0(r, z, 1-r, z'|\delta_0)) = u(\overline{x}) + \beta\, u(\overline{y}).$$

Therefore, from the above three equalities, I obtain

$$u(\mu_0(r, z, (1-r), 0)) + u(\mu_0(r, 0, 1-r, z')) = u(\mu_0(r, z, 1-r, z')). \tag{A.8}$$

Since $(r, z, (1-r), 0) \geq z'$,

$$u(U_0^{-1}(\frac{rU_0(z)}{r+(1-r)(1+\delta_0)})) + u(U_0^{-1}(\frac{(1-r)U_0(z')(1+\delta_0)}{r+(1-r)(1+\delta_0)})) = \tag{A.9}$$

$$u(U_0^{-1}(\frac{rU_0(z)+(1-r)U_0(z')(1+\delta_0)}{r+(1-r)(1+\delta_0)})).$$

Let $A = \frac{rU_0(z)}{r+(1-r)(1+\delta_0)}$ and $B = \frac{(1-r)U_0(z')(1+\delta_0)}{r+(1-r)(1+\delta_0)}$. Therefore, I have

$$u(U_0^{-1}(A)) + u(U_0^{-1}(B)) = u(U_0^{-1}(A+B)) \text{ for any } A, B \geq 0.$$

Since the above is a typical Cauchy functional equation, $u \circ U_0^{-1}$ is a linear function. Moreover, since $u(1) = U_0(1) = 1$, I have $u = U_0$.

Step 3: Let H be a history. By Weak Additivity Tomorrow, $U_H = u$.

Take any $X, (r, z, (1-r), z') \in \Delta(\mathbb{R}_+)$ with $(r, z, (1-r), 0) \geq z'$, and $\overline{x}, \overline{y} \in \mathbb{R}_+$. Let $H = (x_i, X)$. Suppose

$(p_i, (x_i; (r, z, 1-r, 0)), (X_{-i}; 0)) \sim (X; \overline{y})$ and $(p_i, (x_i; (r, 0, (1-r), -z')), (X_{-i}; 0)) \sim (\overline{x}; 0)$;

equivalently, by (A.5),

$$u(\mu_0(X)) + \beta\, p_i\, u(\mu_H(r, z, 1-r, 0)) = u(\mu_0(X)) + \beta\, u(\overline{y})$$

and

$$u(\mu_1(X)) + \beta\, p_i\, u(\mu_H(r, 0, 1-r, z')) = u(\overline{x}).$$

By Weak Additivity, I have

$$(p_i, (x_i; (r, z, 1-r, -z')), (X_{-i}; 0)) \sim (\overline{x}; \overline{y});$$

equivalently, by (A.5),

$$u(\mu_0(X)) + \beta\, p_i\, u(\mu_H(r, z, 1-r, z')) = u(\overline{x}) + \beta\, u(\overline{y}).$$

Therefore, from the above three equalities, I obtain

$$u(\mu_H(r,z,(1-r),0)) + u(\mu_H(r,0,1-r,z')) = u(\mu_H(r,z,1-r,z')). \quad \text{(A.10)}$$

Since $(r,z,(1-r),0) \geq z'$, I have

$$u(U_H^{-1}(\frac{rU_H(z)}{r+(1-r)(1+\delta_H)})) + u(U_H^{-1}(\frac{(1-r)U_{H_i}(z')(1+\delta_H)}{r+(1-r)(1+\delta_H)})) = \quad \text{(A.11)}$$

$$u(U_H^{-1}(\frac{rU_H(z)+(1-r)U_H(z')(1+\delta_H)}{r+(1-r)(1+\delta_H)})).$$

Let $A = \frac{rU_H(z)}{r+(1-r)(1+\delta_H)}$ and $B = \frac{(1-r)U_H(z')(1+\delta_H)}{r+(1-r)(1+\delta_H)}$. Therefore, I have

$$u(U_H^{-1}(A)) + u(U_H^{-1}(B)) = u(U_H^{-1}(A+B)) \text{ for any } A, B \geq 0.$$

Since the above is a typical Cauchy functional equation, $u \circ U_H^{-1}$ is a linear function. Since $u(1) = U_H(1) = 1$, I have $u = U_H$.

■

Appendix B

APPENDIX TO CHAPTER 2

B.1 Proofs

B.1.1 Proof of Proposition 2

First, let us prove that (i) is equivalent to (iii). Take any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$, and $z_2 \in \mathbb{R}_+$ such that $x_1 > y_1 > z_1$ and $x_2 < z_2' < y_2 < z_2$. Note that the alternative $\mathbf{x}$ is chosen over $\mathbf{y}$ from menu $\{\mathbf{x}, \mathbf{y}\}$ iff

$$f(u_1(x_1)-u_1(y_1))+\cancel{f(u_2(x_2)-u_2(x_2))} > \cancel{f(u_1(y_1)-u_1(y_1))}+f(u_2(y_2)-u_2(x_2)).$$

Moreover, $\mathbf{y}$ is chosen from $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ iff

$$f(u_1(x_1)-u_1(z_1))+\cancel{f(u_2(x_2)-u_2(x_2))} < f(u_1(y_1)-u_1(z_1))+f(u_2(y_2)-u_2(x_2)) \text{ and}$$

$$\cancel{f(u_1(z_1)-u_1(z_1))}+f(u_2(z_2)-u_2(x_2)) < f(u_1(y_1)-u_1(z_1))+f(u_2(y_2)-u_2(x_2)).$$

However, the second inequality is implied the first inequality since $\mathbf{x}$ is chosen over $\mathbf{z}$ from $\{\mathbf{x}, \mathbf{z}\}$ iff

$$f(u_2(z_2) - u_2(x_2)) < f(u_1(x_1) - u_1(z_1)).$$

Therefore, when z_2 is small enough, the compromise effect is equivalent to

$$f(u_1(x_1)-u_1(y_1)) > f(u_2(y_2)-u_2(x_2)) > f(u_1(x_1)-u_1(z_1))-f(u_1(y_1)-u_1(z_1)).$$

Second, let us prove that (ii) is equivalent to (iii). Note that $\mathbf{y}$ is chosen over $\mathbf{x}$ from $\{\mathbf{x}, \mathbf{y}, (z_1, z_2')\}$ iff

$$f(u_1(x_1)-u_1(z_1))+\cancel{f(u_2(x_2)-u_2(x_2))} > f(u_1(y_1)-u_1(z_1))+f(u_2(y_2)-u_2(x_2)).$$

Therefore, the attraction effect is also equivalent to

$$f(u_1(x_1)-u_1(y_1)) > f(u_2(y_2)-u_2(x_2)) > f(u_1(x_1)-u_1(z_1))-f(u_1(y_1)-u_1(z_1)).$$

B.1.2 Proof of Proposition 3

By $C_{(f,u_1,u_2)}$, $\mathbf{x} \prec_{(z_1,z_2')} \mathbf{y}$ ($\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$) is equivalent to

$$f(u_1(x_1) - u_1(z_1)) - f(u_1(y_1) - u_1(z_1)) < f(u_2(y_2) - u_2(x_2)).$$

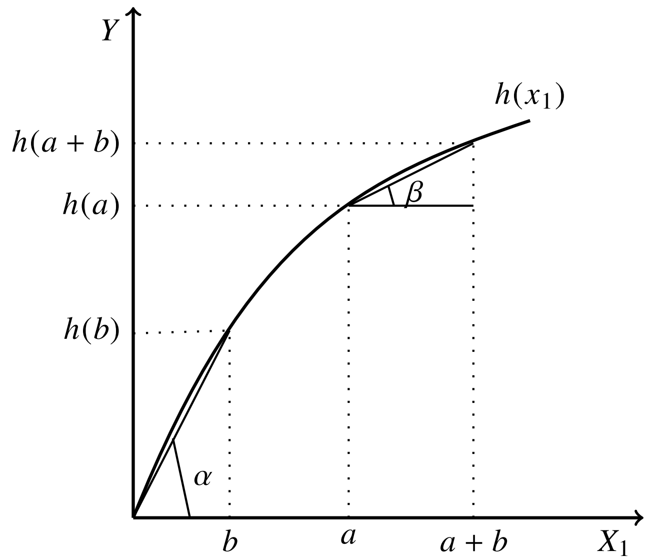

Figure B.1: Strictly Concave h

Since h is strictly increasing, we have

$$h\Big(f(u_1(x_1)-u_1(z_1))-f(u_1(y_1)-u_1(z_1))\Big) < h\Big(f(u_2(y_2)-u_2(x_2))\Big) = f'(u_2(y_2)-u_2(x_2)).$$

Then, in order to obtain $\mathbf{x} \prec'_{(z_1, z_2')} \mathbf{y}$ ($\mathbf{x} \prec'_{\mathbf{z}} \mathbf{y}$), it is enough to prove that

$$h\Big(f(u_1(x_1)-u_1(z_1))\Big)-h\Big(f(u_1(y_1)-u_1(z_1))\Big) < h\Big(f(u_1(x_1)-u_1(z_1))-f(u_1(y_1)-u_1(z_1))\Big).$$

In other words, it is enough to prove that $h(a) + h(b) > h(a + b)$ for any $a, b > 0$ when h is strictly concave.

Without loss of generality, suppose $a \geq b$. Figure B.1 illustrates the intuitive proof. By the strict concavity of h, the angle $\measuredangle\alpha$ is greater than the angle $\measuredangle\beta$. Now since the tangent function is increasing, we have $h(a + b) - h(a) = \tan(\beta)((a + b) - a) < \tan(\alpha)(b - 0) = h(b)$.

B.1.3 Proof of Proposition 4

Without loss of generality, suppose $x_1 > y_1$, $x_2 < y_2$, and $C(\{\mathbf{x}, \mathbf{y}\}) \ni \mathbf{x}$. Then we will prove that $\mathbf{y}$ is a non-extreme option of B. By way of contradiction, suppose b is an extreme option of B, so $y_1 = m_1^B$. Note that $C(\{\mathbf{x}, \mathbf{y}\}) \ni \mathbf{x}$ implies that $f(u_1(x_1) - u_1(y_1)) \geq f(u_2(y_2) - u_2(x_2))$ and $C(B) \ni \mathbf{y}$ and $C(B) \not\ni \mathbf{x}$ imply that $f(u_1(y_1) - u_1(y_1)) + f(u_2(y_2) - u_2(m_2^B)) = f(u_2(y_2) - u_2(p^B)) > f(u_1(x_1) - u_1(y_1)) + f(u_2(x_2) - u_2(m_2^B))$. Therefore, we obtain the following violation of weak diminishing sensitivity:

$$f(u_2(y_2)-u_2(m_2^B)) > f(u_1(x_1)-u_1(y_1)) + f(u_2(x_2)-u_2(m_2^B)) \geq$$

$$\geq f(u_2(y_2)-u_2(x_2)) + f(u_2(x_2)-u_2(m_2^B)).$$

B.1.4 Proof of Proposition 5

First we prove i). Take any $x_1, y_1, z_1 \in \mathbb{R}_+$ such that $x_1 > y_1 > z_1$. Since $f(+\infty) = u_1(+\infty) = u_2(+\infty) = +\infty$, there exist x_2 and y_2 such that $f(u_1(x_1) - u_1(z_1)) = f(u_1(y_1)-u_1(z_1)) + f(u_2(y_2)-u_2(x_2))$; that is, $\mathbf{x} \sim_{(z_1,\, x_2)} \mathbf{y}$. By WBAE, we have $\mathbf{x} \succ \mathbf{y}$; that is, $u_1(x_1) - u_1(y_1) > u_2(y_2) - u_2(x_2)$. Therefore, we have

$$f(u_1(x_1)-u_1(z_1)) = f(u_1(y_1)-u_1(z_1)) + f(u_2(y_2)-u_2(x_2)) <$$

$$< f(u_1(y_1)-u_1(z_1)) + f(u_1(x_1)-u_1(y_1)).$$

Second we prove ii). Take any x_1, y_1, z_1, t_1 with $x_1 > y_1 > z_1 > t_1$. Since $f(+\infty) = u_1(+\infty) = u_2(+\infty) = +\infty$, there exist x_2 and y_2 such that $f(u_1(x_1) - u_1(z_1)) - f(u_1(y_1) - u_1(z_1)) = f(u_2(y_2) - u_2(x_2))$; equivalently, $\mathbf{x} \sim_{(z_1,\, x_2)} \mathbf{y}$. By BAE, we have $\mathbf{x} \prec_{(t_1,\, x_2)} \mathbf{y}$; equivalently,

$$f(u_1(x_1)-u_1(t_1)) - f(u_1(y_1)-u_1(t_1)) < f(u_2(y_2)-u_2(x_2)) =$$

$$= f(u_1(x_1)-u_1(z_1)) - f(u_1(y_1)-u_1(z_1)).$$

B.1.5 Proof of Proposition 6

Let $A \equiv \cup_{n=1}^{N} A_n$ and $\lambda \equiv \max\{\max\{\frac{x_1-y_1}{y_2-x_2}; \frac{y_2-x_2}{x_1-y_1}\} | \mathbf{x}, \mathbf{y} \in A$ with $x_1 > y_1$ and $x_2 < y_2\}$. Moreover, let

$$f(t) \equiv \begin{cases} t & \text{if } t \geq 0 \\ -(\lambda+\epsilon)t & \text{if } t < 0, \end{cases}$$

for some $\epsilon > 0$ and $\mathbf{r}^n \equiv C(A_n)$. Then we will prove that

$$C(A_n) = \arg\max_{\mathbf{x}\in A_n} f(x_1 - r_1^n) + f(x_2 - r_2^n).$$

In other words, we shall prove that for any n and $x \in A_n \setminus C(A_n)$,

$$f(r_1^n - r_1^n) + f(r_2^n - r_2^n) = 0 > f(x_1 - r_1^n) + f(x_2 - r_2^n).$$

It is obvious when $\mathbf{r}^n > \mathbf{x}$ since $0 > f(x_1 - r_1^n)$ and $0 > f(x_2 - r_2^n)$. Moreover, by monotonicity, we cannot have $\mathbf{r}^n \leq \mathbf{x}$.

Therefore, we now consider two cases. When $x_1 \geq r_1^n$ and $x_2 < r_2^n$, $f(x_1 - r_1^i) + f(x_2 - r_2^i) = (x_1 - r_1^n) - (\lambda + \epsilon)(r_2^n - x_2) < 0$, by the definition of λ. Similarly, when $x_1 < r_1^n$ and $x_2 \geq r_2^n$, $f(x_1 - r_1^n) + f(x_2 - r_2^n) = -(\lambda + \epsilon)(r_1^n - x_1) + (y_2 - r_2^n) < 0$, by the definition of λ.

B.1.6 Implications of Regularity and INEA

We first prove the following useful implications of Regularity and INEA.

Lemma 1 *Suppose C satisfies regularity and transitivity. For any $\boldsymbol{y}, \boldsymbol{t} \in X$ and $x_j \in X_j$ with $\boldsymbol{y} > \boldsymbol{t}$ and $x_j < y_j$, there exist x_i and x_i' such that i) $(x_i, x_j) \sim \boldsymbol{y}$ and ii) $(x_i', x_j) \sim_{\boldsymbol{t}} \boldsymbol{y}$.*

Proof 1 (Proof of Lemma 1) *Since there is no essential difference between (i) and (ii), we only prove (i). Take any $\boldsymbol{y} \in X$ and $x_j \in X_j$ with $x_j < y_j$. We shall prove that there exists x_i such that $(x_i, x_j) \sim \boldsymbol{y}$. By solvability, there exists x_i^0 such that $(x_i^0, x_j) \succ \boldsymbol{y}$. Now we construct two sequences $\{x_i^n\}_{n=0}^{\infty}$ and $\{y_i^n\}_{n=0}^{\infty}$ by the induction. Let $y_i^0 \equiv y_i$. Note that $x_i^0 > y_i^0$. Suppose we have constructed $x_i^0, \ldots, x_i^k$ and $y_i^0, \ldots, y_i^k$ and we will define x_i^{k+1} and y_i^{k+1} in the following way: if $(\frac{x_i^k + y_i^k}{2}, x_j) \succeq \boldsymbol{y}$, then let $x_i^{k+1} \equiv \frac{x_i^k + y_i^k}{2}$ and $y_i^{k+1} \equiv y_i^k$, and if $(\frac{x_i^k + y_i^k}{2}, x_j) \prec \boldsymbol{y}$, then let $x_i^{k+1} \equiv x_i^k$ and $y_i^{k+1} \equiv \frac{x_i^k + y_i^k}{2}$. Since $\{x_i^n\}_{n=0}^{\infty}$ is a non-increasing sequence, $\{y_i^n\}_{n=0}^{\infty}$ is a non-decreasing sequence, and $\lim_{k\to\infty} x_i^k - y_i^k = \lim_{k\to\infty} \frac{x_i^0 - y_i^0}{2^k} = 0$, there exists x_i^* such that $\lim_{k\to\infty} x_i^k = \lim_{k\to\infty} y_i^k = x_i^*$. Moreover, by this construction, we have $(x_i^k, x_j) \succeq \boldsymbol{y}$ and $\boldsymbol{y} \succ (y_i^k, x_j)$ for each k. By continuity, $(x_i^*, x_j) \sim \boldsymbol{y}$.*

Lemma 2 *For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$ such that $x_1 > y_1 > z_1$ and $x_2 < y_2 < z_2$,*

i) $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \neq \{\boldsymbol{z}\}$ implies $C(\{\boldsymbol{x}, \boldsymbol{y}, (z_1, x_2)\}) = C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \setminus \{\boldsymbol{z}\}$,

ii) $C(\{(x_1, y_2), (y_1, x_2), \boldsymbol{z}\}) = C(\{(x_1, y_2), (z_1, x_2), \boldsymbol{z}\})$.

Proof 2 (Proof of Lemma 2) *First, we prove i). Let $A = \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}$ and $A' = \{\boldsymbol{x}, \boldsymbol{y}, (z_1, x_2)\}$. Suppose $C(A) \neq \boldsymbol{z}$. Consider a menu $A \cup \{(z_1, x_2)\}$. Since $\boldsymbol{z} > (z_1, x_2)$, by monotonicity, $C(A \cup \{(z_1, x_2)\}) \not\ni (z_1, x_2)$. Therefore, since $(z_1, x_2) \geq \boldsymbol{m}^A$ and $(z_1, x_2) \notin A$, by INEA, $C(A \cup \{(z_1, x_2)\}) = C(A)$. Moreover, by INEA, $C(A' \cup \{\boldsymbol{z}\}) = C(A \cup \{(z_1, x_2)\}) = C(A) \neq \boldsymbol{z}$, $\boldsymbol{z} \geq \boldsymbol{m}^{A'}$, and $\boldsymbol{z} \notin A'$ imply that*

$$C(A) \setminus \{\boldsymbol{z}\} = C(A \cup \{(z_1, x_2)\}) \setminus \{\boldsymbol{z}\} = C(A' \cup \{\boldsymbol{z}\}) \setminus \{\boldsymbol{z}\} = C(A').$$

Second, we prove ii). Consider a menu $\{(x_1, y_2), (y_1, x_2), (z_1, x_2), \boldsymbol{z}\}$. *Since* $(x_1, y_2) > (y_1, x_2) > (z_1, x_2)$, *by monotonicity,* $C(\{(x, q), (y_1, x_2), (z_1, x_2), \boldsymbol{z}\}) \not\ni (z_1, x_2), (y_1, x_2)$. *By INEA,*

$$\text{since } C(\{(x, q), (y_1, x_2), (z_1, x_2), \boldsymbol{z}\}) \not\ni (z_1, x_2) \text{ and } (z_1, x_2) \geq \boldsymbol{m}^{\{(x_1, y_2), (y_1, x_2), \boldsymbol{z}\}},$$

we have

$$C(\{(x_1, y_2), (y_1, x_2), \boldsymbol{z}\}) = C(\{(x_1, y_2), (y_1, x_2), (z_1, x_2), \boldsymbol{z}\}).$$

Moreover, by INEA,

since $C(\{(x_1, y_2), (y_1, x_2), (z_1, x_2), \boldsymbol{z}\}) \not\ni (y_1, x_2)$ *and* $(y_1, x_2) \geq \boldsymbol{m}^{\{(x_1, y_2), (z_1, x_2), \boldsymbol{z}\}}$ *implies that*

$$C(\{(x_1, y_2), (z_1, x_2), \boldsymbol{z}\}) = C(\{(x_1, y_2), (y_1, x_2), (z_1, x_2), \boldsymbol{z}\}).$$

Therefore, we have

$$C(\{(x_1, y_2), (y_1, x_2), \boldsymbol{z}\}) = C(\{(x_1, y_2), (y_1, x_2), (z_1, x_2), \boldsymbol{z}\}) = C(\{(x_1, y_2), (z_1, x_2), \boldsymbol{z}\}).$$

B.1.7 Useful Lemmas for the proof of Theorem 4

Axiom 29 (Transitivity*) *For any* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t} \in X$ *with* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} > \boldsymbol{t}$, *if* $\boldsymbol{x} \succeq_{\boldsymbol{t}} \boldsymbol{y}$ *and* $\boldsymbol{y} \succeq_{\boldsymbol{t}} \boldsymbol{z}$, *then* $\boldsymbol{x} \succeq_{\boldsymbol{t}} \boldsymbol{z}$.

Lemma 3 *If C satisfies INEA and Monotonicity, then it satisfies Transitivity*.*

Proof 3 (Proof of Lemma 3) *Take any* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t} \in X$ *with* $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} > \boldsymbol{t}$ *and* $\boldsymbol{x} \succeq_{\boldsymbol{t}} \boldsymbol{y}$ *and* $\boldsymbol{y} \succeq_{\boldsymbol{t}} \boldsymbol{z}$. *Let us consider* $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\})$.

If $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) \ni \boldsymbol{x}$, *then since* $\boldsymbol{y} > \boldsymbol{t}$ *and* $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) \neq \boldsymbol{y}$, *by INEA, we have* $C(\{\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{t}\}) \ni \boldsymbol{x}$. *In other words,* $\boldsymbol{x} \succeq_{\boldsymbol{t}} \boldsymbol{z}$. *Now suppose* $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) \subseteq \{\boldsymbol{y}, \boldsymbol{z}\}$.

If $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) \ni \boldsymbol{y}$, *then since* $\boldsymbol{z} > \boldsymbol{t}$ *and* $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) \neq \boldsymbol{z}$, *by INEA, we have* $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}\}) = \boldsymbol{y}$. *In other words, we have* $\boldsymbol{y} \succ_{\boldsymbol{t}} \boldsymbol{x}$. *A contradiction.*

If $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) \not\ni \boldsymbol{y}$ *and* $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) \ni \boldsymbol{z}$, *then since* $\boldsymbol{x} > \boldsymbol{t}$ *and* $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) \not\ni \boldsymbol{x}$, *by INEA we have* $C(\{\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}\}) = \boldsymbol{z}$. *In other words,* $\boldsymbol{z} \succ_{\boldsymbol{t}} \boldsymbol{y}$. *A contradiction.*

Now we turn to the proof of Theorem 4. First, we focus on binary choices.

Lemma 4 *If $\succeq$ satisfies regularity, transitivity, and cancellation (i), then there exist strictly increasing and continuous functions u_1 and u_2 such that $u_1(\mathbb{R}_+) = u_2(\mathbb{R}_+) = \mathbb{R}_+$ and for any $\boldsymbol{x}, \boldsymbol{y} \in X$,*

$$\boldsymbol{x} \succeq \boldsymbol{y} \text{ if and only if } u_1(x_1) + u_2(x_2) \geq u_1(y_1) + u_2(y_2). \tag{B.1}$$

Since Lemma 4 is a relatively well-known result, we omitted the proof. For example, see Krantz et al. (1971). Now we assume that we have strictly increasing and continuous functions u_1 and u_2 that satisfy (B.1).

Lemma 5 *If C satisfies regularity, transitivity, cancellation (i), INEA, and RTI, then there exists a strictly increasing and continuous function $W : \mathbb{R}^2_+ \to \mathbb{R}_+$ such that $W(t, 0) = W(0, t) = t$ for any $t \in \mathbb{R}_+$ and for any $A \in \mathscr{A}$,*

$$C(A) = \operatorname*{argmax}_{\boldsymbol{x} \in A} \{ W(u_1(x_1) - u_1(m_1^A), u_2(x_2) - u_2(m_2^A)) \}.$$

Proof 4 (Proof of Lemma 5) *First, take a menu $B = \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}$ with $x_1 > y_1 > z_1$ and $x_2 < y_2 < z_2$. Let $W_B(\boldsymbol{x}) \equiv u_1(x_1) - u_1(z_1)$ and $W_B(\boldsymbol{z}) \equiv u_2(z_2) - u_2(x_2)$. Moreover, by Lemma 1, there exists x'_1 such that $(x'_1, x_2) \sim_{(z_1, x_2)} \boldsymbol{y}$. Then let $W_B(\boldsymbol{y}) \equiv u_1(x'_1) - u_1(z_1)$.*

Fact 1:

$$C(B) = \arg\max_{\boldsymbol{x}' \in B} \{W_B(\boldsymbol{x}')\}.$$

To prove Fact 1, we consider two cases.

Case 1: *When $C(B) \ni \boldsymbol{x}$, we shall prove that*

$$W_B(\boldsymbol{x}) = u_1(x_1) - u_1(z_1) \geq \max\{W_B(\boldsymbol{z}) = u_2(z_2) - u_2(x_2); W_B(\boldsymbol{y}) = u_1(x'_1) - u_1(z_1)\}.$$

By INEA, $\boldsymbol{x} \succeq_{\boldsymbol{y}} \boldsymbol{z}$ implies $\boldsymbol{x} \succeq \boldsymbol{z}$; equivalently, $u_1(x_1) - u_1(z_1) \geq u_2(z_2) - u_2(x_2)$ by (B.1). Moreover, since $C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \neq \{\boldsymbol{z}\}$, by Lemma 2 (i),

$$C(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \setminus \{\boldsymbol{z}\} = C(\{\boldsymbol{x}, \boldsymbol{y}, (z_1, x_2)\}) \ni \boldsymbol{x}.$$

Therefore, we obtain $\boldsymbol{x} \succeq_{(z_1, x_2)} \boldsymbol{y}$. Since $W_B(\boldsymbol{y}) = u_1(x'_1) - u_1(z_1)$ where $(x'_1, x_2) \sim_{(z_1, x_2)} \boldsymbol{y}$, by Transitivity, we have $(x_1, x_2) \succeq_{(z_1, x_2)} (x'_1, x_2)$; equivalently, $x_1 \geq x'_1$. Therefore, $u_1(x_1) - u_1(z_1) \geq u_1(x') - u_1(z_1)$.*

Case 2: *When $C(B) \ni \boldsymbol{y}$, we shall prove that*

$$W_B(\boldsymbol{y}) = u_1(x_1')-u_1(z_1) \geq \max\{W_B(\boldsymbol{z}) = u_2(z_2)-u_2(x_2); W_B(\boldsymbol{x}) = u_1(x_1)-u_1(z_1)\}.$$

First, since $C(\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}) \neq \{\boldsymbol{z}\}$, by Lemma 2 (i),

$$C(\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}) \setminus \{\boldsymbol{z}\} = C(\{\boldsymbol{x},\boldsymbol{y},(z_1,x_2)\}) \ni \boldsymbol{y}.$$

Therefore, we obtain $\boldsymbol{y} \succsim_{(z_1,x_2)} \boldsymbol{x}$. Since $W_B(\boldsymbol{y}) = u_1(x_1')-u_1(z_1)$ where $(x_1',x_2) \sim_{(z_1,x_2)} \boldsymbol{y}$, by Transitivity, we have $(x_1',x_2) \sim_{(z_1,x_2)} \boldsymbol{x}$; equivalently, $x_1' \geq x_1$. Therefore, $u_1(x_1') - u_1(z_1) \geq u_1(x_1) - u_1(z_1)$.*

Second, since $C(\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}) \neq \{\boldsymbol{x}\}$, by Lemma 2 (i),

$$C(\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}) \setminus \{\boldsymbol{x}\} = C(\{(z_1,x_2),\boldsymbol{y},\boldsymbol{z}\}) \ni \boldsymbol{y}.$$

Therefore, we obtain $\boldsymbol{y} \succsim_{(z_1,x_2)} \boldsymbol{z}$. Since $W_B(\boldsymbol{y}) = u_1(x_1')-u_1(z_1)$ where $(x_1',x_2) \sim_{(z_1,x_2)} \boldsymbol{y}$, by Transitivity, we have $(x_1',x_2) \succsim_{(z_1,x_2)} \boldsymbol{z}$. Since $(z_1,x_2) \geq \boldsymbol{m}^{\{(x_1',x_2),\boldsymbol{z}\}}$, by INEA, $(x_1',x_2) \succsim_{(z_1,x_2)} \boldsymbol{z}$ implies $(x_1',x_2) \succsim \boldsymbol{z}$; equivalently, $u_1(x_1')-u_1(z_1) \geq u_2(z_2)-u_2(x_2)$ by (B.1).*

Now we construct W in the following way: for any menu $B = \{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}$ with $x_1 > y_1 > z_1$ and $x_2 < y_2 < z_2$, let $W(u_1(x_1) - u_1(z_1), 0) \equiv W_B(\boldsymbol{x}) = u_1(x_1) - u_1(z_1)$, $W(0, u_2(z_2) - u_2(x_2)) \equiv W_B(\boldsymbol{z}) = u_2(z_2) - u_2(x_2)$, and $W(u_1(y_1) - u_1(z_1), u_2(y_2) - u_2(x_2)) \equiv W_B(\boldsymbol{y})$. Finally, note that $W(u_1(y_1) - u_1(z_1), u_2(y_2) - u_2(x_2))$ is a function of y_1, z_1, x_2, y_2 because $W_B(\boldsymbol{y})$ does not depend on x_1 and z_2. Now we shall prove that W is well-defined.

Fact 2: *W is well-defined; that is, for any y_1, z_1, x_2, y_2 and y_1', z_1', x_2', y_2', if $u_1(y_1) - u_1(z_1) = u_1(y_1') - u_1(z_1')$ and $u_2(y_2) - u_2(x_2) = u_2(y_2') - u_2(y_2')$, then $W(u_1(y_1) - u_1(z_1), u_2(y_2) - u_2(x_2)) = W(u_1(y_1') - u_1(z_1'), u_2(y_2') - u_2(x_2'))$.*

First, by (B.1), note that $u_1(y_1) - u_1(z_1) = u_1(y_1') - u_1(z_1')$ is equivalent to $[y_1,z_1]D_1[y_1',z_1']$ and $u_2(y_2)-u_2(x_2) = u_2(y_2')-u_2(x_2')$ is equivalent to $[y_2,x_2]D_2[y_2',x_2']$.

By the definition of W, $W(u_1(y_1) - u_1(z_1), u_2(y_2) - u_2(x_2)) = u_1(x_1) - u_1(z_1)$ where x_1 satisfies $\boldsymbol{x} \sim_{(z_1,x_2)} \boldsymbol{y}$. Let us find x_1' such that $u_1(x_1) - u_1(z_1) = u_1(x_1') -$

$u_1(z_1')$. By the definition of W, we shall prove that $W(u_1(y_1') - u_1(z_1'), u_2(y_2') - u_2(x_2')) = u_1(x_1') - u_1(z_1')$; i.e., $\boldsymbol{x'} \sim_{(z_1', x_2')} \boldsymbol{y'}$.

Since $u_1(x_1) - u_1(z_1) = u_1(x_1') - u_1(z_1')$ and $u_1(y_1) - u_1(z_1) = u_1(y_1') - u_1(z_1')$, we have $u_1(x_1) - u_1(y_1) = u_1(x_1') - u_1(y_1')$; i.e., $[x_1, y_1]D_1[x_1', y_1']$.

Therefore, since $[x_1, y_1]D_1[x_1', y_1']$, $[y_1, z_1]D_1[y_1', z_1']$, $[y_2, x_2]D_2[y_2', x_2']$, and $[x_2, x_2]D_2[x_2', x_2']$, by RTI, we have $\boldsymbol{x} \sim_{(z_1, x_2)} \boldsymbol{y}$ if and only if $\boldsymbol{x'} \sim_{(z_1', x_2')} \boldsymbol{y'}$.

Now we will prove that we obtained desired W.

Fact 3: *For any menu $A \in \mathscr{A}$,*

$$C(A) = \operatorname*{argmax}_{\boldsymbol{x} \in A} \{ W(u_1(x_1) - u_1(m_1^A), u_2(x_2) - u_2(m_2^A)) \}.$$

Let $|A| = n$. It is obvious when $n = 2$. Now consider the cases when $n \geq 3$. We have already proved the case when $A = B = \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}$ with $x_1 > y_1 > z_1$ and $x_2 < y_2 < z_2$. We now need to consider two other cases when $n = 3$.

Case 1: *$A = \{\boldsymbol{y}, \boldsymbol{y'}, \boldsymbol{t}\}$ with $\boldsymbol{y}, \boldsymbol{y'} > \boldsymbol{t}$ and $C(A) \ni \boldsymbol{y}$.*

By the definition of W, we have $W(u_1(y_1) - u_1(t_1), u_2(y_2) - u_2(t_2)) = u_1(x_1) - u_1(t_1)$ where $(x_1, t_2) \sim_{\boldsymbol{t}} \boldsymbol{y}$ and $W(u_1(y_1') - u_1(t_1), u_2(y_2') - u_2(t_2)) = u_1(x_1') - u_1(t_1)$ where $(x_1', t_2) \sim_{\boldsymbol{t}} \boldsymbol{y'}$. Then by Transitivity, we have $(x_1, t_2) \sim_{\boldsymbol{t}} \boldsymbol{y} \succeq_{\boldsymbol{t}} \boldsymbol{y'} \sim_{\boldsymbol{t}} (x_1', t_2)$. Therefore, we have $x_1 \geq x_1'$. In other words, $W(u_1(y_1) - u_1(t_1), u_2(y_2) - u_2(t_2)) = u_1(x_1) - u_1(t_1) \geq W(u_1(y_1') - u_1(t_1), u_2(y_2') - u_2(t_2)) = u_1(x_1') - u_1(t_1)$.*

Case 2: *$A = \{(x_1, y_2), (y_1, x_2), \boldsymbol{z}\}$ with $x_1 > y_1 > z_1$ and $x_2 < y_2 < z_2$.*

By Lemma 2 (ii), we have $C(\{(x_1, y_2), (y_1, x_2), \boldsymbol{z}\}) = C(\{(x_1, y_2), (z_1, x_2), \boldsymbol{z}\})$. Since $(x_1, y_2), \boldsymbol{z} > (z_1, x_2)$, the proof of Case 1 concludes the proof of Case 2.

Now let us consider menus with more than three alternatives. Let $A = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n\}$ and $C(A) \ni \boldsymbol{x}^k$. We shall prove that for any $i \neq k$,

$$W(u_1(x_1^k) - u_1(m_1^A), u_2(x_2^k) - u_2(m_2^A)) \geq W(u_1(x_1^i) - u_1(m_1^A), u_2(x_2^i) - u_2(m_2^A)).$$

Without loss of generality, let $x_1^1 \geq x_1^2 \geq \ldots \geq x_1^n$ and $x_2^s = m_2^A$ for some s. Then we shall prove that for any $i \neq k$,

$$W(u_1(x_1^k) - u_1(x_1^n), u_2(x_2^k) - u_2(x_2^s)) \geq W(u_1(x_1^i) - u_1(x_1^n), u_2(x_2^i) - u_2(x_2^s)).$$

We consider several cases and in all cases, we eliminate alternatives that do not include minimums (x_1^n and x_2^s) from $A \setminus \{\boldsymbol{x}^i, \boldsymbol{x}^k\}$.

Case 1: *$s \neq i, n, k$ and $k = n$.*

By INEA, $C(A) \ni \boldsymbol{x}^n$ implies that $C(\{\boldsymbol{x}^i, \boldsymbol{x}^s, \boldsymbol{x}^n\}) \ni \boldsymbol{x}^n$. Then by the previous argument for $n = 3$, $u_2(x_2^n) - u_2(x_2^s) \geq W\big(u_1(x_1^i) - u_1(x_1^n), u_2(x_2^i) - u_2(x_2^s)\big)$.

Case 2: *$s \neq i, n$ and $k = s$.*

By INEA, $C(A) \ni \boldsymbol{x}^s$ implies that $C(\{\boldsymbol{x}^i, \boldsymbol{x}^s, \boldsymbol{x}^n\}) \ni \boldsymbol{x}^s$. Then by the previous argument for $n = 3$, $u_1(x_1^s) - u_1(x_1^n) \geq W\big(u_1(x_1^i) - u_1(x_1^n), u_2(x_2^i) - u_2(x_2^s)\big)$.

Case 3: *$s = n \neq k$.*

By INEA, $C(A) \ni \boldsymbol{x}^k$ implies that $C(\{\boldsymbol{x}^i, \boldsymbol{x}^k, \boldsymbol{x}^n\}) \ni \boldsymbol{x}^k$. Then by the previous argument for $n = 3$, $W\big(u_1(x_1^k) - u_1(x_1^n), u_2(x_2^k) - u_2(x_2^n)\big) \geq W\big(u_1(x_1^i) - u_1(x_1^n), u_2(x_2^i) - u_2(x_2^n)\big)$.

Case 4: *$s = i$.*

By INEA, $C(A) \ni \boldsymbol{x}^k$ implies that $C(\{\boldsymbol{x}^i, \boldsymbol{x}^k, \boldsymbol{x}^n\}) \ni \boldsymbol{x}^k$. Then by the previous argument for $n = 3$, $W\big(u_1(x_1^k) - u_1(x_1^n), u_2(x_2^k) - u_2(x_2^i)\big) \geq u_1(x_1^i) - u_1(x_1^n)$.

Case 5: *$s \neq i, n, k$ and $k \neq n$.*

Let $A' = \{\boldsymbol{x}^i, \boldsymbol{x}^s, \boldsymbol{x}^k, \boldsymbol{x}^n\}$. By INEA, $C(A) \ni \boldsymbol{x}^k$ implies that $C(A') \ni \boldsymbol{x}^k$. Let us consider a menu $A' \cup \{(x_1^n, x_2^s)\} = \{\boldsymbol{x}^i, \boldsymbol{x}^s, \boldsymbol{x}^k, \boldsymbol{x}^n, (x_1^n, x_2^s)\}$. By monotonicity, $C(A' \cup \{(x_1^n, x_2^s)\}) \not\ni (x_1^n, x_2^s)$. Therefore, by INEA, $C(A') = C(A' \cup \{(x_1^n, x_2^s)\})$.

Since $C(A' \cup \{(x_1^n, x_2^s)\}) = C(A') \neq \boldsymbol{x}^n$, by INEA,

$$C(A' \cup \{(x_1^n, x_2^s)\} \setminus \{\boldsymbol{x}^n\}) = C(A' \cup \{(x_1^n, x_2^s)\}) \setminus \{\boldsymbol{x}^n\} = C(A') \setminus \{\boldsymbol{x}^n\} \ni \boldsymbol{x}^k.$$

Moreover, since $C(A' \cup \{(x_1^n, x_2^s)\} \setminus \{\boldsymbol{x}^n\}) \neq \boldsymbol{x}^s$, by INEA, we have

$$\begin{aligned} C(A' \cup \{(x_1^n, x_2^s)\} \setminus \{\boldsymbol{x}^n\}) \setminus \{\boldsymbol{x}^s\} &= C(A' \cup \{(x_1^n, x_2^s)\} \setminus \{\boldsymbol{x}^n, \boldsymbol{x}^s\}) \\ &= C(\{\boldsymbol{x}^i, \boldsymbol{x}^k, (x_1^n, x_2^s)\}). \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} C(A') \setminus \{\boldsymbol{x}^n, \boldsymbol{x}^s\} &= C(A' \cup \{(x_1^n, x_2^s)\} \setminus \{\boldsymbol{x}^n, \boldsymbol{x}^s\}) \\ &= C(\{\boldsymbol{x}^i, \boldsymbol{x}^k, (x_1^n, x_2^s)\}) \ni \boldsymbol{x}^k. \end{aligned}$$

Then by the previous argument for $n = 3$, we have $W(u_1(x_1^k) - u_1(x_1^n), u_2(x_2^k) - u_2(x_2^s)) \geq W(u_1(x_1^i) - u_1(x_1^n), u_2(x_2^i) - u_2(x_2^s))$.

B.1.8 Proof of Theorem 4

We focus on the sufficiency part of Theorem 4. See Appendix F for the necessity part of Theorem 4. By Lemma 5, there exist strictly increasing continuous functions u_1, u_2, and $W : \mathbb{R}_+^2 \to \mathbb{R}_+$ such that $W(t, 0) = W(0, t) = t$ for each $t \in \mathbb{R}_+$ and for any $A \in \mathscr{A}$,

$$C(A) = \operatorname*{argmax}_{\mathbf{x} \in A} \{ W(u_1(x_1) - u_1(m_1^A), u_2(x_2) - u_2(m_2^A)) \}.$$

Fact: If C satisfies cancellation (ii), then there exists a strictly increasing continuous function f such that $W(t_1, t_2) = f^{-1}(f(t_1) + f(t_2))$.

Let us consider a binary relation $\succsim_{\mathbf{0}}$. We know that $\mathbf{x} \succsim_{\mathbf{0}} \mathbf{y}$ if and only if $W(u_1(x_1), u_2(x_2)) \geq W(u_1(y_1), u_2(y_2))$. Similar to Lemma 4, by cancellation (ii) and transitivity*, there exist strictly increasing continuous functions f_1 and f_2 such that $f_i(\mathbb{R}_+) = \mathbb{R}_+$ and $\mathbf{x} \succsim_{\mathbf{0}} \mathbf{y}$ if and only if $f_1(x_1) + f_2(x_2) \geq f_1(y_1) + f_2(y_2)$. See Krantz et al. (1971).

Take any $\mathbf{x}, \mathbf{y}$ such that $\mathbf{x} \sim_{\mathbf{0}} \mathbf{y}$. Equivalently, we have $f_1(x_1) + f_2(x_2) = f_1(y_1) + f_2(y_2)$. Given that f_1 is continuous and strictly increasing, we obtain $x_1 = f_1^{-1}(f_1(y_1) + f_2(y_2) - f_2(x_2))$. Therefore, $W(u_1(x_1), u_2(x_2)) = W(u_1(y_1), u_2(y_2))$ implies that

$$W(u_1(f_1^{-1}(f_1(y_1) + f_2(y_2) - f_2(x_2))), u_2(x_2)) = W(u_1(y_1), u_2(y_2)).$$

Since the right hand side does not depend on x_2, the left hand side also should not depend on x_2. Therefore, there exists a continuous and strictly increasing function F such that

$$W(u_1(f_1^{-1}(f_1(y_1)+f_2(y_2)-f_2(x_2))), u_1(x_2)) = F(f_1(y_1)+f_2(y_2)) = W(u_1(y_1), u_2(y_2)).$$

Since $W(t, 0) = W(0, t) = t$, if we set $y_i = 0$, then the above equality implies that $F(f_j(y_j)) = W(u_j(y_j), 0) = u_j(y_j)$. Therefore, $f_j = F^{-1} \circ u_j$. Therefore, $W(t_1, t_2) = F(f_1(u^{-1}(t_1)) + f_2(u^{-2}(t_2))) = F(F^{-1}(t_1) + F^{-1}(t_2))$. Let $f = F^{-1}$, then we have $W(t_1, t_2) = f^{-1}(f(t_1) + f(t_2))$.

Now it is easy to see that for any $A \in \mathscr{A}$,

$$C(A) = \arg \max_{(x,p) \in A} \{ f(u_1(x_1) - u_1(m_1^A)) + f(u_2(x_2) - u_2(m_2^A)) \}.$$

Uniqueness: Take any two vectors of continuous functions (f, u_1, u_2) and (f', u_1', u_2') such that $C = C_{(f,u_1,u_2)} = C_{(f',u_1',u_2')}$, $u_1(1) = u_1'(1)$, and $f(1) = f'(1)$. We prove the uniqueness in two steps.

Step 1. $u_1' = u_1$ and $u_2' = u_2$.

Take any $\mathbf{y} \in X$. By Lemma 1, there exists $x_1 \in \mathbb{R}_+$ such that $(x, 0) \sim \mathbf{y}$. Then we have

$$u_1(x_1) = u_1(y_1) + u_2(y_2) \Leftrightarrow u_1(x_1) = u_1(y_1) + u_2(y_2);\ \ \text{equivalently,}$$

$$x_1 = u_1^{-1}(u_1(y_1) + u_2(y_2)) = u_1'^{-1}(u_1'(y_1) + u_2'(y_2))$$

since all functions are strictly increasing and continuous.

First, let $y_1 = 0$. Then we have $u_1^{-1} \circ u_2 = u_1'^{-1} \circ u_2'$; equivalently, $u_1' \circ u_1^{-1} = u_2' \circ u_2^{-1}$.

Second, let $t_1 \equiv u_1(y_1)$, $t_2 \equiv u_2(y_2)$, and $h(t) \equiv u_1'(u_1^{-1}(t))$. Since u_1 and u_1' are strictly increasing and continuous and $u_1(\mathbb{R}_+) = u_1'(\mathbb{R}_+) = \mathbb{R}_+$, h is also strictly increasing and continuous and $h(\mathbb{R}_+) = \mathbb{R}_+$. Then we have

$$\begin{aligned}
h(t_1 + t_2) = u_1'(u_1^{-1}(t_1 + t_2)) &= u_1'(u_1^{-1}(u_1(y_1) + u_2(y_2))), \text{(by the definitions of } h, t_1, t_2)\\
&= u_1'(u_1'^{-1}(u_1'(y_1) + u_2'(y_2))) = u_1'(y_1) + u_2'(y_2)\\
&= u_1'(u_1^{-1}(t_1)) + u_2'(u_2^{-1}(t_2)), \text{(by the definitions of } t_1, t_2)\\
&= u_1'(u_1^{-1}(t_1)) + u_1'(u_1^{-1}(t_2)) = h(t_1) + h(t_2).
\end{aligned}$$

Since we obtained a typical Cauchy functional equation (see Kuczma 2008), there exists $\alpha > 0$ such that $h(t) = \alpha\, t$; that is, $u_1' = \alpha\, u_1$ and $u_2' = \alpha\, u_2$. Moreover, since $u_1(1) = u_1'(1)$, we have $u_1' = u_1$ and $u_2' = u_2$.

Step 2. $f' = f$.

Take any $\mathbf{y} \in X$. By Lemma 1, there exists $x \in \mathbb{R}_+$ such that $(x, 0) \sim_{(0,0)} \mathbf{y}$. Then we have

$$f(u_1(x_1)) = f(u_1(y_1)) + f(u_2(y_2)) \Leftrightarrow f'(u_1(x_1)) = f'(u_1(y_1)) + f'(u_2(y_2)).$$

Since u_1, u_2, f, f' are strictly increasing and continuous, similar to Step 1, we will obtain a Cauchy functional equation. Therefore, there exists $\beta > 0$ such that $f' \circ u_1 = \beta \cdot f \circ u_1$ and $f' \circ u_2 = \beta \cdot f \circ u_2$. Since $f'(1) = f(1)$, we have $f' = f$.

B.2 Alternatives with n-attributes

Let us briefly discuss how to generalize our result to the n-dimension case. In particular, let us consider the following model:

$$C(A) = \arg\max_{\mathbf{x}\in A} \sum_{i=1}^{n} f(u_i(x_i) - u_i(m_i^A)) \tag{B.2}$$

where $A \subset \mathbb{R}_+^n$ and $\mathbf{m}^A$ is the vector that consists of the minimums of attributes of A; that is, $m_i^A = \min_{\mathbf{x}\in A} x_i$ for each i.

In addition to standard axioms, we need three axioms. First, we need to INEA. Under INEA, $\mathbf{x}$ is chosen over $\mathbf{y}$ from a menu A iff $\mathbf{x}$ is chosen over $\mathbf{y}$ from the menu $\{\mathbf{x}, \mathbf{y}, \mathbf{m}^A\}$ (i.e., $\mathbf{x} \succ_{\mathbf{m}^A} \mathbf{y}$). Therefore, we can focus on $\{\succeq_{\mathbf{t}}\}_{\mathbf{t}\in\mathbb{R}_+^n}$. Now we will use a very standard technique to obtain an additive representation for each $\succeq_{\mathbf{t}}$. Under a general version of cancellation (see Debreu (1960b) and Wakker (1988)), we can find functions $w_{\mathbf{t}}^1, \ldots, w_{\mathbf{t}}^n$ such that for any $\mathbf{x}, \mathbf{y}$ with $\mathbf{x}, \mathbf{y} > \mathbf{t}$,

$$\mathbf{x} \succeq_{\mathbf{t}} \mathbf{y} \text{ iff } \sum_{i=1}^{n} w_{\mathbf{t}}^i(x_i) \geq \sum_{i=1}^{n} w_{\mathbf{t}}^i(y_i).$$

Without loss of generality, for any $w_{\mathbf{t}}^i$, we can write that $w_{\mathbf{t}}^i(x_i) = f_{\mathbf{t}}^i(u_i(x_i) - u_i(t_i))$ where u_i is a utility function that is consistent with relative distance D_i. Therefore, we have

$$\mathbf{x} \succeq_{\mathbf{t}} \mathbf{y} \text{ iff } \sum_{i=1}^{n} f_{\mathbf{t}}^i(u_i(x_i) - u_i(t_i)) \geq \sum_{i=1}^{n} f_{\mathbf{t}}^i(u_i(y_i) - u_i(t_i)).$$

Then we modify RTI in the following way:

Axiom 30 (Reference Translation Invariance* (RTI*)) *Take any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}_+^n$ and any i. For any $x_i', y_i', z_i' \in \mathbb{R}_+$ such that $[x_i, y_i]D_i[x_i', y_i']$ and $[y_i, z_i]D_i[y_i', z_i']$,*

$$(x_i, \boldsymbol{x}_{-i}) \sim_{(z_i, z_{-i})} (y_i, \boldsymbol{y}_{-i}) \textit{ if and only if } (x_i', \boldsymbol{x}_{-i}) \sim_{(z_i', z_{-i})} (y_i', \boldsymbol{z}_{-i}).$$

Under RTI*, we can prove that $f_{\mathbf{t}}^i$ is independent of $\mathbf{t}$. Therefore, we have

$$\mathbf{x} \succeq_{\mathbf{t}} \mathbf{y} \text{ iff } \sum_{i=1}^{n} f^i(u_i(x_i) - u_i(t_i)) \geq \sum_{i=1}^{n} f^i(u_i(y_i) - u_i(t_i)).$$

Then we modify transitivity in the following way.

Axiom 31 (Transitivity)** *For any i, j and $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n_+$,*

$$\text{if } (x_i, x_j, \boldsymbol{0}_{-i,-j}) \succeq (y_i, y_j, \boldsymbol{0}_{-i,-j}) \text{ and } (y_i, y_j, \boldsymbol{0}_{-i,-j}) \succeq (z_i, z_j, \boldsymbol{0}_{-i,-j}), \text{ then}$$

$$(x_i, x_j, \boldsymbol{0}_{-i,-j}) \succeq (z_i, z_j, \boldsymbol{0}_{-i,-j}).$$

It turns out, under transitivity**, we can prove that $f^i = f$ for some f. Therefore, we have

$$\mathbf{x} \succeq_{\mathbf{t}} \mathbf{y} \text{ iff } \sum_{i=1}^{n} f(u_i(x_i) - u_i(t_i)) \geq \sum_{i=1}^{n} f(u_i(y_i) - u_i(t_i)).$$

In other words, under INEA, RTI*, Transitivity*, and a general version of cancellation in addition to standard axioms, we have (B.2).

B.3 Using both Maximums and Minimums

Here we consider models in which reference points not only depend on the minimums of the menu, but also the maximums. Let $\mathbf{M}^A = (M_1^A, M_2^A) \equiv (\max_{\mathbf{x}\in A} x_1, \max_{\mathbf{x}\in A} x_2)$ for each $A \in \mathscr{A}$ (i.e., the join of A). We focus on the following more general model: for any menu $A \in \mathscr{A}$,

$$C(A) = \arg\max_{\mathbf{x}\in A} \{f(u_1(x_1) - u_1(r_1(M_1^A, m_1^A))) + f(u_2(x_2) - u_2(r_1(M_2^A, m_2^A)))\} \tag{B.3}$$

where r_1 and r_2 are strictly increasing reference functions. In Appendix B.3.1-2, we discuss implications of (B.3) by reconsidering section 2.2.2-3.

B.3.1 Attraction Effect and Compromise Effect

First, we discuss a relation between the compromise and attraction effects. Previously, Proposition 2 showed that an agent exhibits the the compromise effect if and only if she exhibits the attraction effect. But in the model (B.3), we have a weaker prediction. In particular, we can show that if the agent exhibits the attraction effect, then she exhibits the compromise effect, which means that the compromise effect is more likely to be exhibited than the attraction effect.

To illustrate, suppose the agent exhibits the attraction effect at some $\mathbf{x}$, $\mathbf{y}$, and (z_1, z_2') with with $x_1 > y_1 > z_1$ and $y_2 > z_2' > x_2$. That is, $\mathbf{x} \succ \mathbf{y}$ and $\mathbf{x} \prec_{(z_1, z_2')} \mathbf{y}$. By the representation (B.3), we have

$$f(u_1(x_1) - u_1(r_1(x_1, y_1))) - f(u_1(y_1) - u_1(r_1(x_1, y_1))) >$$

$$> f(u_2(y_2) - u_2(r_2(y_2, x_2))) - f(u_2(x_2) - u_2(r_2(y_2, x_2)))$$

and

$$f(u_1(x_1) - u_1(r_1(x_1, z_1))) - f(u_1(y_1) - u_1(r_1(x_1, z_1))) <$$

$$< f(u_2(y_2) - u_2(r_2(y_2, x_2))) - f(u_2(x_2) - u_2(r_2(y_2, x_2))).$$

Take any z_2 such that $z_2 > y_2$. Then since $r_2(z_2, x_2) > r_2(y_2, x_2)$, by diminishing sensitivity, we have

$$f(u_2(y_2) - u_2(r_2(y_2, x_2))) - f(u_2(x_2) - u_2(r_2(y_2, x_2))) <$$

$$< f(u_2(y_2) - u_2(r_2(z_2, x_2))) - f(u_2(x_2) - u_2(r_2(z_2, x_2))).$$

Therefore, we have

$$f(u_1(x_1) - u_1(r_1(x_1, z_1))) - f(u_1(y_1) - u_1(r_1(x_1, z_1))) <$$

$$< f(u_2(y_2) - u_2(r_2(z_2, x_2))) - f(u_2(x_2) - u_2(r_2(z_2, x_2)));$$

equivalently, $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$. In other words, the agent exhibits the compromise effect at $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ since $\mathbf{x} \succ \mathbf{y}$ and $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$.

B.3.2 Symmetric Dominance and Two Decoy Effect

Second, we discuss the effect of the symmetrically dominated third alternative (z_1, t_2) as in Observation 1. Consider again two alternatives $\mathbf{x}$ and $\mathbf{y}$ with $\mathbf{x} \succ \mathbf{y}$ and $x_1 > y_1$ and $x_2 < y_2$. Remember that Observation 1 showed that when $t_2 < x_2 < z_2' < y_2$, observing $\mathbf{x} \prec_{(z_1, t_2)} \mathbf{y}$ is less likely than $\mathbf{x} \prec_{(z_1, z_2')} \mathbf{y}$; that is, $\mathbf{x} \prec_{(z_1, t_2)} \mathbf{y}$ implies $\mathbf{x} \prec_{(z_1, z_2')} \mathbf{y}$. We can obtain the same result as in Observation 1.

To illustrate, suppose $\mathbf{x} \prec_{(z_1, t_2)} \mathbf{y}$. By the representation (B.3), we have

$$f(u_1(x_1) - u_1(r_1(x_1, z_1))) - f(u_1(y_1) - u_1(r_1(x_1, z_1))) <$$

$$< f(u_2(y_2) - u_2(r_2(y_2, t_2))) - f(u_2(x_2) - u_2(r_2(y_2, t_2))).$$

By diminishing sensitivity and $r_2(y_2, x_2) > r_2(y_2, t_2)$, we have

$$f(u_2(y_2) - u_2(r_2(y_2, t_2))) - f(u_2(x_2) - u_2(r_2(y_2, t_2))) <$$

$$< f(u_2(y_2) - u_2(r_2(y_2, x_2))) - f(u_2(x_2) - u_2(r_2(y_2, x_2))).$$

Therefore, we have

$$f(u_1(x_1) - u_1(r_1(x_1, z_1))) - f(u_1(y_1) - u_1(r_1(x_1, z_1))) <$$

$$< f(u_2(y_2) - u_2(r_2(y_2, x_2))) - f(u_2(x_2) - u_2(r_2(y_2, x_2)));$$

equivalently, $\mathbf{x} \prec_{(z_1, z_2')} \mathbf{y}$.

Finally, note that the symmetric dominance and the two decoy effect are also equivalent as in Section 2.2.3 since the maximums and the minimums are identical. Therefore, we can summarize Section B.3 in the following way ($\dot{\geq}$ represents the likelihood of observing a preference reversal):

Compromise $\dot{>}$ Attraction (Decoy) $\dot{>}$ Two Decoy $\doteq$ Symmetric Dominance.

B.4 Asymmetry of Two Dimensions and Violations of Transitivity

In Section 2.2.2 and Section 2.3.2, we demonstrated the equivalence between diminishing sensitivity and the compromise and attraction effects. In particular, Proposition 3 shows that if C is more diminishing sensitive (if f is more concave), then the compromise and attraction effects are more likely to be observed. So far we treated two dimensions in a symmetric way in sense that two dimensions have a common distortion function f (recall and compare (2.1) and (2.2)). Now we consider a general representation in which two dimensions have different distortion functions, f and g as in (2.1). More formally,

Definition 21 *A choice correspondence C is a* general additive reference dependent choice *(GARDC) if there are strictly increasing functions f, g, u_1, u_2 such that $f(0) = g(0) = 0$ and for any menu $A \in \mathscr{A}$,*

$$C(A) = \arg\max_{\boldsymbol{x} \in A} \{f(u_1(x_1) - u_1(m_1^A)) + g(u_2(x_2) - u_2(m_2^A))\}. \tag{B.4}$$

Now we have two different diminishing sensitivities for two dimensions and they are defined as in (2.3). Indeed, diminishing sensitivities are equivalent to the strict concavity of f and g. The new representation allows us to compare two dimensions in terms of diminishing sensitivity; i.e., to know which of f and g is more concave. The generalization allows us to obtain the following observed behavior: Heath and Chatterjee (1995) found that, when the dimensions are quality and price, one is less likely to observe the attraction effect when the third alternative is a decoy for the low-quality alternative compared to the high-quality alternative. More precisely, one is more like to have $\mathbf{y} \succ \mathbf{x}$ and $\mathbf{y} \prec_{(k_1,t_2)} \mathbf{x}$ compared to $\mathbf{x} \succ \mathbf{y}$ and $\mathbf{x} \prec_{(z_1,z_2')} \mathbf{y}$ (See Figure B.2). We can have this behavior when g is more diminishing sensitive than f by an argument similar to Proposition 2-3.

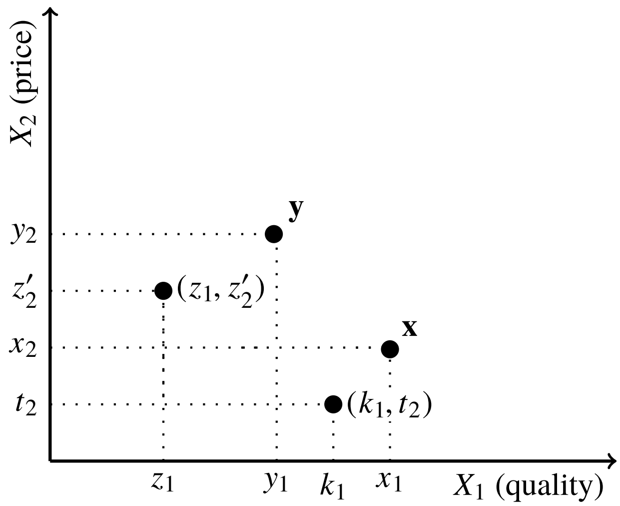

Figure B.2: Asymmetry of Two Dimensions

It turns out the comparison of two dimensions in terms of diminishing sensitivity is closely related to violations of transitivity of binary comparisons.[1] Violations of transitivity are consistently documented in experimental literature.[2]

Now we discuss how to know which of two dimensions is more diminishing sensitivity; that is, which of f and g is more concave, from observed choices. Essentially, a direction of a violation of transitivity tells us which dimension is more diminishing sensitive. Indeed, f can be more concave than g for only a subset of $\mathbb{R}_+$, but in order to illustrate a connection between violations of transitivity and the relative diminishing sensitivity, we suppose $f(t) = t^\alpha$ and $g(t) = t^\beta$ for some α and β. Therefore, we have

$$C(A) = \arg\max_{\mathbf{x}\in A}\{(u_1(x_1) - u_1(m_1^A))^\alpha + (u_2(x_2) - u_2(m_2^A))^\beta\}.$$

Our objective is to know either $\alpha = \beta$, $\alpha > \beta$, or $\alpha < \beta$. Take some alternatives $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ with $x_1 > y_1 > z_1$ and $z_2 > y_2 > x_2$ such that

$$\mathbf{x} \sim \mathbf{y} \text{ and } \mathbf{y} \sim \mathbf{z};\ \text{equivalently,}$$

by our representation,

$$(u_1(x_1) - u_1(y_1))^{\frac{\alpha}{\beta}} = u_2(y_2) - u_2(x_2) \text{ and } (u_1(y_1) - u_1(z_1))^{\frac{\alpha}{\beta}} = u_2(z_2) - u_2(y_2).$$

[1] In fact, we can obtain (B.4) by weakening transitivity of $\succeq$, but we will omit a behavioral foundation for it since it is very similar to Theorem 4.

[2] For example, a violation of transitivity of pairwise comparisons of binary lotteries is documented in Tversky (1969), Lindman and Lyons (1978), Budescu and Weiss (1987), Loomes et al. (1991), and Day and Loomes (2010).

There are three cases for the comparison between $\mathbf{x}$ and $\mathbf{z}$: either i) $\mathbf{x} \sim \mathbf{z}$; ii) $\mathbf{x} \prec \mathbf{z}$; or iii) $\mathbf{x} \succ \mathbf{z}$.

In the first case, we have $\alpha = \beta$ because $\mathbf{x} \sim \mathbf{z}$ implies

$$(u_1(x_1) - u_1(y_1))^{\frac{\alpha}{\beta}} + (u_1(y_1) - u_1(z_1))^{\frac{\alpha}{\beta}} = u_2(y_2) - u_2(x_2) + u_2(z_2) - u_2(y_2) =$$

$$= u_2(z_2) - u_2(x_2) = (u_1(x_1) - u_1(z_1))^{\frac{\alpha}{\beta}}.$$

Therefore, if transitivity is satisfied, two dimensions are the same in terms of diminishing sensitivity ($\alpha = \beta$).

In the second case, we have $\alpha < \beta$ because $\mathbf{x} \prec \mathbf{z}$ implies

$$(u_1(x_1) - u_1(y_1))^{\frac{\alpha}{\beta}} + (u_1(y_1) - u_1(z_1))^{\frac{\alpha}{\beta}} = u_2(y_2) - u_2(x_2) + u_2(z_2) - u_2(y_2) =$$

$$= u_2(z_2) - u_2(x_2) > (u_1(x_1) - u_1(z_1))^{\frac{\alpha}{\beta}}.$$

Therefore, if transitivity is violated in a direction $\mathbf{x} \sim \mathbf{y} \sim \mathbf{z} \succ \mathbf{x}$, then the the first dimension is more diminishing sensitive than the second dimension ($\alpha < \beta$).

Lastly, if transitivity is violated in a direction $\mathbf{x} \sim \mathbf{y} \sim \mathbf{z} \prec \mathbf{x}$, the the second dimension is more diminishing sensitive than the first dimension ($\alpha > \beta$).

APPENDIX TO CHAPTER 3

C.1 Proof of Theorem 5

Since the necessity part is obvious, we only prove the sufficiency part. Suppose a binary relation $\succeq$ is regular and satisfies Separability. We then shall prove that there exist continuous distance-based functions $\{f_i\}_{i=1}^n$ and a continuous aggregator W such that (3.2) holds. First, we will prove the following useful lemma. Recall that a_k and b_k are the infimum and the supremum of $X_k = (a_k, b_k]$, respectively.

Lemma 6 *Suppose $\succeq$ is regular. Take any $\boldsymbol{x} \in X$. For any i and $\boldsymbol{y}_{-i} \in X_{-i}$ with $(b_i, \boldsymbol{y}_{-i}) \succeq \boldsymbol{x}$, there exists $y_i \in X_i$ such that $\boldsymbol{x} \sim \boldsymbol{y} = (y_i, \boldsymbol{y}_{-i})$.*

Proof of Lemma 6 Take any $\mathbf{x} \in X$, i, and $\mathbf{y}_{-i} \in X_{-i}$ such that $(b_i, \mathbf{y}_{-i}) \succeq \mathbf{x}$. We shall find $y_i \in X_i$ such that $\mathbf{x} \sim \mathbf{y}$. By richness, there exists $\overline{y}_i \in X_i$ such that $\mathbf{x} \succ (\overline{y}_i, \mathbf{y}_{-i})$. We now construct two infinite sequences $\{x^n\}_{n=0}^\infty$ and $\{y^n\}_{n=0}^\infty$ by the induction. First, let us set $x^0 = b_i$ and $y^0 = \overline{y}_i$. Suppose we have constructed two sequences $x^0, \ldots, x^k$ and $y^0, \ldots, y^k$. Now we will define x^{k+1} and y^{k+1} in the following way: if $(\frac{x^k+y^k}{2}, \mathbf{y}_{-i}) \succeq \mathbf{x}$, then let $x^{k+1} \equiv \frac{x^k+y^k}{2}$ and $y^{k+1} \equiv y^k$; and if $(\frac{x^k+y^k}{2}, \mathbf{y}_{-i}) \prec \mathbf{x}$, then let $x^{k+1} \equiv x^k$ and $y^{k+1} \equiv \frac{x^k+y^k}{2}$. Note that $\{x^k\}_{k=1}^\infty$ is a non-increasing sequence, $\{y^k\}_{k=1}^\infty$ is a non-decreasing sequence, and $\lim_{k\to\infty} x^k - y^k = \lim_{k\to\infty} \frac{x^0-y^0}{2^k} = 0$. So there exists $y^* \in X_i$ such that $\lim_{k\to\infty} x^k = \lim_{k\to\infty} y^k = y^* \in X_i$. Moreover, by the construction, we have $(x^k, \mathbf{y}_{-i}) \succeq \mathbf{x}$ and $\mathbf{x} \succ (y^k, \mathbf{y}_{-i})$ for all k. By continuity, we have $(y^*, \mathbf{y}_{-i}) \succeq \mathbf{x} \succeq (y^*, \mathbf{y}_{-i})$. Therefore, $\mathbf{x} \sim (y^*, \mathbf{y}_{-i})$. ■

We frequently use the following corollary of Lemma 6.

Corollary 2 *For any i, j with $i \neq j$ and $x_i, y_i \in X_i$ with $x_i \geq y_i$, there exists $x_j \in X_j$ such that $(x_i, x_j, \boldsymbol{b}_{-i,-j}) \sim (y_i, b_j, \boldsymbol{b}_{-i,-j})$. .*

Proof of Corollary 2 Since $x_i \geq y_i$, we have $(x_i, b_j, \mathbf{b}_{-i,-j}) \succeq (y_i, b_j, \mathbf{b}_{-i,-j})$ by strong monotonicity. By Lemma 6, there exists $x_j \in X_j$ such that $(x_i, x_j, \mathbf{b}_{-i,-j}) \sim (y_i, b_j, \mathbf{b}_{-i,-j})$. ■

Now we will prove a lemma which shows that there exist continuous distance-based functions consistent (in some sense) with $\succsim$.

Lemma 7 *If $\succsim$ is regular and satisfies Separability, then there exist continuous distance-based functions $\{f_i\}_{i=1}^n$ such that for any $\boldsymbol{x}, \boldsymbol{y} \in X$, i, and $x_i', y_i' \in X_i$, if $\boldsymbol{x} \sim \boldsymbol{y}$ and $f_i(x_i, y_i) = f_i(x_i', y_i')$, then $(x_i', \boldsymbol{x}_{-i}) \sim (y_i', \boldsymbol{y}_{-i})$.*

Proof of Lemma 7 First, let us construct distance-based functions.

For each $i < n$, let f_i be a function such that for any $x_i, y_i \in X_i$ with $x_i \geq y_i$, $f_i(x_i, y_i) = \frac{b_n - x_n}{b_n - a_n}$ and $f_i(y_i, x_i) = -\frac{b_n - x_n}{b_n - a_n}$ whenever $(x_i, x_n, \mathbf{b}_{-i,-n}) \sim (y_i, b_n, \mathbf{b}_{-i,-n})$. By Corollary 2, f_i is well-defined. Moreover, by continuity and strong monotonicity, f_i is also continuous and strictly increasing in its first argument. Therefore, f_i is a continuous distance-based function.

Now, we will construct f_n. By Corollary 2, for any $x_n, y_n \in X_n$ with $x_n \geq y_n$, there exists $x_1 \in X_1$ such that $(x_1, x_n, \mathbf{b}_{-1,-n}) \sim (y_n, \mathbf{b}_{-n})$. Let f_n be a function such that for any $x_n, y_n \in X_n$ with $x_n \geq y_n$, $f_n(x_n, y_n) = f_1(b_1, x_1)$ and $f_n(y_n, x_n) = f_1(x_1, b_1)$ whenever $(x_1, x_n, \mathbf{b}_{-1,-n}) \sim (y_n, \mathbf{b}_{-n})$. Similarly, f_n is well-defined, continuous, and strictly increasing in its first argument. Also, note that by the definition of f_1, $f_n(b_n, x_n) = \frac{b_n - x_n}{b_n - a_n}$ for all $x_n \in X_n$.

Now, we will prove that we constructed desired distance-based functions. Take any $\mathbf{x}, \mathbf{y} \in X$, i, and $x_i', y_i' \in X_i$ such that $\mathbf{x} \sim \mathbf{y}$ and $f_i(x_i, y_i) = f_i(x_i', y_i')$. We shall prove that $(x_i', \mathbf{x}_{-i}) \sim (y_i', \mathbf{y}_{-i})$. Without loss of generality, suppose $x_i \geq y_i$. We consider two cases.

Case 1: $i < n$.

Take some $\overline{x}_n \in X_n$ such that $f_i(x_i, y_i) = f_i(x_i', y_i') = \frac{b_n - \overline{x}_n}{b_n - a_n}$. By the definition of f_i, we obtain $(x_i, \overline{x}_n, \mathbf{b}_{-i,-n}) \sim (y_i, \mathbf{b}_{-i})$ and $(x_i', \overline{x}_n, \mathbf{b}_{-i,-n}) \sim (y_i', \mathbf{b}_{-i})$. Since $\mathbf{x} \sim \mathbf{y}$, $(x_i, \overline{x}_n, \mathbf{b}_{-i,-n}) \sim (y_i, \mathbf{b}_{-i})$, and $(x_i', \overline{x}_n, \mathbf{b}_{-i,-n}) \sim (y_i', \mathbf{b}_{-i})$, by Separability, $(x_i', \mathbf{x}_{-i}) \sim (y_i', \mathbf{y}_{-i})$ holds.

Case 2: $i = n$.

Take some $\overline{x}_1 \in X_1$ such that $f_n(x_n, y_n) = f_n(x_n', y_n') = f_1(b_1, \overline{x}_1)$. By the definition of f_n, we obtain $(\overline{x}_1, x_n, \mathbf{b}_{-1,-n}) \sim (y_n, \mathbf{b}_{-n})$ and $(\overline{x}_1, x_n', \mathbf{b}_{-1,-n}) \sim (y_n', b_{-n})$.

Since $\mathbf{x} \sim \mathbf{y}$, $(\overline{x}_1, x_n, \mathbf{b}_{-1,-n}) \sim (y_n, \mathbf{b}_{-n})$, and $(\overline{x}_1, x'_n, \mathbf{b}_{-1,-n}) \sim (y'_n, \mathbf{b}_{-n})$, by Separability, $(x'_n, \mathbf{x}_{-n}) \sim (y'_n, \mathbf{y}_{-n})$ holds. ■

Corollary 3 *For any i, the range $f_i(X_i^2) = (-1, 1)$.*

Proof of Corollary 3 When $i = n$, note that $f_n(b_n, x_n) = \frac{b_n - x_n}{b_n - a_n}$ can take any number in $[0, 1)$ by appropriately choosing x_n. Therefore, $f_n(X_n^2) = (-1, 1)$. Now suppose $i < n$. By Corollary 2, for any $x_n \in X_n$, there exists $x_i \in X_i$ such that $(x_i, b_n, \mathbf{b}_{-i,-n}) \sim (b_i, x_n, \mathbf{b}_{-i,-n})$. By the construction of f_i, we have $f_i(b_i, x_i) = f_n(b_n, x_n) = \frac{b_n - x_n}{b_n - a_n}$. Therefore, $f_i(X_i^2) = (-1, 1)$.

■

Finally, we will prove Theorem 5. We shall find an strictly increasing and continuous function W such that $W(t_i, \mathbf{0}_{-i}) = t_i$ and $W(\mathbf{t}) = W(\mathbf{t}_{-n}, 0_n) + t_n$ for any $t \in (-1, 1)^n$. We construct W in the following way. First, for any $\mathbf{t}_{-n} \in (-1, 1)^{n-1}$, we construct $W(\mathbf{t}_{-n}, 0_n)$. Then for any $\mathbf{t} \in (-1, 1)^n$, we set $W(\mathbf{t}) \equiv W(\mathbf{t}_{-n}, 0_n) + t_n$. In order to construct $W(\mathbf{t}_{-n}, 0_n)$, for each $\mathbf{t}_{-n} \in (-1, 1)^{n-1}$, we will find $\overline{x}_n, \overline{y}_n \in X_n$ (to be described later) and set $W(\mathbf{t}_{-n}, 0_n) \equiv f_n(\overline{y}_n, \overline{x}_n)$.

Take $\mathbf{t}_{-n} \in (-1, 1)^{n-1}$. We will construct $\overline{x}_n$ and $\overline{y}_n$ by the following two claims.

Claim 1: For any $j < n$, there exist $\overline{x}_j, \overline{y}_j \in X_j$ such that $t_j = f_j(\overline{x}_j, \overline{y}_j)$.

By Corollary 3, there exists $z_j \in X_j$ such that $f_j(b_j, z_j) = |t_j|$. Then we have $t_j = f_j(\overline{x}_j, \overline{y}_j)$ by setting $(\overline{x}_j, \overline{y}_j) = (b_j, z_j)$ when $t_j \geq 0$ and $(\overline{x}_j, \overline{y}_j) = (z_j, b_j)$ when $t_j < 0$.

Now suppose we have constructed $\overline{x}_1, \ldots, \overline{x}_{n-1}, \overline{y}_1, \ldots, \overline{y}_{n-1}$ by Claim 1.

Claim 2: There exist $\overline{x}_n, \overline{y}_n \in X_n$ such that $\overline{\mathbf{x}} \sim \overline{\mathbf{y}}$.

Now let $W(\mathbf{t}_{-n}, 0_n) \equiv f_n(\overline{y}_n, \overline{x}_n)$ and $W(\mathbf{t}) \equiv W(\mathbf{t}_{-n}, 0_n) + t_n$ for any $\mathbf{t} \in (-1, 1)^n$. By this construction, W is continuous and strictly increasing in all its arguments.

Now we shall prove that $\mathbf{x} \succeq \mathbf{y}$ if and only if $W((f_i(x_i, y_i))_i) \geq 0$. Since W is continuous and strictly increasing in all its arguments and each f_i is continuous and strictly increasing in its first argument, it is enough to prove that $\mathbf{x} \sim \mathbf{y}$ implies $W((f_i(x_i, y_i))_i) = 0$.

Take any $\mathbf{x}, \mathbf{y} \in X$ with $\mathbf{x} \sim \mathbf{y}$. First, let $\mathbf{t}_{-n} = ((f_i(x_i, y_i))_{i<n})$. Then by the above procedure that involves Claims 1-2, we find $\overline{\mathbf{x}}, \overline{\mathbf{y}} \in X$ such that $f_j(\overline{x}_j, \overline{y}_j) = t_j$ for all $j < n$ and $\overline{\mathbf{x}} \sim \overline{\mathbf{y}}$. By the construction of W, we have $W((f_i(x_i, y_i))_i) = W((f_i(x_i, y_i))_i) = W((f_i(x_i, y_i))_{i<n}, 0_n) + f_n(x_n, y_n) = W(\mathbf{t}_{-n}, 0_n) + f_n(x_n, y_n) = f_n(\overline{y}_n, \overline{x}_n) + f_n(x_n, y_n)$.

Since $f_1(x_1, y_1) = f_1(\overline{x}_1, \overline{y}_1) = t_1$, by Lemma 7, $\mathbf{x} \sim \mathbf{y}$ implies $(\overline{x}_1, \mathbf{x}_{-1}) \sim (\overline{y}_1, \mathbf{y}_{-1})$. Similarly, since $f_2(x_2, y_2) = f_2(\overline{x}_2, \overline{y}_2) = t_2$, by Lemma 7, $(\overline{x}_1, \mathbf{x}_{-1}) \sim (\overline{y}_1, \mathbf{y}_{-1})$ implies $(\overline{x}_1, \overline{x}_2, \mathbf{x}_{-1,-2}) \sim (\overline{y}_1, \overline{y}_2, \mathbf{y}_{-1,-2})$. Since $f_i(x_i, y_i) = f_i(\overline{x}_i, \overline{y}_i) = t_i$ for each $i < n$, by repeating this argument $n-1$ times, we will obtain that $(\overline{\mathbf{x}}_{-n}, x_n) \sim (\overline{\mathbf{y}}_{-n}, y_n)$. Moreover, since $\overline{\mathbf{x}} \sim \overline{\mathbf{y}}$, by Lemma 7, $(\overline{\mathbf{x}}_{-n}, x_n) \sim (\overline{\mathbf{x}}_{-n}, x_n) \sim (\overline{\mathbf{y}}_{-n}, y_n)$ implies $f_n(y_n, x_n) = f_n(\overline{y}_n, \overline{x}_n)$. In other words, $W((f_i(x_i, y_i))_i) = f_n(\overline{y}_n, \overline{x}_n) + f_n(x_n, y_n) = 0$.

Lastly, we show that W is an aggregator; that is, $W(t_i, \mathbf{0}_{-i}) = t_i$ for any $t_i \in (-1, 1)$. By the construction of W, it is obvious when $i = n$. Now take any $i < n$ and $t_i \in [0, 1)$. By Corollary 3, there exist $x_i, y_i \in X_i$ such that $f_i(x_i, y_i) = -t_i = f_n(b_n(1 - t_i) + a_n t_i, b_n)$. By the construction of f_i, we have $(x_i, \mathbf{b}_{-i}) \sim (y_i, b_n(1 - t_i) + a_n t_i, \mathbf{b}_{-i,-n})$. Since $W(\mathbf{t}) = W(\mathbf{t}_{-n}, 0_n) + t_n$ for each $\mathbf{t} \in (-1, 1)^n$, we obtain $0 = W(f_i(y_i, x_i), f_n(b_n(1 - t_i) + a_n t_i, b_n), \mathbf{0}_{-i,-n}) = W(f_i(y_i, x_i), \mathbf{0}_{-i}) + f_n(b_n(1 - t_i) + a_n t_i, b_n) = W(t_i, \mathbf{0}_{-i}) - t_i$. A similar argument works for any $t_i \in (-1, 0]$.

Uniqueness: Suppose there are two sets of functions $(W, \{f_i\}_{i=1}^n)$ and $(W', \{f'_i\}_{i=1}^n)$ such that $W(\mathbf{t}) = W(\mathbf{t}_{-n}, 0_n) + t_n$ and $W'(\mathbf{t}) = W'(\mathbf{t}_{-n}, 0_n) + t_n$ for any $\mathbf{t} \in (-1, 1)^n$ and $f_n(b_n, x_n) = f'_n(b_n, x_n) = \frac{b_n - x_n}{b_n - a_n}$ for any $x_n \in X_n$ that satisfy (3.2). We shall prove that $W = W'$ and $f_i = f'_i$ for any i.

Take any $i < n$ and $x_i, y_i \in X_i$ with $x_i > y_i$. By Corollary 2, there exists $x_n \in X_n$ such that $(x_i, x_n, \mathbf{b}_{-i,-n}) \sim (y_i, \mathbf{b}_{-i})$. By (3.2), we have

$$\begin{aligned} 0 &= W(f_i(x_i, y_i), \mathbf{0}_{-i,-n}, f_n(x_n, b_n)) \\ &= W(f_i(x_i, y_i), \mathbf{0}_{-i}) + f_n(x_n, b_n), \text{ (since } W(\mathbf{t}) = W(\mathbf{t}_{-n}, 0_n) + t_n\text{)}, \\ &= f_i(x_i, y_i) - \frac{b_n - x_n}{b_n - a_n}, \text{ (by the definition of aggregator).} \end{aligned}$$

Similarly,

$$W'\left(f_i'(x_i, y_i), \mathbf{0}_{-i,-n}, f_n'(x_n, b_n)\right) = f_i'(x_i, y_i) - \frac{b_n - x_n}{b_n - a_n} = 0.$$

Then we will obtain that $f_i(x_i, y_i) = f_i'(x_i, y_i) = \frac{b_n - x_n}{b_n - a_n}$. Therefore, $f_i = f_i'$ for any $i < n$.

Now we will prove that $f_n = f_n'$. Take any $x_n, y_n \in X_n$ with $x_n < y_n$. By Corollary 2, there exists $y_1 \in X_1$ such that $(x_n, \mathbf{b}_{-n}) \sim (y_1, y_n, \mathbf{b}_{-1,-n})$. By (3.2),

$$\begin{aligned} 0 &= W(f_1(b_1, y_1), \mathbf{0}_{-1,-n}, f_n(x_n, y_n)) \\ &= W(f_1(b_1, y_1), \mathbf{0}_{-1}) + f_n(x_n, y_n), \text{ (since } W'(\mathbf{t}) = W'(\mathbf{t}_{-n}, 0_n) + t_n) \\ &= f_1(b_1, y_1) + f_n(x_n, y_n), \text{ (by the definition of aggregator)} \end{aligned}$$

and

$$W'(f_1'(b_1, y_1), \mathbf{0}_{-1,-n}, f_n'(x_n, y_n)) = f_1'(b_n, y_1) + f_n'(x_n, y_n) = 0.$$

Since $f_1 = f_1'$, we obtain that $f_n(y_n, x_n) = f_n'(y_n, x_n) = f_1(b_1, y_1)$. Therefore, $f_n = f_n'$.

Finally, we shall prove that $W = W'$. Since $W(\mathbf{t}) = W(\mathbf{t}_{-n}, 0_n) + t_n$ and $W'(\mathbf{t}) = W'(\mathbf{t}_{-n}, 0_n) + t_n$, it is enough to prove that $W(\mathbf{t}_{-n}, 0_n) = W'(\mathbf{t}_{-n}, 0_n)$ for any $\mathbf{t}_{-n} \in (-1, 1)^{n-1}$.

Now take any $\mathbf{t}_{-n} \in (-1, 1)^{n-1}$. For each $i < n$, there exist $x_i, y_i \in X_i$ such that $t_i = f_i(x_i, y_i) = f_i'(x_i, y_i)$ by Claim 1. Moreover, by Claim 2, there exist $x_n, y_n \in X_n$ such that $\mathbf{x} \sim \mathbf{y}$. Let $t_n = f_n(x_n, y_n) = f_n'(x_n, y_n)$. Therefore,

$$\begin{aligned} W(\mathbf{t}_{-n}, 0) &= W(\mathbf{t}_{-n}, t_n) - t_n = 0 - t_n, \text{ (by } \mathbf{x} \sim \mathbf{y} \text{ and (3.2))}, \\ &= W'(\mathbf{t}_{-n}, t_n) - t_n, \text{ (by } \mathbf{x} \sim \mathbf{y} \text{ and (3.2))}, \\ &= W'(\mathbf{t}_{-n}, 0). \end{aligned}$$

Therefore, $W = W'$.

Appendix D

APPENDIX TO CHAPTER 4

D.1 Proof of Theorem 6

D.1.1 Necessity

We start by proving the converse statement. Let $(u, \succsim)$ be a regular PALM in which $\succsim$ satisfies Richness. Let $\succsim^*$ be derived revealed perception priority from $\rho_{(u,\succsim)}$. We shall first prove that $\succsim^* = \succsim$. The next lemma is useful throughout this section.

Lemma 8 *If $c \succ a \succ b$, or $a \succ b \succ c$, then* $\dfrac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} < \dfrac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})}$.

Proof: Let $a \succ b$.

Case 1: $c \succ a \succ b$. Since $b \not\sim c$,

$$\begin{aligned}
\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} \Big/ \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} &= \frac{\left(\dfrac{\mu(a,\{a,b,c\})(1-\mu(c,\{a,b,c\}))}{\mu(b,\{a,b,c\})(1-\mu(c,\{a,b,c\}))(1-\mu(a,\{a,b,c\}))}\right)}{\left(\dfrac{\mu(a,\{a,b\})}{\mu(b,\{a,b\})(1-\mu(a,\{a,b\}))}\right)} \\
&= \frac{(1-\mu(a,\{a,b\}))}{(1-\mu(a,\{a,b,c\}))}\left[\frac{u(a)}{u(b)} \Big/ \frac{u(a)}{u(b)}\right] < 1,
\end{aligned}$$

where the last strict inequality is by Luce's regularity on μ; that is, $\mu(a,\{a,b\}) = \frac{u(a)}{u(a)+u(b)+u(x_0)} > \mu(a,\{a,b,c\}) = \frac{u(a)}{u(a)+u(b)+u(c)+u(x_0)}$.

Case 2: $a \succ b \succ c$. Since $b \not\sim c$,

$$\begin{aligned}
\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} \Big/ \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} &= \frac{\mu(a,\{a,b,c\})}{\mu(b,\{a,b,c\})(1-\mu(a,\{a,b,c\}))} \Big/ \frac{\mu(a,\{a,b\})}{\mu(b,\{a,b\})(1-\mu(a,\{a,b\}))} \\
&= \frac{1-\mu(a,\{a,b\})}{1-\mu(a,\{a,b,c\})} < 1;
\end{aligned}$$

where the last strict inequality is by Luce's regularity on μ. ■

First, we prove $a \sim b$ if and only if $a \sim^* b$. Then, we prove $a \succ b$ if and only if $a \succ^* b$.

Lemma 9 *$a \sim b$ if and only if $a \sim^* b$.*

Proof of Lemma 9:

Step 1: If $a \sim b$ then $a \sim^* b$.

Proof of Step 1: Fix $c \in X$ to show $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} \Big/ \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = 1$.

Case 1: $a \sim b \succsim c$.

$$\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} \Big/ \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\mu(a,\{a,b,c\})}{\mu(b,\{a,b,c\})} \Big/ \frac{\mu(a,\{a,b\})}{\mu(b,\{a,b\})} = \frac{u(a)}{u(b)} \Big/ \frac{u(a)}{u(b)} = 1.$$

Case 2: $c \succ a \sim b$.

$$\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} \Big/ \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\mu(a,\{a,b,c\})(1-\mu(c,\{a,b,c\}))}{\mu(b,\{a,b,c\})(1-\mu(c,\{a,b,c\}))} \Big/ \frac{\mu(a,\{a,b\})}{\mu(b,\{a,b\})} =$$

$$= \frac{u(a)}{u(b)} \Big/ \frac{u(a)}{u(b)} = 1.$$

■

Step 2: If $a \succ b$, then $a \nsim^0 b$.

Proof of Step 2: By Richness, there is c with $c \succ a \succ b$ or $a \succ b \succ c$. In either case, by Lemma 8, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} < \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})}$. Hence, $a \nsim^0 b$. ■

Step 3: If $a \succsim^0 b$, then $a \succsim b$.

Proof of Step 3: We show that if $a \not\succsim b$ then $a \not\succsim^0 b$. Let $a \not\succsim b$. Then by completeness, $b \succ a$. By Richness, there is c with $c \succ b \succ a$ or $b \succ a \succ c$. Suppose without loss of generality that $c \succ b \succ a$. By Lemma 8, we have $\frac{\rho(b,\{a,b,c\})}{\rho(a,\{a,b,c\})} < \frac{\rho(b,\{a,b\})}{\rho(a,\{a,b\})}$. Moreover, since $c \succ b$ and $c \succ a$, Step 2 shows that $c \nsim^0 a$ and $c \nsim^0 b$. Hence, $b \succ^0 a$, so that $a \not\succsim^0 b$. ■

Step 4: If $a \sim^* b$ then $a \sim b$.

Proof of Step 4: Let $a \sim^* b$. By the definition of $\sim^*$, $a \succsim^* b$ and $b \succsim^* a$. Then $a \succsim^* b$ implies that there exist $c_1, \dots, c_k$ such that $a = c_1 \succsim^0 c_2 \succsim^0 \dots \succsim^0 c_k = b$. By Step 3 and the transitivity of $\succsim$, we have that $a \succsim b$. Similarly, $b \succsim^* a$ implies that $b \succsim a$. Thus $a \sim b$. ■

In the following, we prove that $a \succ b$ if and only if $a \succ^* b$.

Lemma 10 *If $a \succ^* b$ then $a \succ b$.*

Proof:

Let $a \succ^* b$. It suffices to consider the following two cases.

Case 1: $a \succ^0 b$. Suppose, towards a contradiction, $a \not\succ b$. By the completeness of $\succsim$, $b \succsim a$. Note that $a \succ^0 b$ implies $a \nsim^0 b$, so $a \nsim b$ by Lemma 9. Then $b \succ a$. By Richness there is c such that $c \succ b \succ a$ or $b \succ a \succ c$. In either case, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} / \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > 1$ by Lemma 8, in contradiction with $a \succ^0 b$.

Case 2: By the definition of $\succeq^*$, there exist $c_1, \ldots, c_k \in X$ such that $a \succeq^0 c_1 \succeq^0 \cdots \succeq^0 c_k \succeq^0 b$ (at least one strict relation). Then, by Lemma 9 and Case 1, $a \succeq c_1 \succeq \cdots \succeq c_k \succeq b$ (at least one strict relation). Hence, by transitivity, $a \succ b$.

■

The next lemma shows the converse.

Lemma 11 *If $a \succ b$ then $a \succ^* b$.*

Proof:

Since X is countable, by Richness, we can write $X = \cup_{i \in \mathbb{Z}} X_i$ such that for any $x, x' \in X_i$ and $y \in X_j$ with $i < j$, $x \succ y$ and $x \sim x'$.

Let $a \succ b$. There exists $i, j \in \mathbb{Z}$ with $i < j$ such that $a \in X_i$ and $b \in X_j$.

Case 1: $j = i + 1$. It suffices to show that $a \succ^0 b$. Take any $c \in X$ such that $a \nsim^0 c$ and $b \nsim^0 c$. By Lemma 9, $a \nsim c$ and $b \nsim c$. Therefore, $c \in X \setminus X_i \cup X_{i+1}$. That is, either $c \succ a$ or $b \succ c$. Since $a \succ b$, then $c \succ a \succ b$ or $a \succ b \succ c$. In either case, by Lemma 8, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} / \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} < 1$. Thus $a \succ^0 b$. Hence, $a \succ^* b$.

Case 2: $j > i + 1$. For any $k \in \mathbb{Z}$ with $i < k < j$, let us take $c_k \in X_k$. By the argument in Case 1, $a \succ^0 c_{i+1} \succ^0 \cdots \succ^0 c_{j-1} \succ^0 b$. Therefore, $a \succ^* b$. ■

D.1.2 Sufficiency

In this section, we prove sufficiency. Choose a nondegenerate stochastic choice function ρ that satisfies the axioms in the theorem. Let $\succsim^*$ be the derived revealed perception priority.

Step 1: there exists $u : X_0 \to \mathbb{R}$ such that $q(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a') + u(x_0)}$.

Proof of Step 1: Since q satisfies Luce's IIA and $\sum_{a \in A_0} q(a, A) = 1$, by Luce's theorem (Luce 1959), there exists $u : X \to \mathbb{R}$ such that $q(a, A) = \frac{u(a)}{\sum_{a' \in A} u(a') + u(x_0)}$. Since ρ is nondegenerate, $1 > \rho(a, A) > 0$ for all $a \in A$. Remember that $A_a = \{b \in A | b \succ^* a\}$. Since $a \notin A_a$, $1 - \rho(A_a, A) > 0$. Therefore, $u(a) > 0$ for any $a \in A$. ■

Step 2: $\rho = \rho_{(u, \succsim^*)}$.

Proof of Step 2: Choose any $A \in \mathscr{A}$. Since $\succsim^*$ is a weak order, therefore the indifference relation $\sim^*$ is transitive. Then, the set of equivalence classes $A/_{\succsim^*}$ is well defined and finite. That is, there exists a partition $\{\alpha^1, \alpha^2, \ldots \alpha^k\}$ of A such that $a_j \succ^* a_i$ for all $a_i \in \alpha^i$ and $a_j \in \alpha^j$ with $j > i$ and $a_i \sim^* a_{i'}$ for all $a_i, a_{i'} \in \alpha^i$.

Define $p_i \equiv \rho(\alpha^i, A) = \sum_{a' \in \alpha^i} \rho(a', A)$. Then for $a \in \alpha^i$, $q(a, A) = \frac{\rho(a,A)}{1 - \sum_{j>i} p_j}$. Therefore,

$$\sum_{a \in \alpha^i} q(a, A) = \sum_{a \in \alpha^i} \frac{\rho(a, A)}{1 - \sum_{j>i} p_j} = \frac{\sum_{a \in \alpha^i} \rho(a, A)}{1 - \sum_{j>i} p_j} = \frac{p_i}{1 - \sum_{j=i+1}^k p_j}.$$

Hence,

$$1 - \sum_{a \in \alpha^i} q(a, A) = 1 - \frac{p_i}{1 - \sum_{j=i+1}^k p_j} = \frac{1 - \sum_{j=i+1}^k p_j - p_i}{1 - \sum_{j=i+1}^k p_j} = \frac{1 - \sum_{j=i}^k p_j}{1 - \sum_{j=i+1}^k p_j}.$$

Therefore, for any $s \in \{1, \ldots, k\}$,

$$\prod_{i=s+1}^k (1 - \sum_{a \in \alpha^i} q(a, A)) = \prod_{i=s+1}^k \frac{1 - \sum_{j=i}^k p_j}{1 - \sum_{j=i+1}^k p_j} = \frac{1 - \sum_{j=s+1}^k p_j}{1} = 1 - \rho(A_a, A).$$

For all $a \in A$ and $A \in \mathscr{A}$, define $\mu(a, A) = q(a, A)$.

Choose $a \in A$. Without loss of generality assume that $a \in \alpha^s$. Then, $\rho(a, A) = q(a, A)(1 - \rho(A_a, A)) = \mu(a, A)(1 - \rho(A_a, A)) = \mu(a, A) \prod_{i=s+1}^k (1 - \sum_{a' \in \alpha^i} \mu(a', A)) \equiv \rho_{(u, \succsim^*)}(a, A)$. ■

D.2 Relation to Manzini and Mariotti

The model of Manzini and Mariotti (2014) is specified by a probability measure g on X, and a linear order $\succ_M$. Their representation is then

$$\rho(a, A) = g(a) \prod_{a' \succ_M a} (1 - g(a')).$$

Superficially, this representation looks similar to ours, but it is actually very different: It is incompatible with our model, in the sense that the set of stochastic choices that satisfy our model is disjoint from the set of stochastic choices in Manzini and Mariotti's model. We now proceed to prove this fact.

Let ρ have a Manzini and Mariotti (2014) representation as above and let X have at least three elements. Suppose, towards a contradiction that it also has a representation using our model.

We are going to prove that the two models differ in a strong sense, because we are going to show that there is no subset of X of three elements on which the two models can coincide.

Let $a, b, c \in X$. The preference relation $\succ_M$ is a linear order. Suppose, without loss of generality, that $a \succ_M b \succ_M c$. Given the Manzini-Mariotti representation, then

$$\rho(a, \{a, b, c\}) = \rho(a, \{a, b\}) = \rho(a, \{a, c\}) = g(a),$$

and

$$\rho(b, \{a, b, c\}) = \rho(b, \{a, b\}) = g(b)(1 - g(a)).$$

We have assumed that ρ has a PALM representation given by some $(u, \succsim)$. Now consider how a, b, c are ordered by $\succsim$.

There are seven cases to consider; each one of these cases end in a contradiction.

1. $a \succsim b$, $a \succsim c$, and $b \nsim c$: By Regularity, since $b \nsim c$, $\rho(a, \{a, b, c\}) = q(a, \{a, b, c\}) < \rho(a, \{a, b\}) = q(a, \{a, b\})$.

2. $b \succsim a$, $b \succsim c$ and $a \nsim c$: By Regularity, since $a \nsim c$, $\rho(b, \{a, b, c\}) = q(b, \{a, b, c\}) < \rho(b, \{a, b\}) = q(b, \{a, b\})$.

3. $c \succ a \succsim b$: By Regularity, $\rho(a, \{a, b, c\}) = q(a, \{a, b, c\})(1 - q(c, \{a, b, c\})) < q(a, \{a, b, c\}) \leq q(a, \{a, b\}) = \rho(a, \{a, b\})$.

4. $a \succ b \sim c$: By Regularity, since $\rho(a, \{a, b, c\}) = q(a, \{a, b, c\}) = \rho(a, \{a, b\}) = q(a, \{a, b\})$ and $q(b, \{a, b, c\}) < q(b, \{a, b\})$ because $a \nsim c$, $\rho(b, \{a, b, c\}) = q(b, \{a, b, c\})(1 - q(a, \{a, b, c\})) < \rho(b, \{a, b\}) = q(b, \{a, b\})(1 - q(a, \{a, b\}))$.

5. $b \succ a \sim c$: By Regularity, $\rho(a, \{a, b, c\}) = q(a, \{a, b, c\})(1 - q(b, \{a, b, c\})) < q(a, \{a, b, c\}) \leq q(a, ac) = \rho(a, ac)$.

6. $c \succ b \succ a$: By Regularity, $\rho(b,\{a,b,c\}) = q(b,\{a,b,c\})(1-q(c,\{a,b,c\})) < q(b,\{a,b,c\}) \leq \rho(b,\{a,b\}) = q(b,\{a,b\})$.

7. $a \sim b \sim c$: In this case, Luce's IIA cannot be violated in PALM. However, in Manzini and Marriott's Model, there is always at least one violation of Luce's IIA.

D.3 Proof of Proposition 10

Observation 11: For any PALM ρ,

$$\rho(x_0,\{x,y\}) > \rho^f(x_0,\{x,y\}) \text{ if and only if } u(x_0) > u^f(x_0)$$

Proof:

By a direct calculation, $\rho(x_0,\{x,y\}) = \frac{\left(u(x)+u(x_0)\right)\left(u(y)+u(x_0)\right)}{\left(u(x)+u(y)+u(x_0)\right)^2}$ and $\rho^f(x_0,\{x,y\}) = \frac{\left(u(x)+u^f(x_0)\right)\left(u(y)+u^f(x_0)\right)}{\left(u(x)+u(y)+u^f(x_0)\right)^2}$. Let $g(t) = \frac{\left(u(x)+t\right)\left(u(y)+t\right)}{\left(u(x)+u(y)+t\right)^2}$. Since $g'(t) = \frac{t\left(u(x)+u(y)\right)+u^2(x)+u^2(y)}{\left(u(x)+u(y)+t\right)^3}$, g is increasing in t when $t > -\frac{u^2(x)+u^2(y)}{u(x)+u(y)}$.

Now it is enough to prove that $u(x_0)$ is larger than $-\frac{u^2(x)+u^2(y)}{u(x)+u(y)}$. First, $\rho(y,\{x,y\}) = \frac{u(y)\left(u(y)+u(x_0)\right)}{\left(u(x)+u(y)+u(x_0)\right)^2} > 0$ implies that $u(x_0) > -u(y)$. Second,

$$\rho(x_0,\{x,y\}) = \frac{\left(u(x)+u(x_0)\right)\left(u(y)+u(x_0)\right)}{\left(u(x)+u(y)+u(x_0)\right)^2} \geq 0$$

and $u(x_0) > -u(y)$ imply $u(x_0) \geq -u(x)$. Then we obtain $u(x_0) > -\frac{u(x)+u(y)}{2} \geq -\frac{u^2(x)+u^2(y)}{u(x)+u(y)}$. ■

Observation 12: For any PALM ρ,

$$\rho(x_0,\{x,y,z\}) > \rho^f(x_0,\{x,y,z\}) \text{ if and only if } u(x_0) > u^f(x_0).$$

Proof: By a direct calculation,

$$\rho(x_0,\{x,y,z\}) = \frac{\left(u(x)+u(y)+u(x_0)\right)\left(u(x)+u(z)+u(x_0)\right)\left(u(y)+u(z)+u(x_0)\right)}{\left(u(x)+u(y)+u(z)+u(x_0)\right)^3}$$

and $p^f(x_0,\{x,y,z\}) = \frac{\left(u(x)+u(y)+u^f(x_0)\right)\left(u(x)+u(z)+u^f(x_0)\right)\left(u(y)+u(z)+u^f(x_0)\right)}{\left(u(x)+u(y)+u(z)+u^f(x_0)\right)^3}$.

Let $s(t) = \frac{\left(u(x)+u(y)+t\right)\left(u(x)+u(z)+t\right)\left(u(y)+u(z)+t\right)}{\left(u(x)+u(y)+u(z)+t\right)^3}$. Also, let $A \equiv u(x)+u(y)+u(z)$, $B \equiv u^2(x)+u^2(y)+u^2(z)+u(x)u(y)+u(x)u(z)+u(y)u(z)$, and $C \equiv u^3(x)+u^3(y)+u^3(z)+u^2(x)u(y)+u^2(x)u(z)+u^2(y)u(x)+u^2(y)u(z)+u^2(z)u(x)+u^2(z)u(y)+3u(x)u(y)u(z)$.

Then we obtain $s'(t) = \frac{t^2 \cdot A + 2t \cdot B + C}{(t+A)^4}$. Therefore, s is increasing when $t > -\frac{B-\sqrt{B^2-AC}}{A}$.

Now it is enough to prove that $u(x_0)$ is larger than $-\frac{B-\sqrt{B^2-AC}}{A}$. First, $\rho(y, \{x, y, z\}) = \frac{u(y)\left(u(y)+u(z)+u(x_0)\right)}{\left(u(x)+u(y)+u(z)+u(x_0)\right)^2} > 0$ implies $u(x_0) > -(u(y)+u(z))$. Second,

$\rho(z, \{x, y, z\}) = \frac{u(z)\left(u(y)+u(z)+u(x_0)\right)\left(u(x)+u(z)+u(x_0)\right)}{\left(u(x)+u(y)+u(z)+u(x_0)\right)^3} > 0$ and $u(x_0) > -(u(y)+u(z))$ imply $u(x_0) > -(u(x)+u(z))$. Lastly,

$$\rho(x_0, \{x, y, z\}) = \frac{(u(x)+u(y)+u(x_0))(u(x)+u(z)+u(x_0))(u(y)+u(z)+u(x_0))}{(u(x)+u(y)+u(z)+u(x_0))^3} \geq 0$$

implies that $u(x_0) \geq -(u(x)+u(y))$. When $u(x) = u(y) = u(z) = t$, it is obvious that $u(x_0) > -\frac{B-\sqrt{B^2-AC}}{A} = -2t$. Now it is enough to prove that

$$-\min\left(u(x)+u(z);\ u(x)+u(z);\ u(x)+u(z)\right) \geq -\frac{B-\sqrt{B^2-AC}}{A}.$$

Since the inequality is completely symmetric, without loss of generality, let us assume that $u(z) \geq u(y) \geq u(x)$. Now we shall prove that $\frac{B-\sqrt{B^2-AC}}{A} \geq u(x)+u(y)$.

$B-\sqrt{B^2-AC} \geq A(u(x)+u(y))$ if and only if $u^2(z)-u(x)u(y) \geq \sqrt{B^2-AC}$ if and only if

$u^4(z) - 2u^2(z)u(x)u(y) + u^2(x)u^2(y) \geq u^2(x)u^2(y) + u^2(y)u^2(z) + u^2(x)u^2(z) - u(x)u(y)u(z)(u(x)+u(y)+u(z))$; equivalently,

$$u^4(z)+u^2(x)u(y)u(z)+u^2(y)u(x)u(z) \geq u^2(x)u^2(z)+u^2(y)u^2(z)+u^2(z)u(x)u(y)$$

if and only if $u(z)(u(z)-u(y))(u(z)-u(x))(u(x)+u(y)+u(z)) \geq 0$.

■

D.4 Proof of Proposition 12

By direct calculations, we obtain

$\frac{\rho(x, \{x, y\})}{\rho(y, \{x, y\})} \Big/ \frac{\rho(x, \{x, y, z\})}{\rho(y, \{x, y, z\})} = 1 + \frac{u(x)\,u(z)}{(u(y)+u(x_0))(u(x)+u(y)+u(z)+u(x_0))}$ and

$$\frac{\rho^f(x,\{x,y\})}{\rho^f(y,\{x,y\})}\Big/\frac{\rho^f(x,\{x,y,z\})}{\rho^f(y,\{x,y,z\})} = 1+\frac{u(x)\,u(z)}{(u(y)+u^f(x_0))\,(u(x)+u(y)+u(z)+u^f(x_0))}.$$

Therefore, $\frac{\rho^f(x,\{x,y\})}{\rho^f(y,\{x,y\})}/\frac{\rho^f(x,\{x,y,z\})}{\rho^f(y,\{x,y,z\})} > \frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})}/\frac{\rho(x,\{x,y,z\})}{p(\rho,\{x,y,z\})}$ if and only if

$$1 > \frac{(u(y)+u^f(x_0))\,(u(x)+u(y)+u(z)+u^f(x_0))}{(u(y)+u(x_0))\,(u(x)+u(y)+u(z)+u(x_0))} \text{ iff } u(x_0) > u^f(x_0).$$

D.5 Proof of Proposition 13

First, we show that if $u(a) > u(b)$ then $a \succ b$. By way of contradiction, suppose $b \succ a$. By calculation, $\rho(a,\{a,b\}) = \frac{u(a)(u(a)+u(x_0))}{(u(a)+u(b)+u(x_0))^2}$, $\rho(b,\{a,b\}) = q(b,\{a,b\}) = \frac{u(b)}{u(a)+u(b)+u(x_0)}$, and $\rho(x_0,\{a,b\}) = 1 - \rho(b,\{a,b\}) - \rho(a,\{a,b\}) = \frac{(u(a)+u(x_0))(u(b)+u(x_0))}{(u(a)+u(b)+u(x_0))^2}$.

First consider the case when $\rho(a,\{a,b\}) = \min(\rho(a,\{a,b\}),\rho(b,\{a,b\}))$. Then $\rho(x_0,\{a,b\}) \geq \rho(a,\{a,b\})$ if and only if $\frac{(u(a)+u(\{a,b\}))(u(b)+u(x_0))}{(u(a)+u(b)+u(x_0))^2} \geq \frac{u(a)(u(a)+u(x_0))}{(u(a)+u(b)+u(x_0))^2}$ if and only if $u(b)+u(x_0) \geq u(a)$. Therefore, since $u(x_0) \leq 0$, $\rho(x_0,\{a,b\}) \geq \rho(a,\{a,b\})$ implies $u(b) \geq u(a)$. Contradiction.

Second consider the case when $\rho(b,\{a,b\}) = \min(\rho(a,\{a,b\}),\rho(b,\{a,b\}))$. Then $\rho(x_0,\{a,b\}) \geq \rho(b,\{a,b\})$ if and only if $\frac{(u(a)+u(x_0))(u(b)+u(x_0))}{(u(a)+u(b)+u(x_0))^2} \geq \frac{u(b)}{u(a)+u(b)+u(x_0)}$ if and only if $(u(a)+u(x_0))(u(b)+u(x_0)) \geq u(b)(u(a)+u(b)+u(x_0))$. Therefore, since $u(x_0) \leq 0$, $\rho(x_0,\{a,b\}) \geq \rho(b,\{a,b\})$ implies $(u(a)+u(x_0))(u(b)+u(x_0)) \geq (u(b)+u(x_0))(u(a)+u(b)+u(x_0))$, i.e., $u(a)+u(x_0) \geq u(a)+u(b)+u(x_0)$. Contradiction. Therefore, we proved that $a \succ b$.

Finally, we show that if $a \succ b$, then $u(a) > u(b)$. Suppose $u(b) > u(a)$. Then by the previous part, $u(b) > u(a)$ implies $b \succ a$. Contradiction.

D.6 Proof of Proposition 14

By calculation, we obtain $\rho(a,\{a,b\}) = q(a,\{a,b\}) = \frac{u(a)}{u(a)+u(b)+u(x_0)}$, $\rho(b,\{a,b\}) = q(b,\{a,b\})(1-q(a,\{a,b\})) = \frac{u(b)(u(b)+u(x_0))}{(u(a)+u(b)+u(x_0))^2}$, and $\rho(x_0,\{a,b\}) = 1-\rho(a,\{a,b\}) - \rho(b,\{a,b\}) = \frac{(u(a)+u(x_0))(u(b)+u(x_0))}{(u(a)+u(b)+u(x_0))^2}$.

First, $\rho(a,\{a,b\}) > \rho(x_0,\{a,b\})$ if and only if $\frac{u(a)}{u(a)+u(b)+u(x_0)} > \frac{(u(a)+u(x_0))(u(b)+u(x_0))}{(u(a)+u(b)+u(x_0))^2}$ if and only if $u(a)(u(a)+u(b)+u(x_0)) > (u(a)+u(x_0))(u(b)+u(x_0))$. Since $u(a)+u(b)+u(x_0) > u(b)+u(x_0)$, we obtain $(u(a)+u(x_0))(u(b)+u(x_0)) \geq (u(a)+u(x_0))(u(a)+u(b)+u(x_0))$. Therefore, $\rho(a,\{a,b\}) > \rho(x_0,\{a,b\})$. Second, $\rho(x_0,\{a,b\}) \geq \rho(b,\{a,b\})$ if and only if $\frac{(u(a)+u(x_0))(u(b)+u(x_0))}{(u(a)+u(b)+u(x_0))^2} \geq \frac{u(b)(u(b)+u(x_0))}{(u(a)+u(b)+u(x_0))^2}$ if and only if $u(a)+u(x_0) \geq u(b)$. Therefore, $\rho(x_0,\{a,b\}) \geq \rho(b,\{a,b\}) = \min(\rho(a,\{a,b\}),\rho(b,\{a,b\}))$.

D.7 Finite X

Here we consider the case where X is finite. In order to obtain the sufficiency part of Theorem 6, we will modify $\succsim^*$ in the following way:

Definition 22 *Let a and b be arbitrary elements in X.*

(i)

$$a \sim^0 b \quad \text{if} \quad \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} = \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})},$$

for all $c \in X$;

(ii)

$$aPb \quad \text{if} \quad \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})},$$

for all $c \in X$ such that $c \nsim^0 a$ and $c \nsim^0 b$, and if there is at least one such c.

(iii) $a \succ^0 b$ if aPb, but there is no $c_1, \ldots, c_k \in X$ such that $aPc_1P \cdots c_kPb$. We write $a \succsim^0 b$ if $a \sim^0 b$ or $a \succ^0 b$.

(iv) Define $\succsim^*$ be the transitive closure *of $\succsim^0$; that is, $a \succsim^* b$ if there exist $c_1, \ldots, c_k \in X$ such that*

$$a \succsim^0 c_1 \succsim^0 \cdots c_k \succsim^0 b.$$

The binary relation $\succsim^*$ is called the revealed perception priority *derived from ρ.*

Now we can prove Theorem 6 when X is finite. In particular, we use very weak version of Richness.

Richness*: There exist $a, b, c \in X$ such that $a \succ b \succ c$.

Theorem 10 *If a nondegenerate stochastic choice function ρ satisfies Weak Order and Hazard Rate IIA, then there is a PALM $(u, \succsim)$ such that $\succsim^* = \succsim$ and $\rho = \rho_{(u,\succsim)}$.*

Conversely, for a given PALM $(u, \succsim)$, if $\succsim$ satisfies Richness, then $\rho_{(u,\succsim)}$ satisfies Weak Order and Hazard Rate IIA, and $\succsim = \succsim^*$.*

Proof:

The proof of the sufficiency part of Theorem 10 is identical to that of Theorem 6. For the necessity part of Theorem 10, we want to prove lemmas similar to Theorem 6.

Lemma 8*: If $c \succ a \succ b$, or $a \succ b \succ c$, then $\frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > \frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})}$.

The proof is identical to the proof of Lemma 8 of Theorem 6.

Lemma 9*: $a \sim b$ if and only if $a \sim^* b$.

The proof is identical to the proof of Lemma 9 of Theorem 6.

Lemma 10*: If $a \succ^* b$, then $a \succ b$.

Let $a \succ^* b$. It suffices to consider the following two cases.

Case 1: $a \succ^0 b$. Suppose, towards a contradiction, $a \not\succ b$. By the completeness of $\succsim$, $b \succsim a$. Note that $a \succ^0 b$ implies $a \nsim^0 b$, so $a \nsim b$ by Lemma 9*. Then $b \succ a$.

By Richness*, there is c such that $c \succ b \succ a$ or $b \succ a \succ c$. In either case, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} / \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > 1$ by Lemma 8*, in contradiction with $a \succ^0 b$.

Case 2: There exist $c_1, \ldots, c_k \in X$ such that $a \succeq^0 c_1 \succeq^0 \cdots \succeq^0 c_k \succ^0 b$. Then, by the proof in Case 1, $a \succ c_1 \succ \cdots \succ c_k \succ b$. Hence, by transitivity, $a \succ b$.

Lemma 11*: If aPb, then $a \succ b$.

Let aPb. Suppose, towards a contradiction, $a \not\succ b$. By the completeness of $\succsim$, $b \succsim a$. Note that aPb implies $a \nsim^0 b$, so $a \nsim b$ by Lemma 9*. Then $b \succ a$.

By Richness*, there is c such that $c \succ b \succ a$ or $b \succ a \succ c$. In either case, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} / \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} > 1$ by Lemma 8*, in contradiction with aPb.

Lemma 12*: If $a \succ b$, then $a \succ^* b$.

Let $a \succ b$. To simplify the exposition, we use the following notation in this proof: $a \vdash b$ if $a \succ b$ and there is no $c \in X$ with $a \succ c \succ b$.

Case 1: $a \vdash b$. It suffices to show that $a \succ^0 b$. By Richness*, there exists c such that $a \nsim c$ and $b \nsim c$. By Lemma 9*, $a \nsim^0 c$ and $b \nsim^0 c$.

Choose any $d \in X$ such that $a \nsim^0 d$ and $b \nsim^0 d$. By Lemma 9*, $a \nsim d$ and $b \nsim d$. Since $a \vdash b$, it is not true that $a \succ d \succ b$. That is, either $d \succ a$ or $b \succ d$. Since $a \succ b$, then $d \succ a \succ b$ or $a \succ b \succ d$. In either case, by Lemma 8*, $\frac{\rho(a,\{a,b,c\})}{\rho(b,\{a,b,c\})} / \frac{\rho(a,\{a,b\})}{\rho(b,\{a,b\})} < 1$. Thus aPb.

We now shall prove that there is no $c_1, \ldots, c_k \in X$ such that $aPc_1P \ldots Pc_kPb$. By way of contradiction, suppose there exist $c_1, \ldots, c_k \in X$ such that $aPc_1P \ldots Pc_kPb$.

By Lemma 11*, $aPc_1P\ldots Pc_kPb$ implies $a \succ c_1 \succ \ldots \succ c_k \succ b$, in contradiction with $a \vdash b$. Thus $a \succ^0 b$. Hence, $a \succ^* b$.

Case 2: $a \nvdash b$. There exist $c_1, \ldots, c_k \in X$ such that $a \vdash c_1 \vdash \cdots \vdash c_k \vdash b$. By the argument in Case 1, $a \succ^0 c_1 \succ^0 \cdots \succ^0 c_k \succ^0 b$. Therefore, $a \succ^* b$.

■

D.8 A modification without the outside option

In this section we show that by modifying PALM, we can dispense with the outside option. In the modified model, whenever the agent chooses no alternative, she repeats the sequential procedure of PALM until she chooses some alternatives. This modified PALM is represented by the following representation:

$$\rho(a, A) = \frac{\mu(a, A) \prod_{\alpha \in A/\succsim:\, \alpha \succ a} (1 - \mu(A_\alpha, A))}{\sum_{b \in A} \mu(b, A) \prod_{\alpha \in A/\succsim:\, \alpha \succ b} (1 - \mu(A_\alpha, A))}$$

where

$$\mu(a, A) = \frac{u(a)}{\sum_{b \in A} u(b)}.$$

Let us now show that this modified PALM can rationalize the compromise and attraction effects. First, we show that this modified PALM can rationalize the compromise effect. In fact, we obtain the following observation which is very similar to Proposition 8 (note that Equation (D.1) is very similar to Equation (4.8)).

Observation 13: *When $x \succ y \succsim z$, $\rho_{(u,\succsim)}$ exhibits the compromise effect (i.e., (4.7)) if and only if $u(y) > u(x)$ and*

$$u(z) > \frac{u^2(x) - u^2(y) + u(x)u(y)}{u(y) - u(x)} \geq 0. \tag{D.1}$$

Proof of Observation 13: We have $\frac{\rho(x,\{x,y,z\})}{\rho(y,\{x,y,z\})} < 1 \leq \frac{\rho(x,\{x,y\})}{\rho(y,\{x,y\})}$ iff

$$\frac{\mu(x, \{x, y, z\})}{\mu(y, \{x, y, z\})(1 - \mu(x, \{x, y, z\}))} < 1 \leq \frac{\mu(x, \{x, y\})}{\mu(y, \{x, y\})(1 - \mu(x, \{x, y\}))} \text{ iff}$$

$$\frac{u(x)}{u(y)} \cdot \frac{u(x) + u(y) + u(z)}{u(y) + u(z)} < 1 \leq \frac{u(x)}{u(y)} \cdot \frac{u(x) + u(y)}{u(y)}.$$

By direct calculations, we obtain (D.1). ■

Similarly, the following observation which is very similar to Proposition 9 shows that the modified PALM can rationalize the attraction effect.

Observation 14: *If $x \succ y \succ z$ and $u(x)$ is large enough, then $\rho_{(u,\succsim)}$ exhibits the attraction effect (i.e., (4.9)).*

Proof of Observation 14: We have $\rho(y,\{x,y,z\}) > \rho(y,\{x,y\})$ iff

$$\frac{\mu(y,\{x,y,z\})(1-\mu(x,\{x,y,z\}))}{\mu(x,\{x,y,z\})+\mu(y,\{x,y,z\})(1-\mu(x,\{x,y,z\}))+\mu(z,\{x,y,z\})(1-\mu(x,\{x,y,z\}))(1-\mu(y,\{x,y,z\}))}$$

$$> \frac{\mu(y,\{x,y\})(1-\mu(x,\{x,y\}))}{\mu(x,\{x,y\})+\mu(y,\{x,y\})(1-\mu(x,\{x,y\}))} \text{ iff}$$

$$\frac{u(y)(1-\mu(x,\{x,y,z\}))}{u(x)+u(y)(1-\mu(x,\{x,y,z\}))+u(z)(1-\mu(x,\{x,y,z\}))(1-\mu(y,\{x,y,z\}))}$$

$$> \frac{u(y)(1-\mu(x,\{x,y\}))}{u(x)+u(z)(1-\mu(x,\{x,y\}))}.$$

By direct calculations, we obtain that $\rho(y,\{x,y,z\}) > \rho(y,\{x,y\})$ iff

$$u^2(x) > u(y)\big(u(y)+u(z)\big)\frac{u(x)+u(z)}{u(x)+u(y)+u(z)}.$$

Since $1 > \frac{u(x)+u(z)}{u(x)+u(y)+u(z)}$, if $u(x) > \sqrt{u(y)\big(u(y)+u(z)\big)}$, then we have $\rho(y,\{x,y,z\}) > \rho(y,\{x,y\})$. Therefore, when $u(x)$ is large enough, we can have the attraction effect.

■

The above two observations illustrate that the outside option does not really play a role in explaining the two effects, but the sequential procedure does.

Appendix E

APPENDIX TO CHAPTER 5

E.1 Proof of Theorem 7

Since the necessity part is obvious, we only prove the sufficiency part. Suppose p satisfies R-IIA and R-CI. We shall construct a utility function $u : X \times \mathbb{N} \to \mathbb{R}_{++}$ such that $p = p_{(u,R)}$. Without loss of generality, let $X = \{a_1, a_2, \ldots, a_n\}$ where $n \geq 3$ and $R(a_i, X) = i$ for each $i = 1, 2, \ldots, n$. Let us define u as follows.

First, let

$$u(a_1, 1) \equiv 1.$$

Second, for each i such that $2 \leq i \leq n-1$, let

$$u(a_i, 1) \equiv \frac{p(a_i, \{a_i, a_n\}) \cdot p(a_n, \{a_1, a_n\})}{p(a_n, \{a_i, a_n\}) \cdot p(a_1, \{a_1, a_n\})}.$$

Lastly, for any k and i such that $2 \leq i \leq n$ and $2 \leq k \leq i$, let

$$u(a_i, k) \equiv \frac{p(a_i, A(k,i))}{p(a_1, A(k,i))}$$

where $A(k,i) \equiv \{a_1, \ldots, a_{k-1}, a_i\}$.

Now we prove that we in fact constructed the desired u. In particular, we shall prove that for any $A = \{a_{i_1}, a_{i_2}, \ldots, a_{i_m}\}$ with $i_1 < i_2 < \ldots < i_m$,

$$p(a_{i_s}, A) = \frac{u(a_{i_s}, s)}{\sum_{k=1}^{m} u(a_{i_k}, k)} \text{ for all } s.$$

Let us consider two cases.

Case 1: When $i_1 = 1$. Take any s. By the construction of u, $u(a_{i_s}, s) = \frac{p(a_{i_s}, A(s,i_s))}{p(a_1, A(s,i_s))}$. Since $R(a_1, A) = R(a_1, A(s, i_s)) = 1$ and $R(a_{i_s}, A) = R(a_{i_s}, A(s, i_s)) = s$, by R-IIA,

$$\frac{p(a_{i_s}, A)}{p(a_1, A)} = \frac{p(a_{i_s}, A(s, i_s))}{p(a_1, A(s, i_s))} = u(a_{i_s}, s).$$

Therefore,

$$\frac{u(a_{i_s}, s)}{\sum_{k=1}^{m} u(a_{i_k}, k)} = \frac{\frac{p(a_{i_s}, A)}{p(a_1, A)}}{\sum_{k=1}^{m} \frac{p(a_{i_k}, A)}{p(a_1, A)}} = \frac{\frac{p(a_{i_s}, A)}{p(a_1, A)}}{\frac{1}{p(a_1, A)}} = p(a_{i_s},\ A).$$

Case 2: When $i_1 \geq 2$. If we can prove that $\frac{p(a_{i_s}, A)}{p(a_{i_1}, A)} = \frac{u(a_{i_s}, s)}{u(a_{i_1}, 1)}$ for each s, then we obtain

$$\frac{u(a_{i_s}, s)}{\sum_{k=1}^{m} u(a_{i_k}, k)} = \frac{\frac{p(a_{i_s}, A)}{p(a_{i_1}, A)} \cdot u(a_{i_1}, 1)}{\sum_{k=1}^{m} \frac{p(a_{i_k}, A)}{p(a_{i_1}, A)} \cdot u(a_{i_1}, 1)} = \frac{\frac{p(a_{i_s}, A)}{p(a_{i_1}, A)}}{\frac{1}{p(a_{i_1}, A)}} = p(a_{i_s}, A).$$

Take any s. By the construction of u, $u(a_{i_s}, s) = \frac{p(a_{i_s}, A(s, i_s))}{p(a_1, A(s, i_s))}$. Moreover, since $a_1 R a_{i_1} R a_n$ and $a_1 R a_{i_1} R a_{i_2}$, by R-CI,

$$\frac{p(a_{i_2}, \{a_1, a_{i_2}\})}{p(a_1, \{a_1, a_{i_2}\})} / \frac{p(a_{i_2}, \{a_{i_1}, a_{i_2}\})}{p(a_{i_1}, \{a_{i_1}, a_{i_2}\})} = \frac{p(a_n, \{a_1, a_n\})}{p(a_1, \{a_1, a_n\})} / \frac{p(a_n, \{a_{i_1}, a_n\})}{p(a_{i_1}, \{a_{i_1}, a_n\})} = u(a_{i_1}, 1)$$

where the second equality is from the definition of u. Therefore, we obtain

$$\frac{u(a_{i_s}, s)}{u(a_{i_1}, 1)} = \frac{p(a_{i_s}, A(s, i_s))}{p(a_1, A(s, i_s))} \cdot [\frac{p(a_{i_2}, \{a_{i_1}, a_{i_2}\})}{p(a_{i_1}, \{a_{i_1}, a_{i_2}\})} \cdot \frac{p(a_1, \{a_1, a_{i_2}\})}{p(a_{i_2}, \{a_1, a_{i_2}\})}]. \tag{E.1}$$

Moreover, since a_{i_1}, a_{i_2} are the top two alternatives in A, by R-IIA,

$$\frac{p(a_{i_2}, \{a_{i_1}, a_{i_2}\})}{p(a_{i_1}, \{a_{i_1}, a_{i_2}\})} = \frac{p(a_{i_2}, A)}{p(a_{i_1}, A)}. \tag{E.2}$$

Combining equations (E.1) and (E.2),

$$\frac{p(a_{i_s}, A)}{p(a_{i_1}, A)} = \frac{u(a_{i_s}, s)}{u(a_{i_1}, 1)} \text{ if and only if } \frac{p(a_{i_s}, A)}{p(a_{i_1}, A)} = \frac{p(a_{i_s}, A(s, i_s))}{p(a_1, A(s, i_s))} \cdot \frac{p(a_{i_2}, A)}{p(a_{i_1}, A)} \cdot \frac{p(a_1, \{a_1, a_{i_2}\})}{p(a_{i_2}, \{a_1, a_{i_2}\})}$$

if and only if

$$\frac{p(a_{i_s}, A)}{p(a_{i_2}, A)} = \frac{p(a_{i_s}, A(s, i_s))}{p(a_1, A(s, i_s))} \cdot \frac{p(a_1, \{a_1, a_{i_2}\})}{p(a_{i_2}, \{a_1, a_{i_2}\})}. \tag{E.3}$$

So it is enough to prove (E.3).

Let $A' \equiv \{a_1, a_{i_2}, \ldots, a_{i_m}\}$. Since $R(a_{i_2}, A) = 2 = R(a_{i_2}, A')$ and $R(a_{i_s}, A) = s = R(a_{i_s}, A')$, by R-IIA,

$$\frac{p(a_{i_s}, A)}{p(a_{i_2}, A)} = \frac{p(a_{i_s}, A')}{p(a_{i_2}, A')}.$$

Moreover, since a_1 and a_{i_2} are the top two alternatives in A', by R-IIA,

$$\frac{p(a_1, A')}{p(a_{i_2}, A')} = \frac{p(a_1, \{a_1, a_{i_2}\})}{p(a_{i_2}, \{a_1, a_{i_2}\})}.$$

Also, since $R(a_1, A') = 1 = R(a_1, A(s, i_s))$ and $R(a_{i_s}, A') = s = R(a_{i_s}, A(s, i_s))$, by R-IIA,

$$\frac{p(a_{i_s}, A')}{p(a_1, A')} = \frac{p(a_{i_s}, A(s, i_s))}{p(a_1, A(s, i_s))}.$$

Combining the last three equalities:

$$\frac{p(a_{i_s}, A)}{p(a_{i_2}, A)} = \frac{p(a_{i_s}, A')}{p(a_{i_2}, A')} = \frac{p(a_1, A')}{p(a_{i_2}, A')} \cdot \frac{p(a_{i_s}, A')}{p(a_1, A')} = \frac{p(a_{i_s}, A(s, i_s))}{p(a_1, A(s, i_s))} \cdot \frac{p(a_1, \{a_1, a_{i_2}\})}{p(a_{i_2}, \{a_1, a_{i_2}\})}.$$

Uniqueness: Now we shall prove the uniqueness of u. Suppose there exist functions u and u' such that $p = p_{(u,R)}$, $p = p_{(u',R)}$, and $u(a_1, 1) = u'(a_1, 1)$.

First, we will prove that for any i and k such that $2 \le i \le n$ and $k \le i$, $u(a_i, k) = u'(a_i, k)$. Let's prove it by induction on i. By assumption, it is true when $i = 1$. Suppose for any $i \le (t-1)$ and $2 \le k \le i$, $u(a_i, k) = u'(a_i, k)$. Now we need to prove that $u(a_t, k) = u'(a_t, k)$ for all $2 \le k \le t$. Let $A = \{a_1, \ldots, a_{k-1}, a_t\}$. Note that

$$p(a_t, A) = \frac{u(a_t, k)}{\sum_{j=1}^{k-1} u(a_j, j) + u(a_t, k)} = \frac{u'(a_t, k)}{\sum_{j=1}^{k-1} u'(a_j, j) + u'(a_t, k)}.$$

Since $u(a_j, j) = u'(a_j, j)$ for all $j \le (k-1) \le (t-1)$, above implies that $u(a_t, k) = u'(a_t, k)$.

Second, we will prove that $u(a_i, 1) = u'(a_i, 1)$ for all $n > i \ge 2$. Note that

$$p(a_i, \{a_i, a_n\}) = \frac{u(a_i, 1)}{u(a_i, 1) + u(a_n, 2)} = \frac{u'(a_i, 1)}{u'(a_i, 1) + u'(a_n, 2)}.$$

Since $u(a_n, 2) = u'(a_n, 2)$ by the previous argument, we obtain $u(a_i, 1) = u'(a_i, 1)$.

E.2 Proof of Proposition 15

Take an ODLM (u, R) and suppose $p = p_{(u,R)}$ satisfies R-Increasing and $\frac{p(a,\{a,b\})}{p(b,\{a,b\})} \cdot \frac{p(b,\{b,c\})}{p(c,\{b,c\})} \ge \frac{p(a,\{a,c\})}{p(c,\{a,c\})}$ for any $a, b, c \in X$ with $aRbRc$. Without loss of generality, let $X = \{a_1, a_2, \ldots, a_n\}$ where $n \ge 3$ and $R(a_i, X) = i$ for each i and $u(a_1, 1) = 1$.

First, let us prove that $u(a_i, k) \le u(a_i, k-1)$ for any $i \ge 3$ and $i \ge k \ge 3$. Let $A = \{a_1, \ldots, a_{k-1}, a_i\}$. Then, by R-Increasing, $a_1 R a_{k-1} R a_i$ implies

$$\frac{p(a_i, A)}{p(a_1, A)} = u(a_i, k) \le \frac{p(a_i, A \setminus a_{k-1})}{p(a_1, A \setminus a_{k-1})} = u(a_i, k-1).$$

Second, we will prove that $u(a_i, 1) \ge u(a_i, 2)$ for all $(n-1) \ge i \ge 2$. Note that

$$u(a_i, 1) = \frac{p(a_i, \{a_i, a_n\})}{p(a_n, \{a_i, a_n\})} \cdot \frac{p(a_n, \{a_1, a_n\})}{p(a_1, \{a_1, a_n\})} \text{ and } u(a_i, 2) = \frac{p(a_i, \{a_1, a_i\})}{p(a_1, \{a_1, a_i\})}.$$

Therefore, $\frac{p(a_1,\{a_1,a_i\})}{p(a_i,\{a_1,a_i\})} \cdot \frac{p(a_i,\{a_i,a_n\})}{p(a_n,\{a_i,a_n\})} \ge \frac{p(a_1,\{a_1,a_n\})}{p(a_n,\{a_1,a_n\})}$ implies $u(a_i, 1) \ge u(a_i, 2)$.

E.3 Proof of Proposition 16

Asymmetry: Take any $a, b \in X$. Without loss of generality, suppose aR^0b; that is, there exists $c \in X \setminus \{a, b\}$ such that $b \in L(a, c)$. Then we cannot have bR^0a because by Asymmetry there is no c' such that $a \in L(b, c')$.

Transitivity: Take any $a, b, c \in X$ such that aR^0b and bR^0c. By Asymmetry, aR^0b implies $a \notin L(b, c)$ and bR^0c implies $b \notin L(a, c)$. By Transitivity, $a \notin L(b, c)$ and $b \notin L(a, c)$ imply $c \in L(a, b)$. Therefore, aR^0c.

Almost Complete: Suppose that there is a pair (a^*, b^*) such that neither $a^*R^0b^*$ nor $b^*R^0a^*$; that is, for any c such that $a^* \notin L(b^*, c)$ and $b^* \notin L(a^*, c)$. By transitivity, we have $c \in L(a^*, b^*)$. Therefore, a^*R^0c and b^*R^0c.

Now we will prove that for any a and b such that $(a, b) \neq (a^*, b^*)$, either aR^0b or bR^0a. We consider three cases.

Case 1: $a \in \{a^*, b^*\}$ and $b \notin \{a^*, b^*\}$. Since $b \in L(a^*, b^*)$, $\{a^*, b^*\} \ni aR^0b$.

Case 2: $a \notin \{a^*, b^*\}$ and $b \in \{a^*, b^*\}$. Since $a \in L(a^*, b^*)$, $\{a^*, b^*\} \ni bR^0a$.

Case 3: $a \notin \{a^*, b^*\}$ and $b \notin \{a^*, b^*\}$. Since $a \in L(a^*, b^*)$, by Asymmetry we obtain $a^* \notin L(a, b)$. Then by Transitivity, $a^* \notin L(a, b)$ implies either $a \in L(a^*, b)$; that is, bR^0a, or $b \in L(a^*, a)$; that is, aR^0b.

E.4 Proof of Proposition 17

Take any ODLM (u, R) such that $u(a, 1) > u(a, 2)$ for any $a \in X$ and $u(b, 2) \neq u(b, 3)$ and $\frac{u(a,1)}{u(b,2)} \neq \frac{u(a,2)}{u(b,3)}$ for all $a, b \in X$ with aRb. Without loss of generality, let $X = \{a_1, a_n, \ldots, a_n\}$ with $a_1Ra_2 \ldots Ra_n$.

First, let us prove that $R^* = R$. Take any i, j with $1 < i < j$. Then $a_1R^0a_j$ and $a_iR^0a_j$ because

$$\frac{p(a_1, \{a_1, a_i, a_j\})}{p(a_i, \{a_1, a_i, a_j\})} = \frac{u(a_1, 1)}{u(a_i, 2)} = \frac{p(a_1, \{a_1, a_i\})}{p(a_i, \{a_1, a_i\})} \Rightarrow a_j \in L(a_1, a_i).$$

Now we shall prove that $a_1R^*a_2$. For any $i > 2$, since $u(a_i, 2) \neq u(a_i, 3)$,

$$\frac{p(a_1, \{a_1, a_2, a_i\})}{p(a_i, \{a_1, a_2, a_i\})} = \frac{u(a_1, 1)}{u(a_i, 3)} \neq \frac{p(a_1, \{a_1, a_i\})}{p(a_i, \{a_1, a_i\})} = \frac{u(a_1, 1)}{u(a_i, 2)} \Rightarrow a_2 \notin L(a_1, a_i).$$

Similarly, for any $i > 2$, since $\frac{u(a_2,2)}{u(a_i,3)} \neq \frac{u(a_2,1)}{u(a_i,2)}$,

$$\frac{p(a_2, \{a_1, a_2, a_i\})}{p(a_i, \{a_1, a_2, a_i\})} = \frac{u(a_2, 2)}{u(a_i, 3)} \neq \frac{p(a_2, \{a_2, a_i\})}{p(a_i, \{a_2, a_i\})} = \frac{u(a_2, 1)}{u(a_i, 2)} \Rightarrow a_1 \notin L(a_2, a_i).$$

Moreover, for any $i > 2$, $\frac{p(a_1,\{a_1,a_2\})\cdot p(a_2,\{a_2,a_i\})\cdot p(a_i,\{a_1,a_i\})}{p(a_2,\{a_1,a_2\})\cdot p(a_i,\{a_2,a_i\})\cdot p(a_1,\{a_1,a_i\})} = \frac{u(a_2,1)}{u(a_2,2)} > 1$. Therefore, $a_1 R^* a_2$.

Now we prove the necessity of Asymmetry and Transitivity.

Asymmetry: Take any $a, b \in X$. Suppose there exists $c \in X \setminus \{a, b\}$ such that $b \in L(a, c)$. We shall prove that $a \notin L(b, c')$ for all $c' \in X \setminus \{a, b\}$. First of all, $b \in L(a, c)$ implies $aR^0 b$. Since $R^* = R$, we have aRb. By way of contradiction, suppose there is $c' \in X \setminus \{a, b\}$ such that $a \in L(b, c')$. Then $bR^0 a$. A contradiction.

Transitivity: Take any $a, b, c \in X$ such that $b \notin U(a, c)$ and $a \notin L(b, c)$. We shall prove that $c \in L(a, b)$. If aRb and cRb, then $b \in L(a, c)$. A contradiction. Similarly, if bRa and cRa, then $a \in L(b, c)$. A contradiction. Therefore, we have aRc and bRc which imply $c \in L(a, b)$.

E.5 Completing revealed order R^0

Proposition 16 shows that R^0 is almost complete. In particular, there are only one pair that is not comparable by R^0. Now we will discuss how to complete R^0 using the following observation.

Observation 13: Take any ODLM (u, R). For any $a, b, c, d \in X$ with $aRbRcRd$,

$$\frac{p(a, \{a, b, d\})}{p(d, \{a, b, d\})} = \frac{u(a, 1)}{u(d, 3)} = \frac{p(a, \{a, c, d\})}{p(d, \{a, c, d\})}$$

and

$$\frac{p(b, \{a, b, d\})}{p(d, \{a, b, d\})} = \frac{u(b, 2)}{u(d, 3)} \neq \frac{u(b, 1)}{u(d, 3)} = \frac{p(b, \{b, c, d\})}{p(d, \{b, c, d\})}.$$

Observation 13 says that since the rankings of the highest ranking alternative a and the lowest ranking alternative d does not change, replacing b with c cannot affect the probability of choosing a relative to that of choosing d. However, replacing a with c affects the probability of choosing b relative to that of choosing d as long as $u(b, 1) \neq u(b, 2)$.

Now we discuss how to complete R^0 using the idea of Observation 3. Suppose there is a pair (a^*, b^*) such that neither $a^* R^0 b^*$ nor $b^* R^0 a^*$. Take any c, d such that $c, d \in X \setminus \{a^*, b^*\}$. By Proposition 16, $a^*, b^* R^0 c, d$ and either $cR^0 d$ or $dR^0 c$. Without loss of generality, let us assume $cR^0 d$.

By Observation 13, if

$$\frac{p(a^*, \{a^*, b^*, d\})}{p(d, \{a^*, b^*, d\})} = \frac{p(a^*, \{a^*, c, d\})}{p(d, \{a^*, c, d\})},$$

then we complete R^0 in the way that a^* has a higher ranking than b^*. However, if

$$\frac{p(a^*, \{a^*, b^*, d\})}{p(d, \{a^*, b^*, d\})} \neq \frac{p(a^*, \{a^*, c, d\})}{p(d, \{a^*, c, d\})},$$

then we complete R^0 in the way that b^* has a higher ranking than a^*. As long as $u(x, 1) \neq u(x, 2)$ for any $x \in X$, we can uniquely identify an unknown R.

BIBLIOGRAPHY

ACKERT, L. F., N. CHARUPAT, B. K. CHURCH, AND R. DEAVES (2006): "An experimental examination of the house money effect in a multi-period setting," *Experimental Economics*, 9, 5–16.

ALLAIS, M. (1953): "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine," *Econometrica: Journal of the Econometric Society*, 503–546.

APESTEGUIA, J. AND M. A. BALLESTER (2009): "A theory of reference-dependent behavior," *Economic Theory*, 40, 427–455.

——— (2013): "Choice by sequential procedures," *Games and Economic Behavior*, 77, 90–99.

ARIELI, A., Y. BEN-AMI, AND A. RUBINSTEIN (2011): "Tracking Decision Makers under Uncertainty," *American Economic Journal: Microeconomics*, 3, 68–76.

ARIELY, D. AND T. S. WALLSTEN (1995): "Seeking subjective dominance in multidimensional space: An explanation of the asymmetric dominance effect," *Organizational Behavior and Human Decision Processes*, 63, 223–232.

ARROW, K. J. (1959): "Rational choice functions and orderings," *Economica*, 26, 121–127.

ATTANASIO, O. P. (1999): "Consumption," *Handbook of macroeconomics*, 1, 741–812.

ATTANASIO, O. P. AND G. WEBER (2010): "Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy," *Journal of Economic Literature*, 48, 693–751.

BARBERÁ, S. AND P. K. PATTANAIK (1986): "Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings," *Econometrica*, 54, pp. 707–715.

BARBERIS, N., M. HUANG, AND T. SANTOS (2001): "Prospect Theory and Asset Prices," *The Quarterly Journal of Economics*, 116, 1–53.

BARBERIS, N. C. (2013): "Thirty years of prospect theory in economics: A review and assessment," *Journal of Economic Perspectives*, 27, 173–96.

BENARTZI, S. AND R. H. THALER (1995): "Myopic loss aversion and the equity premium puzzle," *The quarterly journal of Economics*, 110, 73–92.

BLACKORBY, C., D. PRIMONT, AND R. ROBERT RUSSELL (2008): "Separability," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and L. E. Blume, Basingstoke: Palgrave Macmillan.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2012): "Salience theory of choice under risk," *The Quarterly Journal of Economics*, 127, 1243–1285.

——— (2013): "Salience and Consumer choice," *Journal of Political Economy*, 121, 803–843.

BOUYSSOU, D. AND M. PIRLOT (2002): "Nontransitive decomposable conjoint measurement," *Journal of Mathematical Psychology*, 46, 677–703.

BRUNI, L. AND R. SUGDEN (2007): "The road not taken: how psychology was removed from economics, and how it might be brought back*," *The Economic Journal*, 117, 146–173.

BUDESCU, D. AND W. WEISS (1987): "Reflection of transitive and intransitive preferences: A test of prospect theory," *Organizational Behavior and Human Decision Processes*, 39, 184–202.

CAMERER, C. F. AND T.-H. HO (1994): "Violations of the betweenness axiom and nonlinearity in probability," *Journal of risk and uncertainty*, 8, 167–196.

CAMPBELL, J. Y. (2003): "Consumption-based asset pricing," *Handbook of the Economics of Finance*, 1, 803–887.

CAMPBELL, J. Y. AND J. H. COCHRANE (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *The Journal of Political Economy*, 107, 205–251.

CAMPBELL, J. Y. AND N. G. MANKIW (1991): "The response of consumption to income: a cross-country investigation," *European economic review*, 35, 723–756.

Cherepanov, V., T. Feddersen, and A. Sandroni (2013): "Rationalization," *Theoretical Economics*, 8, 775–800.

Chernev, A. (2003): "When more is less and less is more: The role of ideal point availability and assortment in consumer choice," *Journal of consumer Research*, 30, 170–183.

——— (2004): "Extremeness Aversion and Attribute-Balance Effects in Choice," *Journal of Consumer Research*, 31, 249–263.

Cochrane, J. H. (2011): "Presidential address: Discount rates," *The Journal of Finance*, 66, 1047–1108.

Constantinides, G. M. (1990): "Habit formation: A resolution of the equity premium puzzle," *Journal of political Economy*, 519–543.

Day, B. and G. Loomes (2010): "Conflicting violations of transitivity and where they may lead us," *Theory and decision*, 68, 233–242.

De Clippel, G. and K. Eliaz (2012): "Reason-based choice: A bargaining rationale for the attraction and compromise effects," *Theoretical Economics*, 7, 125–162.

Deaton, A. (1991): "Saving and liquidity constraints," *The Quarterly Journal of Economics*, 59, 1221–48.

Debreu, G. (1960a): "Review of RD Luce, Individual choice behavior: A theoretical analysis," *American Economic Review*, 50, 186–188.

——— (1960b): "Topological methods in cardinal utility theory," *Mathematical methods in the social sciences*, 1959, 16–26.

Dhar, R. and I. Simonson (2003): "The effect of forced choice on choice," *Journal of Marketing Research*, 146–160.

Dillenberger, D. and K. Rozen (2015): "History-dependent risk attitude," *Journal of Economic Theory*, 157, 445–477.

Doyle, J. R., D. J. O'Connor, G. M. Reynolds, and P. A. Bottomley (1999): "The robustness of the asymmetrically dominated effect: Buying frames, phantom alternatives, and in-store purchases," *Psychology & Marketing*, 16, 225–243.

Dreze, X., S. J. Hoch, and M. E. Purk (1995): "Shelf management and space elasticity," *Journal of Retailing*, 70, 301–326.

Echenique, F., K. Saito, and G. Tserenjigmid (2013): “The Perception-Adjusted Luce Model,” *Caltech SS Working Paper 1379*.

Ehlers, L. and Y. Sprumont (2008): “Weakened WARP and top-cycle choice rules,” *Journal of Mathematical Economics*, 44, 87–94.

Eliaz, K., M. Richter, and A. Rubinstein (2011): “Choosing the two finalists,” *Economic Theory*, 46, 211–219.

Epstein, L. G. and S. E. Zin (1989): “Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework,” *Econometrica: Journal of the Econometric Society*, 937–969.

——— (2001): “The independence axiom and asset returns,” *Journal of Empirical Finance*, 8, 537–572.

Falmagne, J. (1978): “A representation theorem for finite random scale systems,” *Journal of Mathematical Psychology*, 18, 52–72.

Fishburn, P. C. and A. Rubinstein (1982): “Time preference,” *International Economic Review*, 23, 677–694.

Fudenberg, D., R. Iijima, and T. Strzalecki (2013a): “Stochastic Choice and Revealed Perturbed Utility,” Tech. rep., Mimeo Harvard University.

——— (2013b): “Stochastic Choice and Revealed Perturbed Utility,” Mimeo, Harvard University.

——— (2015): “Stochastic choice and revealed perturbed utility,” *Econometrica*, 83, 2371–2409.

Guiso, L., P. Sapienza, and L. Zingales (2013): “Time varying risk aversion,” Tech. rep., National Bureau of Economic Research.

Gul, F. (1991): “A theory of disappointment aversion,” *Econometrica: Journal of the Econometric Society*, 667–686.

Gul, F., P. Natenzon, and W. Pesendorfer (2010): “Random Choice as Behavioral Optimization,” *Unpublished manuscript*.

——— (2014): “Random Choice as Behavioral Optimization,” Forthcoming, Econometrica.

Harrison, G. W. (2007): "House money effects in public good experiments: Comment," *Experimental Economics*, 10, 429–437.

Hayashi, T. (2008): "Regret aversion and opportunity dependence," *Journal of Economic Theory*, 139, 242–268.

Heath, T. B. and S. Chatterjee (1995): "Asymmetric decoy effects on lower-quality versus higher-quality brands: Meta-analytic and experimental evidence," *Journal of Consumer Research*, 268–284.

Herne, K. (1997): "Decoy alternatives in policy choices: Asymmetric domination and compromise effects," *European Journal of Political Economy*, 13, 575–589.

——— (1998): "Testing the reference-dependent model: An experiment on asymmetrically dominated reference points," *Journal of Behavioral Decision Making*, 11, 181–192.

——— (1999): "The effects of decoy gambles on individual choice," *Experimental Economics*, 2, 31–40.

Hey, J. D. (2001): "Does repetition improve consistency?" *Experimental economics*, 4, 5–54.

Ho, D. E. and K. Imai (2008): "Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment The California Alphabet Lottery, 1978–2002," *Public Opinion Quarterly*, 72, 216–240.

Huber, J., J. W. Payne, and C. Puto (1982): "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of consumer research*, 90–98.

Iyengar, S. S. and M. R. Lepper (2000): "When choice is demotivating: Can one desire too much of a good thing?" *Journal of personality and social psychology*, 79, 995.

Jappelli, T. (1990): "Who is credit constrained in the US economy?" *The Quarterly Journal of Economics*, 105, 219–234.

Kahneman, D. and A. Tversky (1979): "Prospect theory: An analysis of decision under risk," *Econometrica: Journal of the Econometric Society*, 263–291.

Kamenica, E. (2008): "Contextual inference in markets: On the informational content of product lines," *The American Economic Review*, 98, 2127–2149.

Kaustia, M. and S. Knüpfer (2008): "Do investors overweight personal experience? Evidence from IPO subscriptions," *The Journal of Finance*, 63, 2679–2702.

Keeney, R. L. and H. Raiffa (1976): "Decision analysis with multiple conflicting objectives," *Wiley& Sons, New York*.

Kim, K. H. (2014): "Counter-cyclical risk aversion," *Journal of Empirical Finance*, 29, 384–401.

Knüpfer, S., E. H. Rantapuska, and M. Sarvimäki (2014): "Labor market experiences and portfolio choice: Evidence from the Finnish Great Depression," *Available at SSRN 2275930*.

Koopmans, T. C. (1960): "Stationary ordinal utility and impatience," *Econometrica: Journal of the Econometric Society*, 287–309.

Koszegi, B. and M. Rabin (2006): "A MODEL OF REFERENCE-DEPENDENT PREFERENCES." *Quarterly journal of economics*, 121.

——— (2007): "Reference-dependent risk attitudes," *The American Economic Review*, 97, 1047–1073.

——— (2009): "Reference-dependent consumption plans," *The American Economic Review*, 99, 909–936.

Krajbich, I. and A. Rangel (2011): "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions," *Proceedings of the National Academy of Sciences*, 108, 13852–13857.

Krantz, D., D. Luce, P. Suppes, and A. Tversky (1971): "Foundations of measurement, Vol. I: Additive and polynomial representations," .

Kreps, D. M. and E. L. Porteus (1978): "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica: journal of the Econometric Society*, 185–200.

Kuczma, M. (2008): *An introduction to the theory of functional equations and inequalities: Cauchy's equation and Jensen's inequality*, Springer.

Lindman, H. and J. Lyons (1978): “Stimulus complexity and choice inconsistency among gambles,” *Organizational Behavior and Human Performance*, 21, 146–159.

Liu, Y.-J., C.-L. Tsai, M.-C. Wang, and N. Zhu (2010): “Prior consequences and subsequent risk taking: New field evidence from the Taiwan Futures Exchange,” *Management Science*, 56, 606–620.

Lleras, J. S., Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay (2010): “When More is Less: Choice with Limited Consideration,” Mimeo, U. of Michigan.

Loomes, G., C. Starmer, and R. Sugden (1991): “Observing violations of transitivity by experimental methods,” *Econometrica: Journal of the Econometric Society*, 425–439.

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

Lucas, J. R. E. (1978): “Asset prices in an exchange economy,” *Econometrica: Journal of the Econometric Society*, 1429–1445.

Luce, R. D. (1959a): *Individual Choice Behavior a Theoretical Analysis*, John Wiley and sons.

——— (1959b): *Individual Choice Behavior a Theoretical Analysis*, John Wiley and sons.

——— (1990): “Rational versus plausible accounting equivalences in preference judgments,” *Psychological Science*, 1, 225–234.

Machina, M. J. (1982): “" Expected Utility" Analysis without the Independence Axiom,” *Econometrica: Journal of the Econometric Society*, 277–323.

——— (1985): “Stochastic choice functions generated from deterministic preferences over lotteries,” *The economic journal*, 575–594.

——— (1989): “Dynamic consistency and non-expected utility models of choice under uncertainty,” *Journal of Economic Literature*, 1622–1668.

Malmendier, U. and S. Nagel (2011): “Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?*.” *Quarterly Journal of Economics*, 126.

Manzini, P. and M. Mariotti (2007): "Sequentially rationalizable choice," *The American Economic Review*, 1824–1839.

——— (2012a): "Categorize then choose: Boundedly rational choice and welfare," *Journal of the European Economic Association*, 10, 1141–1165.

——— (2012b): "Choice by lexicographic semiorders," *Theoretical Economics*, 7, 1–23.

——— (2014): "Stochastic Choice and Consideration Sets," *Econometrica*, 82, 1153–1176.

Manzini, P., M. Mariotti, and C. J. Tyson (2013): "Two-stage threshold representations," *Theoretical Economics*, 8, 875–882.

Markowitz, H. (1952): "The utility of wealth," *The Journal of Political Economy*, 151–158.

Marley, A. (1991): "Context dependent probabilistic choice models based on measures of binary advantage," *Mathematical Social Sciences*, 21, 201–231.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al. (1995): *Microeconomic theory*, vol. 1, Oxford university press New York.

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012): "Revealed attention," *The American Economic Review*, 102, 2183–2205.

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

——— (2013): "A canonical model of choice with initial endowments," *The Review of Economic Studies*.

Masatlioglu, Y. and N. Uler (2013): "Understanding the reference effect," *Games and Economic Behavior*, 82, 403–423.

Massa, M. and A. Simonov (2005): "Behavioral biases and investment," *Review of Finance*, 9, 483–507.

McFadden, D. (2001): "Economic choices," *The American Economic Review*, 91, 351–378.

McFadden, D. and M. Richter (1990): "Stochastic rationality and revealed stochastic preference," in *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz, Westview Press: Boulder, CO*, 161–186.

Mehra, R. and E. C. Prescott (1985): "The equity premium: A puzzle," *Journal of monetary Economics*, 15, 145–161.

Meredith, M. and Y. Salant (2013): "On the causes and consequences of ballot order effects," *Political Behavior*, 35, 175–197.

Munro, A. and R. Sugden (2003): "On the theory of reference-dependent preferences," *Journal of Economic Behavior & Organization*, 50, 407–428.

Natenzon, P. (2010): "Random choice and learning," Tech. rep., Working paper, Princeton University.

Ok, E. and Y. Masatlioglu (2007): "A theory of (relative) discounting," *Journal of Economic Theory*, 137, 214–245.

Ok, E. A., P. Ortoleva, and G. Riella (2014): "Revealed (p) reference theory," *Forthcoming, The American Economic Review.*

Pagel, M. (2013): "Expectations-Based Reference-Dependent Life-Cycle Consumption," .

Peng, J., D. Miao, and W. Xiao (2013): "Why are gainers more risk seeking," *Judgment and Decision Making*, 8, 150–160.

Pollak, R. A. (1970): "Habit formation and dynamic demand functions," *The Journal of Political Economy*, 745–763.

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

Ratneshwar, S., A. D. Shocker, and D. W. Stewart (1987): "Toward understanding the attraction effect: The implications of product stimulus meaningfulness and familiarity," *Journal of Consumer Research*, 520–533.

Ravid, D. (2015): "Focus, Then Compare," Working Paper, Princeton University.

Reutskaja, E., R. Nagel, C. F. Camerer, and A. Rangel (2011): "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study," *The American Economic Review*, 101, 900–926.

Rieskamp, J., J. R. Busemeyer, and B. A. Mellers (2006): "Extending the bounds of rationality: evidence and theories of preferential choice," *Journal of Economic Literature*, 631–661.

Roe, R. M., J. R. Busemeyer, J. T. Townsend, et al. (2001): "Multialternative decision field theory: A dynamic connectionist model of decision making," *PSYCHOLOGICAL REVIEW-NEW YORK-*, 108, 370–392.

Routledge, B. R. and S. E. Zin (2010): "Generalized disappointment aversion and asset prices," *The Journal of Finance*, 65, 1303–1332.

Rubinstein, A. (1988): "Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?)," *Journal of Economic Theory*, 46, 145–153.

Rubinstein, A., Y. Salant, et al. (2006): "A model of choice from lists," .

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

Saito, K. (2015): "Preferences for flexibility and randomization under uncertainty," *American Economic Review*, 105, 1246–1271.

Salant, Y. and A. Rubinstein (2008): "(A, f): Choice with Frames," *The Review of Economic Studies*, 75, 1287–1296.

Segal, U. (1990): "Two-stage lotteries without the reduction axiom," *Econometrica: Journal of the Econometric Society*, 349–377.

Selden, L. (1978): "A new representation of preferences over" certain x uncertain" consumption pairs: The" ordinal certainty equivalent" hypothesis," *Econometrica: Journal of the Econometric Society*, 1045–1060.

Sharpe, K. M., R. Staelin, and J. Huber (2008): "Using extremeness aversion to fight obesity: policy implications of context dependent demand," *Journal of Consumer Research*, 35, 406–422.

Sheng, S., A. M. Parker, and K. Nakamoto (2005): "Understanding the mechanism and determinants of compromise effects," *Psychology & Marketing*, 22, 591–609.

Simonson, I. (1989): "Choice based on reasons: The case of attraction and compromise effects," *Journal of consumer research*, 158–174.

Simonson, I. and A. Tversky (1992): "Choice in context: tradeoff contrast and extremeness aversion." *Journal of marketing research*.

Sivakumar, K. and J. Cherian (1995): "Role of product entry and exit on the attraction effect," *Marketing Letters*, 6, 45–51.

Sue O'Curry, Y. P. and R. Pitts (1995): "The attraction effect and political choice in two elections," *Journal of Consumer Psychology*, 4, 85–101.

Sugden, R. (2003): "Reference-dependent subjective expected utility," *Journal of economic theory*, 111, 172–191.

Suppes, P. and M. Winet (1955): "An axiomatization of utility based on the notion of utility differences," *Management Science*, 1, 259–270.

Tang, Y. and R. F. Whitelaw (2011): "Time-varying Sharpe ratios and market timing," *The Quarterly Journal of Finance*, 1, 465–493.

Teppan, E. C. and A. Felfernig (2009): "Minimization of product utility estimation errors in recommender result set evaluations," in *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 01*, IEEE Computer Society, 20–27.

Thaler, R. H. and E. J. Johnson (1990): "Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice," *Management science*, 36, 643–660.

Tserenjigmid, G. (2013): "The Order Dependent Luce Model," Mimeo, Caltech.

Tversky, A. (1969): "Intransitivity of preferences." *Psychological Review; Psychological Review*, 76, 31.

——— (1972a): "Choice by elimination," *Journal of mathematical psychology*, 9, 341–367.

——— (1972b): "Elimination by aspects: A theory of choice." *Psychological review*, 79, 281.

TVERSKY, A. AND D. KAHNEMAN (1991): "Loss aversion in riskless choice: A reference-dependent model," *The Quarterly Journal of Economics*, 106, 1039–1061.

——— (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 5, 297–323.

TVERSKY, A. AND I. SIMONSON (1993): "Context-dependent preferences," *Management science*, 39, 1179–1189.

USHER, M. AND J. L. MCCLELLAND (2004): "Loss aversion and inhibition in dynamical models of multialternative choice." *Psychological review*, 111, 757.

WAKKER, P. (1988): "The algebraic versus the topological approach to additive representations," *Journal of Mathematical Psychology*, 32, 421–435.

WEIL, P. (1989): "The equity premium puzzle and the risk-free rate puzzle," *Journal of Monetary Economics*, 24, 401–421.

WERNERFELT, B. (1995): "A rational reconstruction of the compromise effect: Using market data to infer utilities," *Journal of Consumer Research*, 627–633.

WIEBACH, N. AND L. HILDEBRANDT (2012): "Explaining customers' switching patterns to brand delisting," *Journal of Retailing and Consumer Services*, 19, 1–10.

YILDIZ, K. (2012): "List rationalizable choice," *Available at SSRN 2147231*.

ZELDES, S. P. (1989): "Consumption and liquidity constraints: an empirical investigation," *The Journal of Political Economy*, 97, 305.