ESSAYS IN BEHAVIORAL DECISION THEORY

Thesis by

Matthew L. Kovach

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2015

(Defended May 15, 2015)

I dedicate this project to my wife, Alexandra Fritz. Without her love and support this would not be possible.

Acknowledgements

I am especially grateful for the guidance and encouragement of Federico Echenique and Pietro Ortoleva. I also benefited greatly from discussions with Kota Saito, Jakša Cvitani□, Leeat Yariv, Colin Camerer, John Ledyard, and Marina Agranov.

I would also like to thank my fellow classmates, Jonathan Chapman, Matthew Chao, Khai Chiong, Euncheol Shin, Allyson Pellissier, Samantha Myers, Kirill Pogorelskiy, Gerelt Tserenjigmid, and Liam Clegg, for their support and friendship.

Abstract

This thesis studies decision making under uncertainty and how economic agents respond to information. The classic model of subjective expected utility and Bayesian updating is often at odds with empirical and experimental results; people exhibit systematic biases in information processing and often exhibit aversion to ambiguity. The aim of this work is to develop simple models that capture observed biases and study their economic implications.

In the first chapter I present an axiomatic model of cognitive dissonance, in which an agent's response to information explicitly depends upon past actions. I introduce novel behavioral axioms and derive a representation in which beliefs are *directionally* updated. The agent twists the information and overweights states in which his past actions provide a higher payoff. I then characterize two special cases of the representation. In the first case, the agent distorts the likelihood ratio of two states by a function of the utility values of the previous action in those states. In the second case, the agent's posterior beliefs are a convex combination of the Bayesian belief and the one which maximizes the conditional value of the previous action. Within the second case a unique parameter captures the agent's sensitivity to dissonance, and I characterize a way to compare sensitivity to dissonance between individuals. Lastly, I develop several simple applications and show that cognitive dissonance contributes to the equity premium and price volatility, asymmetric reaction to news, and belief polarization.

The second chapter characterizes a decision maker with *sticky* beliefs. That is, a decision maker who does not update enough in response to information, where enough means as a Bayesian decision maker would. This chapter provides axiomatic foundations for sticky beliefs by weakening the standard axioms of dynamic consistency and consequentialism. I derive a representation in which updated beliefs are a convex combination of the prior and the Bayesian posterior. A unique parameter captures the weight on the prior and is interpreted as the agent's measure of belief stickiness or conservatism bias. This parameter is endogenously identified from preferences and is easily elicited from experimental data.

The third chapter deals with updating in the face of ambiguity, using the framework of Gilboa and Schmeidler. There is no consensus on the correct way way to update a set of priors. Current methods either do not allow a decision maker to make an inference about her priors or require an extreme level of inference. In this chapter I propose and axiomatize a general model of updating a set of priors. A decision maker who updates her beliefs in accordance with the model can be thought of as one that chooses a threshold that is used to determine whether a prior is *plausible*, given some observation. She retains the plausible priors and applies Bayes' rule. This model includes generalized Bayesian updating and maximum likelihood updating as special cases.

Contents

List of Figures

Chapter 1

Twisting the Truth: A Model of Cognitive Dissonance and Information

1.1 Introduction

This paper provides a theory of how an agent sensitive to cognitive dissonance incorporates new information into his beliefs. Cognitive dissonance refers to the psychological discomfort that arises when two *cognitions* are in conflict. In this paper, cognitive dissonance arises when information and a previous action are in conflict. That is, if the agent updated his beliefs according to Bayes' rule, his beliefs and past actions would create dissonance. The agent assuages this cognitive dissonance by incorporating information in a non-Bayesian manner and distorting his beliefs to support his previous actions.

Consider an agent who chooses actions at two points in time, where these actions may have uncertain payoffs depending on some unknown or unrealized state of the world. Between times 1 and 2 the agent observes information about the state of the world. The standard rational model assumes that an agent uses Bayes' rule to update his beliefs, but if the agent is sensitive to cognitive dissonance he has an inherent desire to justify past actions. This paper investigates how the agent resolves the tension between Bayesian learning and the desire to view past actions as optimal.

As a simple example, consider an investor deciding which company's stock to purchase. After the investor makes his initial purchase, he receives some information that is relevant to the company's valuation. The investor now must take the new information into account and decide again which stocks to purchase (or sell). However, our investor is sensitive to cognitive dissonance and thus experiences psychic distress if the new information, combined with his original beliefs, suggests that he originally made a poor investment decision. Hence the investor's original decision and new information jointly determine his updated beliefs, and consequently his new decision.

While the concept of cognitive dissonance is well known, the psychology literature has not provided a precise notion of how it affects an agent's future decision making, which is necessary to apply the model to economic problems. The main contribution of this paper is to answer these questions in a way suitable for an economist. I utilize a standard economic setup to study the effect of information on preferences. Within this framework I introduce behavioral conditions, in the form of axioms on preferences, that capture cognitive dissonance and then derive a utility representation. Thus this paper answers the question of how cognitive dissonance affects an agent's response to information.

More formally, I assume the agent's preferences over actions as a primitive, where actions have consequences that depend on some state of the world. Time 1 preferences are both before information and before an action choice, whereas time 2 preferences are conditional on both the information received and some chosen action, jointly referred to as a *scenario*. The main behavioral condition in the paper is the axiom of *Dynamic Cognitive Dissonance*. For a simple intuition behind this axiom, consider an investor that invests in company X at time 1. In particular, suppose company X may either yield a high return or a low return. Say H and L are the events in which X provides a high or low return, respectively. Then after any observation, at time 2 the agent prefers investments that also have high payoffs in H. The resulting utility representation is one in which the agent's time 2 beliefs shift probability from states in which the time 1 action, denoted by f, is relatively poor to states in which it is relatively good.

In addition to the general model I characterize two special cases, each of which is derived by imposing one additional axiom. The first case I characterize is the *proportional* distortion. Under this representation the agent distorts the relative likelihood between states by the payoff of f in those states. Thus it is as if the agent views his original action f as being informative about the relative likelihood of states of the world. The proportional distortion is characterized by the addition of a *scenario independence* property, which states that whenever two scenarios share a common event in which each action provides the same state-wise payoff on the common event, then the agent's ranking of acts that vary only on the common event are the same in each scenario. That is, the relative distortion between any two states is independent of the payoffs in any other states.

The second case I characterize, referred to as the *best-case binary* distortion, is one in which time 2 beliefs are a convex combination of the Bayesian posterior and the posterior that maximizes the value of the time 1 action. In contrast to the proportional distortion, the best-case binary model generally violates the scenario independence condition. Specifically, this is because the posterior that maximizes the value of the time 1 action is conditional on the realized event, hence the justifying posterior varies across scenarios. Thus one can think of the agent as compromising between two selves—a rational Bayesian self and a self that only considers the best possible states of the world, where best is defined relative to the original action. The weight that each self receives in the representation is endogenously captured by a unique parameter,δ, which I interpret as the agent's sensitivity to cognitive dissonance. An agent that is very sensitive to dissonance, or has a large δ, puts greater weight on the justifying states.

The model produces some interesting, testable implications. First, risk free actions induce no belief distortion and thus the agent appears Bayesian in some situations. This is intuitive since there is no payoff variation in a risk free action and hence there is no possible revision of beliefs that could make the action appear better. Second, the agent will exhibit an asymmetric reaction to good and bad news, which is consistent with empirical evidence on financial analysts' forecasts (see Easterwood and Nutt, 1999). That is, if we consider an agent's monetary valuation of some action, the agent always overvalues the time 1 action compared to a Bayesian agent. Thus the agent *over-reacts* to the good news and *under-reacts* to the bad news. Neither of these implications can result from models of non-Bayesian updating that do not also condition on an agent's past action.

1.1.1 The Psychology of Cognitive Dissonance

The theory of cognitive dissonance, developed by Leon Festinger [22], states that people tend to adjust beliefs to enhance the attractiveness of their past actions. In particular, Festinger proposed that conflict or tension between beliefs and actions creates psychological discomfort. He termed this resulting discomfort *dissonance* and states that the only way to eliminate this discomfort is to eliminate the conflict and achieve *consonance*. Thus after taking some action people are motivated to change their beliefs about the desirability of that action.

For example, suppose an agent invests in company X. The action of investing reveals beliefs that company X will provide high returns in the future. If negative information is released about the company, then the agent now experiences the following conflicting thoughts, *I bought X expecting high returns* and *this information suggests X was a bad investment*, and hence suffers the discomfort caused by cognitive dissonance. In order to achieve consonance the agent incorporates the new information into his original beliefs in a biased manner. This bias causes the agent to increase the conditional likelihood of high returns and hence to view investment in X more favorably than an outside Bayesian would.

1.1.1.1 Experimental Evidence

In an early and influential laboratory experiment, Festinger and Carlsmith [23] asked students to perform a long and boring task and then to recruit more participants. Some students were paid a substantial amount while others were paid very little. Those who were paid very little reported the task as more interesting than students who were paid a more substantial amount. This suggests that those who were paid little manipulate their beliefs in order to justify performing the task for very little pay. Similarly, students who gave speeches advocating an ideological position were more likely to align their beliefs with their speech the lower their pay (see Aronson [5] for an overview of this and other experiments).

In an investment experiment by Kuhnen and Knutson [43], subjects could either purchase a bond or a stock. Stocks could be good or bad and could pay a high or low dividend, with good stocks more likely to pay high dividends. Subjects were told that good and bad stocks are equally likely and were also given objective information about the likelihood of a good (bad) stock paying a high or low dividend. After deciding whether to purchase a stock or bond, subjects observed the dividend payments and were then asked to provide their beliefs about the probability that the stock was good. They found that, relative to subjects that purchased the bond, those that purchased the stock over-estimated the probability that the stock was good by 10%. This effect was present even after controlling for the objective probability and for the Bayesian posterior of the subject's reported beliefs from previous periods. This suggests that subjects update their beliefs in a non-Bayesian manner dependent on their actions.

1.1.1.2 Empirical Evidence

The specific role of cognitive dissonance in voter preferences was studied by Mullainathan and Washington [49]. They measured the effect of voting for a candidate on a voter's future opinion of that candidate. To control for the selection problem, the authors compared the opinion ratings of voting age eligible and ineligible voters two years after the 1996 presidential election. They found that eligible voters showed 2-3 times greater polarization than ineligible voters, supporting the relevance of cognitive dissonance in shaping political attitudes.

A more recent paper by Kaplan and Mukand [37] shows that political party registration seems to be excessively persistent. They also utilize a discontinuity design based on voting age while also utilizing the 9/11/01 terrorist attacks as an exogenous shock to party registration. Party registration is persistent even for those registered near universities, suggesting that this persistence is not easily explained by lack of access to information.

1.1.2 Relation to the Literature

Akerlof and Dickens [2] developed perhaps the earliest model of cognitive dissonance in economics. They allow for the agent to choose his beliefs while considering both the cost of making the wrong decision and a psychological cost of believing that his past choice was suboptimal.[1] Their main result shows that cognitive

[1] That is, they implicitly assume that the agent holds the belief: "I make good decisions and would not choose a risky career". Because of this, any level of risk is suboptimal and the associated cost is increasing in the perceived risk level.

dissonance may cause workers to forgo efficient safety equipment. An application demonstrating that this result generally holds is developed later in the paper.

Perhaps the most closely related papers are Yariv [63] and Epstein and Kopylov [18]. Yariv considers as a primitive preferences over pairs of actions and beliefs. She then provides axiomatic foundations for a linear representation over infinite streams of action and beliefs pairs. However, her setup and axioms do not deliver any specific way in which actions and beliefs interact. Additionally, it may be that the action has no impact on the next periods belief. Epstein and Kopylov consider a model with preferences over menus of acts and impose modifications of the Gul and Pesendorfer [32] axioms. They derive a representation where the agent chooses a menu *ex-ante* to balance his preferences under commitment—expected utility with respect to true beliefs—with a temptation utility—where the agent minimizes (or maximizes) the value of an action over a set of priors. While conceptually similar, [18] is distinct in two important ways. First, the representation suggests *ex-post* choice from the menu is of a max-min form while this paper imposes expected utility. Second, the set of priors Q is independent of any choices the agent makes and hence *ex-post* beliefs are independent of the agent's actions, while in this paper posterior beliefs explicitly depended upon the agent's action.

Mayraz [46] studies a model of payoff-dependent beliefs. His paper assumes preferences conditional on some real-valued payoff function, but does not have a notion of *ex-ante* vs *ex-post* preferences or information and thus does make the connection to non-Bayesian updating or weakening dynamic consistency. Additionally, he requires distortions to take a specific functional form, whereas this paper studies a general model that allows for a variety of distortions.

Other axiomatic models of non-Bayesian beliefs include Epstein [16] and Epstein et al. [21]. Epstein (2006) also utilizes a modification of the temptation model of Gul and Pesendorfer [32]. In this paper Epstein focuses on modeling general non-Bayesian behavior (consequences of which are explored in [21]) but does not allow for updating to depend on anything other than the information received. In contrast I propose a single behavioral axiom, along with some regularity conditions, and generate a representation such that updated beliefs depend explicitly upon past choice.

Ortoleva [52] develops an axiomatic model of updating that is at the intersection of Bayesian and non-Bayesian models. That is, he introduces the *hypothesis testing representation* which holds whether or not the agent utilizes Bayes' rule to update beliefs. However, the Bayesian model is embedded as a special case. An agent in his model chooses a new prior only in the case of an unexpected (low probability) event, and may or may not use Bayes' rule otherwise. The model presented in this paper is distinct since the deviation from Bayes' rule depends on the interaction of a past choice and information, not purely on information. Thus an agent in this model can behave as a Bayesian when responding to an unexpected event and violate Bayes' rule for an expected event.

A closely related, non-axiomatic paper is Yariv [64]. She considers an agent represented by an instrumental utility (a classical utility over consequences) and a belief utility, where the belief utility captures the agent's innate preference for belief consistency. Her agent is forward looking, though the agent may incorrectly forecast the weight placed on the belief utility. However, in each period the agent's choice is over beliefs, subject to the constraint that the agent will take an action consistent with his beliefs and suffers a cost of changing his beliefs. This is contrasted with my model in which the belief change is not necessarily a conscious procedure and manipulations force beliefs to be *more* consistent with past actions, rather than past beliefs. (see also Bénabou and Tirole [9], Bénabou [8]).

Other papers of relevance include Caplin and Leahy [12] and Brunnermeier and Parker [10]. Caplin and Leahy consider a two period model in which the agent's utility is defined over both prizes and psychological states. However, the model designed to study the role of *anticipatory* feelings. Such an agent chooses his beliefs *ex-ante* to balance his *instrumental* (prize) utility and his utility from anticipation (anxiety). Thus their model is not well suited to study cognitive dissonance, since cognitive dissonance is not a forward looking emotion but a retrospective one. Brunnermeier and Parker consider a dynamic model in which the agent balances how much he distorts his beliefs from the truth with his taste for optimistic expectations about future utility. Once the initial belief is chosen however, the agent acts as a Bayesian in all future periods. [31, see also]

A psychological concept closely related to cognitive dissonance is *motivated reasoning*. An agent engaging in this behavior reasons so that he may support his favored ideas or actions, perhaps by only acknowledging some information (see Kruglanski [42] Kunda [44]). In a sense, motivated reasoning can be seen as a mechanism by which cognitive dissonance is reduced. With this view, the model of cognitive dissonance in this paper is also a model of motivated reasoning. Motivated reasoning in political science has been studied by Redlawsk [57] and by Taber and Lodge [62].

The literature on Bayesian updating is large and I do not to attempt provide a complete survey. The Bayesian model (subjective expected utility and Bayesian updating) is known to be equivalent to preferences satisfying the traditional savage axioms plus two axioms: consequentialism and dynamic consistency (see Ghirardato [25] for a review). Consequentialism is a weak rationality condition which states that for any event, preferences conditional on that event only depend on how acts perform within that event. That is, the agent is not concerned with impossibilities. Dynamic consistency requires that preferences conditional on an event respect unconditional rankings of acts that differ only within that event. Specifically, this rules out changing one's mind about the relative likelihoods of states within an event. The literature on alternatives to the Bayesian model is perhaps larger; for a sample, see Camerer [11], Kahneman and Tversky [36], Mullainathan et al. [50], Rabin and Schrag [55], Rabin and Vayanos [56].

1.2 Setup and Foundations

1.2.1 Formal Setup

I adopt a standard setup for studying the effect of information on preferences. There is a finite set S of states of the world, with $|S| \geq 3$.[2] Events are denoted $A, B, C \in \Sigma = 2^S \backslash \{S, \varnothing\}$[3] and X denotes the set of consequences, assumed to be a convex subset of a vector space. For example, X could be the set of monetary prizes (e.g. $X = \mathbb{R}_+$) or X could be the set of lotteries over some set Y (which corresponds to the classic Anscombe–Aumann setup [4]). Let $\mathscr{F}$ denote the set of all acts, which are functions $f : S \to X$. Following a standard abuse of notation, let $x \in \mathscr{F}$ denote the constant act that returns $x \in X$ in every state. For any event A and acts $f, g \in \mathscr{F}$, let fAg denote the act h such that $h(s) = f(s)$ for $s \in A$ and $h(s) = g(s)$ for $s \in A^c$.

Let $\Delta(S)$ denote the set of probability distributions over S, which is identified with the $|S| - 1$ dimensional simplex in $\mathbb{R}^{|S|}$. For any $\mu \in \Delta(S)$ and any $A \in \Sigma$, let $\mu_{|A}$ (or sometimes $\mu(\cdot|A)$) denote the Bayesian update of μ given A[4].

Definition 1.1 (Scenario)**.** I will refer to an information-choice pair, (A, f), as a *scenario*.

I take as a primitive a class of preference relations $\{\succsim, \succsim_{A,f}\}_{(A,f)\in\Sigma\times\mathscr{F}}$ over $\mathscr{F}$.

[2]The assumption of finite S is merely for convenience. All results are unchanged if I assume an infinite state space and restrict attention to non-null events. What is crucial is the existence of at least three non-null events.

[3]I assume that the agent's information is in fact informative.

[4] That is, for all $B \in \Sigma$, $\mu(B|A) = \frac{\mu(B\cap A)}{\mu(A)}$.

Here $\succsim$ represents the agent's *ex-ante* preferences, while $\succsim_{A,f}$ is interpreted as his preference after making choice f and receiving information A; his preferences in scenario (A, f).[5]

The literature offers two interpretations for conditional preferences: 1) $\succsim_{A,f}$ represents what the agent thinks his preferences would be if he later faced scenario (A, f); and 2) $\succsim_{A,f}$ are his actual preferences when facing scenario (A, f). For this paper I adopt the second interpretation and assume that $\succsim_{A,f}$ is a representation of how the agent actually responds *ex-post* in scenario (A, f).

The statement $h \succsim_{A,f} g$ may be interpreted as follows: after having chosen f and learning A, the agent prefers h to g. In this way preference statements may be connected to choice data by observing choices from binary menus.[6] Thus the primitives may be interpreted as follows (i) at time 1 the experimenter elicits the agent's preferences, then (ii) during an interim period the agent chooses an alternative and then receives information, and finally (iii) at time 2 the experimenter elicits the agent's preferences again. The interim choice is not modeled and could be made from some subset of acts (a menu) or possibly utilize some form of randomization.[7]

[5]Alternatively, I could simplify notation by writing the above conditions only in terms of $\succsim_{A,f}$, for $A \in \Sigma^* = \Sigma \cup S$, and imposing the condition that $\succsim_{S,f} = \succsim_{S,\hat{f}}$, for all $f, \hat{f}$, in which case we identify $\succsim$ with $\succsim_{S,f}$. This changes the interpretation of the model, in that I interpret $\succsim$ as preferences before both information and choice, whereas in the new formulation all preferences are conditional on some action. However, it is reasonable to argue that even in the presence of cognitive dissonance, if no new information is received, (i.e., the agent observes S) then the fact that a choice was made does not immediately impact preferences. That is, I argue that dissonance requires information. Additionally, the fact that the model implicitly assumes $\succsim_{S,f} = \succsim_{S,\hat{f}}$ makes for a clear distinction between the model here and models of status-quo bias.

[6]It is simple to translate the framework and axioms into an inter-temporal choice setting. I utilize preferences as a primitive for axiomatic transparency.

[7]The specifics of how the choice is made are not important, since I am interested in how the existence of a choice affects conditional preferences. Many of the early psychology experiments on cognitive dissonance utilized a method of *forced compliance*. This suggests that it is the performance of the action that is relevant, rather than the specifics of the choice environment, and the belief changes are mechanical reactions.

1.3 Axioms

1.3.1 The Standard Axioms

The first axiom, Consistent Expected Utility, is a collection of classic axioms which are known to be equivalent to subjective expected utility maximization, plus ordinal preference consistency, which is a regularity condition between the *ex-ante* and *ex-post* preferences. Ordinal preference consistency is sensible in this environment, even in the presence of cognitive dissonance. In particular, suppose X is a set of monetary prizes, say $[0, 10]$. Then the condition merely imposes that if $10 is preferred to $1 before information, it is also preferred after information, regardless of what the agent has done in the past.

Axiom 1.1 (Consistent Expected Utility)**.** For all $A \in \Sigma$ and $g, h, f, \hat{f} \in \mathscr{F}$.

Weak Order: $\succsim$ and $\succsim_{A,f}$ are complete and transitive binary relations on $\mathscr{F}$.

Independence: For all $\alpha \in (0, 1)$.

(i) $f \succsim g$ if and only if $\alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h$.

(ii) $f \succsim_{A,\hat{f}} g$ if and only if $\alpha f + (1-\alpha)h \succsim_{A,\hat{f}} \alpha g + (1-\alpha)h$.

Strict Monotonicity: If $h(s) \succsim g(s)$ for all s, then $h \succsim g$. In addition, if for some s, $h(s) \succ g(s)$, then $h \succ g$. Similarly, if $h(s) \succsim_{A,f} g(s)$ for all s, then $h \succsim_{A,f} g$, and if in addition $h(s) \succ_{A,f} g(s)$ for some s, then $h \succ_{A,f} g$.

Continuity: The sets $\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succsim h\}$,$\{\alpha \in [0,1] : h \succsim \alpha f + (1-\alpha)g\}$, $\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succsim_{A,\hat{f}} h\}$ and $\{\alpha \in [0,1] : h \succsim_{A,\hat{f}} \alpha f + (1-\alpha)g\}$ are closed.

Ordinal Preference Consistency: $x \succsim y$ if and only if $x \succsim_{A,f} y$.

Non-triviality: There are $x, y \in X$ such that $x \succ y$.

Consequentialism is also a standard axiom, which ensures that the agent *believes* the information. That is, once the agent learns states outside of A are impossible, he is only concerned with how actions perform within A. That is, the agent's posterior beliefs put probability 1 on A.

Axiom 1.2 (Consequentialism)**.** For all $A \in \Sigma$ and $f \in \mathscr{F}$,

$$h(s) = g(s) \text{ for all } s \in A \implies h \sim_{A,f} g.$$

The novel axioms will be concerned with *how* the agent changes his preferences after receiving information. Before introducing them I first introduce another classic axiom.

Axiom 1.3 (Dynamic Consistency)**.** For all $h, g \in \mathscr{F}$ and $A \in \Sigma$,

$$hAg \succsim g \iff h \succsim_{A,f} g.$$

This axiom requires that the ranking of two acts, after the arrival of information A, only depends on their variation in A *and* is consistent with the agent's *ex-ante* ranking between acts that only vary within A. In other words, the agent does not adjust the relative probabilities of states in A. While the axiom has normative appeal, it is too restrictive and rules out any sensitivity to cognitive dissonance. In particular, an agent sensitive to cognitive dissonance allows yesterday's action to influence today's preferences. However, dynamic consistency requires preferences to be independent of yesterday's action. The following example illustrates this.

Example 1.1 (Investing)**.** Consider an example similar to the experimental setup in [43]. An investor is deciding between a stock, s, and a bond, b. The stock can be *good* or *bad* and the stock's terminal value depends on its type. In the interim, the

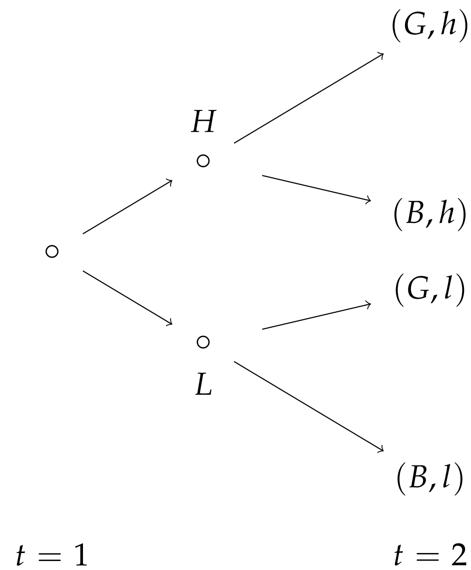

Figure 1.1: States and Information

stock can pay either a high or low dividend. Good stocks are more likely to pay high dividends.

Formally, the state space is $S = \{(G,h),(G,l),(B,h),(B,l)\}$ and μ is the agent's prior. Suppose that good and bad are equally likely but a high dividend correlates with a good stock—$\mu(G,h) = \mu(B,l) = \frac{3}{8}$ and $\mu(B,h) = \mu(G,l) = \frac{1}{8}$.[8] Let $H(L)$ be the event that a high (low) dividend is observed.

If agents satisfy dynamic consistency, then $\mu_{H,s}(G) = \mu_{H,b}(G)$. However, the experimental evidence finds that subjects report $\mu_{H,s}(G) > \mu_{H,b}(G)$. Formulated in terms of observables, we can consider preferences between two the following two bets:

	(G,h)	(B,h)	(G,l)	(B,l)
f	\$3	\$0	\$0	\$0
g	\$0	\$9	\$0	\$0

Before the dividend is observed the agent is indifferent between the two bets, $f \sim g$. Dynamic consistency requires the agent to also be indifferent after the div-

[8]These probabilities match those used in [43]

idend observation—$f \sim_{H,s} g$ and $f \sim_{H,b} g$—regardless of which asset the agent originally choose. However after observing the dividend, the stock holder is more optimistic than the bond holder and thus we expect $f \succ_{H,s} g$, a violation of dynamic consistency

1.3.2 Behavioral Axioms

As seen in Example 1.1, Dynamic Consistency is too strong and rules out cognitive dissonance. This is because it requires that both those that voted for the hawk and those that voted for the dove have the same posterior after observing A. It requires that relative likelihoods between states are constant *and* independent of prior actions, whereas cognitive dissonance implies that relative likelihoods change in favor of *states that favor prior actions*. The following behavioral axiom, Dynamic Cognitive Dissonance, is the precise weakening of Axiom 2.2 needed for the behavior seen in Example 1.1.

Axiom 1.4 (Dynamic Cognitive Dissonance)**.** For all (A, f) and $B, C \subset A$, such that for every $s \in B$ and $\tilde{s} \in C$, $f(s) \succsim f(\tilde{s})$, then for any $x, y, z \in X$, where $x, y \succsim z$,

$$xBz \succsim yCz \implies xBz \succsim_{A,f} yCz.$$

Axiom 1.4 states that for any pair of events, where one event is always better according to the previous action, the agent weakly prefers to bet on the better event. That is, think about an agent committed to action f. This induces a preference over what events might occur, where the agent prefers s to $\tilde{s}$ if $f(s) \succ f(\tilde{s})$. Since cognitive dissonance causes the agent to align his conditional beliefs with f, if he were to learn that $\{s, \tilde{s}\}$ has occurred he ought to prefer betting on s to $\tilde{s}$. In the context of the voting example, what this axiom states is that the voter will become (weakly) more convinced of war if he voted for the hawk and (weakly) more convinced of peace if he voted for the dove. Since agents are expected utility maximizers, such a belief distortion will cause the agent to support more extreme policy positions in future elections.

1.4 The General Representation

This section introduces the main model for cognitive dissonance. I discuss several properties of the general model and then present a representation theorem, showing that the model is equivalent to some standard postulates and the behavioral weakening of Dynamic Consistency—Dynamic Cognitive Dissonance. I then discuss the model's uniqueness properties.

Definition 1.2 (Cognitive Dissonance Representation)**.** There exists utility function $u : X \to \mathbb{R}$, a prior belief $\mu \in \Delta(S)$, and for each (A, f), an increasing distortion function $\delta_{A,f} : u(X) \to \mathbb{R}_+$ such that

$\succsim$ is represented by:

$$V(g) = \sum_{s \in S} u(g(s))\mu(s),$$

$\succsim_{A,f}$ is represented by:

$$V_{A,f}(g) = \sum_{s \in A} u(g(s))\mu_{A,f}(s),$$

where

$$\mu_{A,f}(s) = \delta_{A,f}(u(f(s)))\mu(s|A).$$

This is the most general model of belief distortions, where the distortion function depends upon the scenario (A, f) in a general way, not just through the payoff profile of f on A and a normalizing constant. In particular, the distortion magnitudes may vary considerably between scenarios, even if state-wise payoffs are identical. This scenario sensitivity may or may not be sensible in certain contexts, both of which will be explored later. This model embeds the standard Bayesian model as a special case, where $\delta_{A,f}(a) = 1$ for all $a \in u(X)$.

To clarify, the agent does not privilege the past choice, f, *per se*. Rather, he privileges those states of the world in which, given the partition he finds himself in, his past choice would do best. Thus the agent may in fact move away from f after information. That is, he views f as a commitment to certain states of the world and hence his beliefs become biased in favor of those states. Alternatively, one may interpret this as if the agent views his action f as an additional piece of information and uses this information to adjust the relative probabilities of states $s \in A$. This last interpretation is reminiscent of Bénabou and Tirole [9], which considers a model of an agent who infers their beliefs from their past actions.

Consider some scenario (A, f) and suppose $f(s) \succsim f(\tilde{s})$ for some $s, \tilde{s} \in A$. Then the agent's *ex-post* subjective relative likelihood of state s to state $\tilde{s}$ is given by

$$\frac{\mu_{A,f}(s)}{\mu_{A,f}(s)} = \frac{\delta_{A,f}(u(f(s)))}{\delta_{A,f}(u(f(\tilde{s})))} \times \frac{\mu(s)}{\mu(s)}.$$

Since $\delta_{A,f}$ is increasing, and $f(s) \succsim f(\tilde{s})$, it follows that $\frac{\delta_{A,f}(u(f(s)))}{\delta_{A,f}(u(f(\tilde{s})))} \geq 1$, hence the agent believes s to be *relatively* more likely than $\tilde{s}$ when compared to a Bayesian agent.

Notice that if $f(s) \sim f(\tilde{s})$, then $\delta_{A,f}(u(f(s))) = \delta_{A,f}(u(f(\tilde{s})))$ and hence $\frac{\mu_{A,f}(s)}{\mu_{A,f}(s)} = \frac{\mu(s)}{\mu(s)}$. Thus the relative likelihood between states that provide identical payoffs under f is undistorted. This is actually quite intuitive, since whatever feelings the agent has toward s, since both states s and $\tilde{s}$ are equally good according to his action, he should have precisely the same feelings toward $\tilde{s}$.

To further this intuition, consider an agent having taken a constant action x.

Then regardless of the information he learns, there is no possible distortion of this information that could increase the agent's valuation of x, hence there is no distortion of beliefs at all. This reasoning extends to all scenarios that are equivalent to having taken a constant action. The following definition precisely identifies which scenarios are equivalent to a constant action.

Definition 1.3. A scenario (A, f) is constant if for all $s, \tilde{s} \in A, f(s) \sim f(\tilde{s})$.

A constant scenario is one in which the agent's initial action does not vary, conditional on event A. Thus conditional on A, there is no distortion of beliefs that can improve the valuation of f. Let $\mathcal{C}$ denote the set of constant scenarios.

Observation 1.1. *For all $(A, f) \in \mathcal{C}$, the agent's posterior beliefs are derived via Bayes rule, $\mu_{A,f} = \mu_{|A}$.*

1.4.1 Representation and Uniqueness

This section presents the representation theorem and the uniqueness properties of the representation. The following theorem connects the representation to the axioms.

Theorem 1.1 (Representation)**.** *The following are equivalent:*

(i) $\{\succsim, \succsim_{A,f}\}$ satisfy Consistent Expected Utility, Consequentialism, Dynamic Cognitive Dissonance,

(ii) The agent admits a Cognitive Dissonance Representation.

The uniqueness properties of the representation are illustrated in the following theorem. It is standard to show that the utility index u is unique up to positive affine transformations. The uniqueness of μ is also standard. The uniqueness of δ follows from the uniqueness of the subjective probabilities. That is, from standard results we have a unique $\mu_{A,f}$ for every scenario, while uniqueness of δ follows from the decomposition into the product $\delta_{A,f}(f(s))\mu(s|A)$.

Theorem 1.2 (Uniqueness)**.** *If* $(u, \mu, \delta_{A.f})$ *and* $(u', \mu', \delta'_{A.f})$ *both represent* $\{\succsim, \{\succsim_{A,f}$ $\}\}$ *then*

(i) $u' = \alpha u + \beta$ *for* $\alpha > 0, \beta \in \mathbb{R}$

(ii) $\mu' = \mu$

(iii) $\delta'_{A,f}(u'(x)) = \delta_{A,f}(u(x))$ *for all* $x \in f(A)$.

1.5 Proportional Distortions

This section introduces and characterizes the first special case by introducing a *scenario independence* property. While the general representation allows for some arbitrariness between scenarios, it is natural to think that the distortions between any two states should only depend on the relative payoffs between those states. To this end I define the proportional distortion.

Definition 1.4 (Proportional distortion)**.** The belief distortion function is a *proportional distortion* if there exists an increasing function $v : u(X) \to \mathbb{R}_+$ such that

$$\delta_{A,f}(a) := \frac{v(a)}{\sum_{s \in A} v(u(f(s)))\mu(s|A)}. \tag{1.1}$$

In the case of a proportional distortion, the belief distortion only depends on the scenario up to a normalizing constant. This becomes clear when looking at the probability ratio of any two states. Suppose $f(s) \succ f(\tilde{s})$, then after learning some A containing both states, it is as if the agent takes (A, f) as a signal that s has been made more likely than $\tilde{s}$ by an amount proportional to $\frac{v(u(f(s)))}{v(u(f(\tilde{s})))}$.

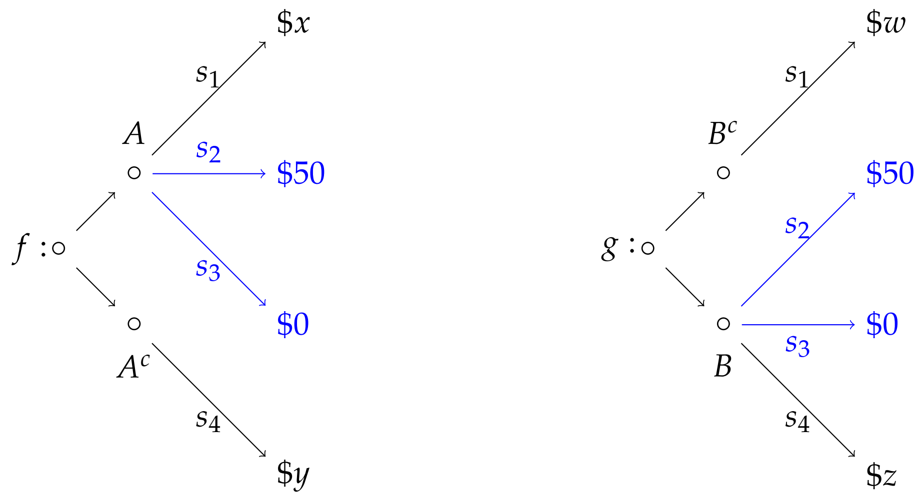

Figure 1.2: Scenarios (A, f) and (B, g).

Example 1.2. More precisely, suppose $S = \{s_1, s_2, s_3, s_4\}$, and consider the following two events $A = \{s_1, s_2, s_3\}$ and $B = \{s_2, s_3, s_4\}$ and the following two acts:

	s_1	s_2	s_3	s_4
f	$\$x$	$\$50$	$\$0$	$\$y$
g	$\$w$	$\$50$	$\$0$	$\$z$

Figure 1.2 illustrates scenarios (A, f) and (B, g). Suppose $C = \{s_2, s_3\} = A \cap B$ (shown in blue) and consider the agent's preferences over acts that vary only within C. More specifically, consider an agent placing bets on a given state, say bets of the form $(\$a, s_2)$ and $(\$b, s_3)$. If an agent is in scenario (A, f), then since $50 > 0$, an agent sensitive to cognitive dissonance may distort the relative probabilities of states s_2 and s_3. It is natural to think that when determining preferences over the binary bets, the *only* relevance of f is through how it performs in states s_2 and s_3. That is, the specific value of y is irrelevant. If this is the case, then since $g(s_2) = f(s_2)$ and $g(s_3) = f(s_3)$, the agent should report the same preferences over binary bets of the form $(\$a, s_2)$ and $(\$b, s_3)$ when in scenario (B, g).

The following scenario independence axiom imposes precisely the intuition from the previous example.

Axiom 1.5 (Scenario Independence)**.** For all (A, f), (B, g) and $C \subset A \cap B$, if $f(s) \sim g(s)$ for all $s \in C$, then for all $h, j \in \mathscr{F}$ and any $z \in X$,

$$hCz \succsim_{A,f} jCz \Longleftrightarrow hCz \succsim_{B,g} jCz.$$

For any two scenarios, if there is some event in which both action f and action g are payoff equivalent, then the preference ordering between any two acts that vary only on that event is the same in either scenario. The next theorem shows that the addition of Scenario Independence completely characterizes proportional distortions.

Theorem 1.3 (Representation)**.** *Suppose* $\{\succsim, \succsim_{A,f}\}$ *satisfy Consistent Expected Utility, Consequentialism, Dynamic Cognitive Dissonance, then the following are equivalent,*

(i) $\{\succsim, \succsim_{A,f}\}$ *satisfy Scenario Independence*

(ii) $\delta_{A,f}$ *is a proportional distortion*

While Theorem 2 shows the uniqueness of $\delta_{A,f}$, the same uniqueness does not extend to the value function determining a proportional distortion. That is, v is only identified up to the ratio of $\delta_{A,f}(a)$ and $\delta_{A,f}(b)$, as shown in the following theorem.

Theorem 1.4 (Uniqueness)**.** *Suppose* $\{\succsim, \succsim_{A,f}\}$ *has a cognitive dissonance representation with a proportional distortion. Then the value function* v *is unique up to a positive scalar.*

So far I have not required the distortion to be continuous. However, one may impose continuity of the distortion function through the following axiom.

Axiom 1.6 (Commitment Continuity)**.** For all A and all $f, g, h\mathscr{F}$ and any $(f_n), (g_n), (h_n) \in \mathscr{F}^\infty$ such that $f_n \to f, g_n \to g, h_n \to h$:

$$\text{if } g_n \succsim_{A,f_n} h_n \text{ for all } n, \text{ then } g \succsim_{A,f} h.$$

As will be shown later however, there is an interesting class of distortions that are not continuous in the sense of Axiom 1.6.

Corollary 1.1. *Suppose* $\{\succsim, \succsim_{A,f}\}$ *satisfy Consistent Expected Utility, Consequentialism, Dynamic Cognitive Dissonance, Scenario Independence, then the following are equivalent:*

(i) The value function v *is continuous*

(ii) $\{\succsim, \succsim_{A,f}\}$ *satisfy Commitment Continuity*

1.5.1 Examples

The following examples may help illustrate the distinction between continuous and discontinuous proportional distortions.

Example 1.3 (Step Distortion)**.** Fix a parameter $\theta \in (0, 1)$ and some $a^* \in u(X)$. Then

$$v(a) = \begin{cases} 1+\theta & \text{if } a \geq a^* \\ 1-\theta & \text{if } a < a^* \end{cases}$$

An agent described by a step distortion is one that divides the consequence space into *good* and *bad* outcomes, defined relative to some threshold a^*. The agent's relative beliefs about states that yield consequences of the same type are undistorted, while the agent overweights good states relative to bad states. If $X = \mathbb{R}$, then a natural division is at $a^* = 0$. In this case the agent will overweight the probability of gains relative to losses. In this case, any scenario (A, f) that yields only gains or only losses will result in posterior beliefs that coincide with using Bayes' rule. Non-Bayesian behavior in this example would only be observed when observing *mixed* scenarios - those in which (A, f) allows for both gains and losses.

Example 1.4 (Logistic Distortion)**.** Fix a parameter $\lambda \in [0, \infty)$. Then $v(a) = e^{\lambda a}$ is a logistic distortion, where

$$\delta_{A,f}(u(x)) = \frac{e^{\lambda u(x)}}{\sum_{\tilde{s} \in A} e^{\lambda u(f(\tilde{s}))} \mu(s|A)}$$

and the likelihood ratios are given by

$$\frac{\mu_{A,f}(s)}{\mu_{A,f}(\tilde{s})} = \frac{\mu(s)}{\mu(\tilde{s})} e^{\lambda[u(f(s)) - u(f(\tilde{s}))]}$$

A version of the logistic distortion was studied by Mayraz [46]. The logistic distortion includes the Bayesian model as the special case $\lambda = 0$.

Example 1.5 (Distortion with Decreasing Sensitivity)**.** Suppose $u(X) = [0, \infty)$. Then define $v : u(X) \to \mathbb{R}$ by

$$v(a) = \ln(1 + a)$$

.

An agent with such a distortion function will be approximately Bayesian on high stakes events (relative beliefs between high payoff states are accurate), while he will dramatically under-weight the probability of low payoff states.

1.6 Binary Distortions

While continuity is often considered an attractive property, there are interesting cases in which the distortion is not continuous. For example, the distortion may take the form of a step function. This may be interpreted as an agent separating the event A into *good* and *bad* states, and only overweighting the probability of good states relative to bad states. In contrast to the proportional distortion, in which the scenario only matters for a normalization, the class of binary distortions are often most sensible when there is scenario dependence.

That is, the agent determines some threshold depending on (A, f), and states which are better than the threshold are classified as *good*. The following definition introduces a special case of the binary distortion.

Definition 1.5 (Best-case Binary Distortion)**.** The belief distortion is a *Best-case Binary Distortion*[9] if there exists a non-constant, affine utility function $u : X \to \mathbb{R}$, a probability distribution $\mu \in \Delta(S)$, and a function $\delta : \Sigma \times \mathscr{F} \to [0,1]$ such that:

$\succsim$ is represented by:

$$V(g) = \sum_{s \in S} u(g(s))\mu(s),$$

$\succsim_{A,f}$ is represented by:

$$V_{A,f}(g) = \sum_{s \in A} u(g(s))\mu_{A,f}(s)$$

where

$$\mu_{A,f}(s) = (1 - \delta(A,f))\mu(s|A) + \delta(A,f)\mu(s|\mathscr{D}(A,f))$$

[9]This representation fits into the general model as a binary distortion given by the pair (θ, t), where $\theta : \Sigma \times \mathscr{F} \to [0,1]$ and $t : \Sigma \times \mathscr{F} \to X$ such that

$$\delta_{A,f}(x) := \begin{cases} 1 - \theta(A,f) & \text{if } x \prec t(A,f) \\ 1 - \theta(A,f) + \theta(A,f)\frac{\mu(A)}{\mu(\{s \in A | f(s) \succsim t(A,f)\})} & \text{if } x \succsim t(A,f) \end{cases}$$

and

$$\mathscr{D}(A, f) = \{s \in A | f(s) \succsim f(s') \text{ for all } s' \in A\}.$$

1.6.1 Interpretations

Here V represents the *ex-ante* preference and $V_{A,f}$ represents the agent's preferences after information A and the choice of f. Behaviorally, δ represents the agent's *sensitivity to dissonance*. Presented in this form, it looks similar to the "α-maxmin expected utility" model of Arrow and Hurwicz [6]. In a way they are similar, as they both allow the agent to "average" the utility of an act according to a "pessimistic (rational)" belief and an "optimistic" belief. The similarity is only superficial, however, since in this model the agent's preferences satisfy the subjective expected utility axioms. Hence the agent being studied has beliefs represented by a unique probability distribution over the states, where $\mu_{A,f}(s) = (1-\delta)\mu_{|A}(s) + \delta\mu_{\mathscr{D}(A,f)}(s)$.

One can think of an agent that satisfies the above axioms as one who systematically mixes his beliefs with what he wishes would actually occur. That is, while he knows the true state lies within the event A, he believes, given his previous choice of f, that states in which f is good must be more likely than he originally supposed, because he must have chosen f for good reason. Note that states in which f is good is defined relative to how f performs in other states *within* A. Alternatively, one could think of the agent as being composed of two selves: one of which is a realist (and is Bayesian) and one of which is a *justifier*, who interprets the information to maximize the conditional value of the action. When the agent receives information, he must balance the desires of both selves, and δ is the weight that the justifying self receives when making decisions. Since the preferences are represented by expectation with respect to a weighting between the Bayesian belief and a justifying belief, the agent will generally make choices that are seen as a *compromise* between (i) accepting the information and taking the correct action, and (ii) "sticking to your guns" by repeating the previous choice.

This representation suggests a specific "cognitive mechanism" that underlies multi-period choice and updating. The agent considers a scenario, an information-choice pair, and partitions the event into good and bad states. As mentioned earlier, good states are states in which f performs best given the known information. This is a rather blunt cognitive rule, since even if two states are very close in utility space, they may be classified as "distinct states." Hence preferences are generally discontinuous between scenarios.

Example 1.6. In order to get a better understanding of how this bias affects preferences, consider the following simple example. Say $X = [w, b] \subset \mathbb{R}$ and $u(x) = x$. Suppose $S = \{s_1, s_2, s_3\}$ with prior $\mu = (1/3, 1/4, 5/12)$, and suppose $A = \{s_1, s_2\}$. Consider an act $f = (y, x, z)$, where $x \succ y \succ z$. Then the Bayesian posterior is given by $\mu_{|A} = (4/7, 3/7)$ and the corresponding indifference curve in utility space, illustrated in Figure 1.3, corresponds to the solid line ($\delta = 0$). Since $x \succ y$, then $\mathscr{D}(A, f) = \{s_2\}$ and $\mu_{|\mathscr{D}(A,f)} = (0, 1)$. The horizontal dotted line denotes the indifference curve of an agent that has taken action f with $\delta = 1$, whereas the intermediate dashed line represents the corresponding indifference curve for $\delta = \frac{1}{2}$. For clarity, the curves all intersect at the constant utility line $y = x$.

1.6.2 Characterization

Before introducing the next axiom I first introduce a few definitions. The first is comonotonicity, which is standard in the literature.

Definition 1.6. Say that h and f are *strongly A-comonotonic* if $h(s) \succsim h(s')$ if and only if $f(s) \succsim f(s')$ for all $s, s' \in A$. Denote this by $h \asymp_A f$.

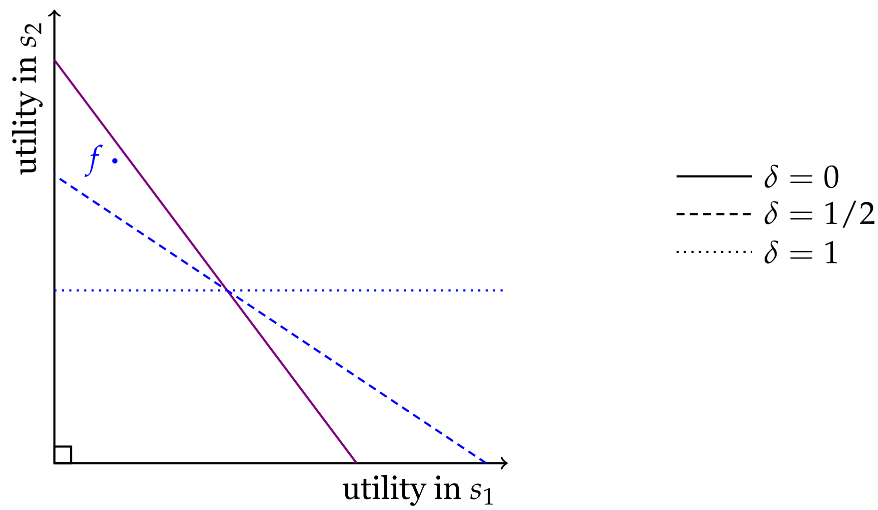

Figure 1.3: Indifference curves for $\succsim_{A,f}$.

Consider an agent that has taken action f and is sensitive to cognitive dissonance. Then one would expect the agent to express an inflated view about the value of f. However, dissonance theory suggests more than this. That is, the agent will seek out a *consistent* view of the world, and hence will similarly express an inflated view about the value of actions that are similar to f. Thus we seek to impose a weakening of dynamic consistency that takes into consideration the impact of the action f on all similar acts. This leads to the following axiom.

Axiom 1.7 (Best-Case Dominance)**.** For all A, and all $h, g \preceq_A f$:

$$\left.\begin{array}{l} hAg \succsim g \\ h(s) \succsim g(s'), \text{ for some } s \in A \text{ and all } s' \in A \end{array}\right\} \implies h \succsim_{A,f} g$$

This is a three part weakening of dynamic consistency. First, I only apply a single direction. That is, I only seek to regulate when an *ex-ante* preference is sufficient to make an *ex-post* comparison. Second, in addition to the classic conditioning requirement, I require the additional property stated. To understand this requirement, consider an agent originally choosing action f over action g, where $fAg \sim g$ held *ex-ante*. That is, before the agent took an action, if he knew he would learn A, he would say he was indifferent between taking either action. However, after taking an action and having observed A, the agent feels the need to justify having taken f as opposed to g. Thus we come to the second condition — if the best possible payoff of action f is better than the best possible payoff of action g, then the agent can use that as a justification to declare a preference for f over g. Third, I extend this logic to all actions that are strongly A-comonotonic with f. That is, imagine an agent who is contractually obligated to some action f. Then this commitment to f induces a preference over states of the world. However, an action h that is strongly A-comonotonic with f induces the same preference over states, and hence in a sense they are equivalent. That is, they induce the same desires about *how the world turns out* and thus if commitment to f induces a dynamic inconsistency, any arguments or justifications used for f also apply to h.

Theorem 1.5 (Representation)**.** *The following are equivalent:*

(i) $\{\succsim, \succsim_{A,f}\}$ *satisfy Consistent Expected Utility, Consequentialism, Dynamic Cognitive Dissonance, and Best-Case Dominance.*

(ii) The agent has a Best-case Binary Distortion representation.

Theorem 1.6 (Uniqueness)**.** *Moreover, if* (u, μ, δ) *and* (u', μ', δ') *represent the same preferences, then there is some* $\alpha > 0, \beta \in \mathbb{R}$*, such that* $u' = \alpha u + \beta$*,* $\mu = \mu'$*, and* $\delta(A, f) = \delta'(A, f)$ *for all* (A, f)*.*

To gain some additional intuition for Axiom 1.7, consider an agent who displays an *extreme* level of dissonance. That is, such an agent maintains a preference for f over g if there exists a possibility of f being better than anything g might return. One may think of the agent reasoning as follows: *I must have chosen* f *for good reason, and if state* s *is the true state, the* f *is better than* g*. Hence it must be that state* s *will be realized.* This is formalized in the following axiom:

Axiom 1.8 (Extreme Dissonance)**.** For all $f, g \in \mathscr{F}$, if there is some $s \in A$ such that $f(s) \succsim g(s')$ for all $s' \in A$, then

$$f \succsim_{A,f} g$$

Extreme Dissonance is essentially the second condition of Best-Case Dominance. Thus one can think of Best-Case Dominance as *softening* Extreme Dissonance and asserting that whenever both a Bayesian agent and an agent that is *maximally sensitive* to dissonance prefers f to g, then an agent of any sensitivity to dissonance also prefers f to g. The following theorem shows that Extreme Dissonance indeed does characterize a *maximally sensitive* agent, while the Bayesian agent is the opposite extreme.

Theorem 1.7. *Suppose* $\{\succsim, \succsim_{A,f}\}$ *satisfies Axiom 1.1, Axiom 3.2, then*

(i) $\{\succsim, \succsim_{A,f}\}$ *satisfy Axiom 2.2 if and only if for all* (A, f), $\delta(A, f) = 0$.

(ii) $\{\succsim, \succsim_{A,f}\}$ *satisfy Axiom 1.4, Axiom 1.8 if and only if for all* (A, f), $\delta(A, f) = 1$.

1.7 Connecting the Two Cases

So far I have presented two special cases of the general model, which each require a single additional axiom. Both the proportional and best-case binary distortions have intuitive appeal. The proportional distortion allows for belief distortions to be somewhat independent of the previous action and allows for beliefs to vary in a continuous sense. The best-case binary model is suggestive of a simple cognitive mechanism in which the agent simply thinks of *good* or *bad* states and gives greater weight to the good states.

However, it remains to see how the proportional distortion and the best-case binary distortion relate to each other. That is, this section asks what model of behavior is consistent with both Scenario Independence (Axiom 1.5) and Best-Case Dominance (Axiom 1.7) holding. It turns out that both special cases are distinct in a very strong sense - an agent may satisfy both conditions only if the agent is in fact a Bayesian.

Theorem 1.8. *Suppose $\{\succsim, \succsim_{A,f}\}$ satisfy Axiom 1.1, Axiom 3.2, Axiom 1.4. Then the following are equivalent*

(i) $\{\succsim, \succsim_{A,f}\}$ satisfy Scenario Independence and Best-Case Dominance

(ii) $\{\succsim, \succsim_{A,f}\}$ satisfy Dynamic Consistency

This theorem therefore shows that there is a trade-off between scenario independence and continuity and cognitive simplicity. Further, continuity of the agent's beliefs is not a purely technical assumption because it is violated by the best-case binary model. Finally, that fact that there is a sharp distinction between the two models allows us to design experimental procedures to distinguish between the two cases and gain a much deeper understanding of the mechanism through which beliefs are distorted.

Despite this strong distinction, they also *almost* coincide in the extreme case. That is, the Best-case Binary distortion with $\delta(A, f) = 1$ is the limit of a proportional representation as sensitivity to dissonance increases without bound. This result is illustrated by the following corollary.

Corollary 1.2. *Let $\mu^{\lambda}_{A,f}$ denote logistic distorted beliefs with sensitivity parameter λ. Then*

$$\lim_{\lambda \to \infty} \mu^{\lambda}_{A,f} = \mu_{|\mathscr{D}(A,f)}$$

Thus it makes sense to consider $\delta(A, f)$ as a measure of dissonance, since as the parameter increases behavior approaches the most extreme version of dissonance.

1.8 Comparative Dissonance

This section considers comparing individuals' sensitivity to cognitive dissonance. That is, this section asks when can an experimenter conclude that one agent is *more sensitive* to dissonance than another. Consider two agents that satisfy the conditions of Theorem 1. For $i = 1, 2$, let $\{\succsim^i, \succsim^i_{(A,f)}\}$ denote i's preferences. The following definition is similar in spirit to definitions of *more ambiguity averse* or *more status-quo biased*.

Definition 1.7. Given two agents, with preferences $\{\succsim^1, \succsim^1_{A,f}\}$ and $\{\succsim^2, \succsim^2_{A,f}\}$, agent 2 is *more sensitive to dissonance* than agent 1 if $\succsim^2 = \succsim^1$ and for all (A, f), $f \succsim^1_{A,f} x \Rightarrow f \succsim^2_{A,f} x$

The following result relates the preference based definition of *more sensitive to dissonance* to model parameters.

Theorem 1.9. *Suppose agents* 1 *and* 2 *have cognitive dissonance representations and agent 2 is more sensitive to dissonance than agent 1. Then*

(i) If both agents have best-case binary distortions, $\delta^2(A, f) \geq \delta^1(A, f)$ *for all* (A, f).

The theory of cognitive dissonance has previously lacked a method for measuring dissonance within individuals and comparing between individuals. The framework presented here, in which information is observable to the experimenter, provides a precise way to do both while the above theorem demonstrates that the comparative measure in fact corresponds to a sensible, preference characterization of *more sensitive to dissonance*.

1.9 Applications

1.9.1 A Simple Asset Pricing Problem

For simplicity, uncertainty is represented by four states, $S = \{u^h, u^l, d^h, d^l\}$, and interim information is given by: $\{\{u^h, u^l\}, \{d^h, d^l\}\} = \{U, D\}$. That is, the agent will receive news of the form *the asset will go up* or *the asset will go down*. There is a single risk-free asset, b, which pays b at time 3. There is single unit of risky asset in each period $f : S \to \mathbb{R}$ such that $f(d^l) < f(d^h) \leq f(u^l) < f(u^h)$. There is no discounting and no short selling, so that the agent may buy a single unit of the risky asset at each of time 1 or 2. Assume the initial prior $\mu \in \text{int}(\Delta(S))$ and for simplicity, that $\mu(U) = \mu(D)$.

1.9.1.1 Asset Pricing Without Dissonance: Rational Benchmark

As a benchmark, first consider prices when an agent is a standard Bayesian. In this case the bond and stock must both offer the same expected return to be traded in equilibrium. Hence

$$b - P_b = \mathbb{E}_\mu(f) - P_f^1$$

For simplicity, normalize the bond return to zero ($P_b = b$), hence $P_f^1 = \mathbb{E}_\mu(f)$. The conditional prices at time 2 are found similarly, and thus equal to the conditional expected payoff under Bayesian updating. The priced are illustrated in Figure 1.4

1.9.1.2 Prices with Cognitive Dissonance and a Naive Agent

Now I consider an agent that is sensitive to cognitive dissonance but at time 1 does not anticipate how his beliefs will be biased at time 2. Hence if agent is naive, the time $t=1$ valuation is equal to the rational price: $P^1_f = \mathbb{E}_\mu(f)$. Now, consider what happens if the agent purchase the risky asset. At time 2, after the news is released but before the final states are revealed, price must again equal the agent's expected valuation, $P^{1\delta}_f(U) = V_{U,f}$. However, under dissonance the agent's valuation is as follows:

$$V_{U,f}(f) = \delta_{U,f}(f(u^h))\frac{\mu(u^h)}{\mu(U)}f(u^h) + \delta_{U,f}(f(u^l))\frac{\mu(u^l)}{\mu(U)}f(u^l).$$

For simplicity, we can define δ by $1-\delta = \delta_{U,f}(f(u^l))$ and simple algebra yields the following pricing equation:

$$P^{2\delta}_f(U) = \left[(1-\delta)\frac{\mu(u^h)}{\mu(U)} + \delta\right] f(u^h) + \left[(1-\delta)\frac{\mu(u^l)}{\mu(U)}\right] f(u^l). \quad (1.2)$$

When $\delta = 0$ the agent is a Bayesian. As δ increases towards one (the agent is more sensitive to dissonance) then $V_{U,f}(f)$ increases towards $f(u^h)$, and hence the market price increases.

1.9.1.3 Prices with Cognitive Dissonance and Sophisticated Agent

In this case I consider an agent that anticipates the belief distortion after information and hence knows that by buying at $t=1$, he will overpay at $t=2$. In this case the agent will price the asset via backwards induction. Since the time 2 prices are given from above, all that remains is to determine a price at time 1 such that the agent is willing to buy the risky asset. Thus the agent takes time 2 prices as given and sets total expected return equal to purchasing the bond today.

$$b - P_b = \mathbb{E}_\mu(f) - P^{1\delta}_f + \frac{1}{2}\left[\mathbb{E}_\mu(f|U) - P^{2\delta}_f(U)\right] + \frac{1}{2}\left[\mathbb{E}_\mu(f|D) - P^{2\delta}_f(D)\right]. \quad (1.3)$$

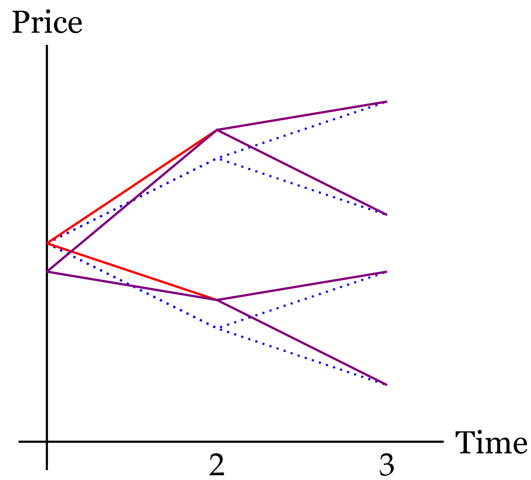

Figure 1.4: Predicted price paths for $f = (1, 3, 4, 6)$ and (i) the agent is rational (blue dashes), (ii) a naive agent experiences cognitive dissonance, $\delta = \frac{1}{2}$ (red), and (iii) a sophisticated agent experiences cognitive dissonance $\delta = \frac{1}{2}$ (purple)

Some algebra provides

$$P_f^{1\delta} = \mathbb{E}_\mu(f) + \delta\left(\mathbb{E}_\mu(f) - [\mu(U)f(u^h) + \mu(D)f(d^h)]\right). \tag{1.4}$$

When $\delta = 0$ the time 1 price corresponds to the rational (and naive) price. As δ increases P_f^δ decreases, since $\mathbb{E}_\mu(f) - [\mu(U)f(u^h) + \mu(D)f(d^l)] < 0$. Thus the equity premium at time 1 is given by

$$-\delta(P_f - [\mu(U)f(u^h) + \mu(D)f(d^h)]), \tag{1.5}$$

which is positive and increasing in δ. In this simple model a positive equity premium arises, even with risk neutrality. The cause of the premium is not simply due to a behavioral trait, but due to sophisticated agent trying to protect himself from future mistakes. That is, the sophisticated agent demands the premium today as a buffer for his inevitable bias tomorrow. Thus Cognitive Dissonance potentially contributes toward explaining the equity premium puzzle, Mehra and Prescott [47].

1.9.2 Response to Information

This section considers an agent satisfying the conditions of the best-case binary distortion and studies his response to information. I find that such an agent deviates from a Bayesian in a rather systematic way. Posterior beliefs are influenced by the agent's time 1 choice and are such that the agent always believes his time 1 choice is better than a Bayesian would. It is in this way that the agent exhibits an asymmetric reaction to news. By generally over-valuing his original action it is as if he over-reacts to news that is good for f, while under-reacting to news that is bad for f.

Research by Easterwood and Nutt [13] suggests that this behavior in fact occurs in financial markets. They study analysts' forecasts and find that analysts systematically under-react to negative information and overreact to positive information. Since analysts exhibit both under and overreaction (depending on information), this cannot be due to generic over(under)-reaction to information. For example, a model in which people systematically overreact to information predicts that, after bad news, they should have beliefs *more negative* than the information warrants, whereas the opposite is observed. This phenomena, however, is consistent with the model presented in this paper, under the presumption that an analyst's decision to cover a stock is seen as an implicit endorsement of the stock.

Definition 1.8. Say that A is *good news for* f if $fAx \succ x$ for some constant act satisfying $x \sim f$. Similarly, A is *bad news for* f if $x \succ fAx$.

By fixing a choice f, one can compare how an agent sensitive to cognitive dissonance reacts to various events. For any event A, if the agent has chosen f then he always believes that f provides greater *ex-post* expected utility than it would if the agent had used Bayes' rule. Since the best-case binary distortion is a special case of the general cognitive dissonance representation, the agent's posterior satisfies Bayes' rule for all constant scenarios. Let $\mathcal{N}$ denote the set of non-constant scenarios.[10]

Theorem 1.10. *If* $(A, f) \in \mathcal{N}$, *then* $\mathbb{E}_{\mu_{A,f}}(u(f)) > \mathbb{E}_{\mu_{|A}}(u(f))$.

Thus an agent that originally chose f will overreact to good news (for f) and under-react to bad news (for f). That is, whenever A is good news for f and f is non-constant on A, the agent overvalues f (relative to a Bayesian). Hence he is willing to pay more for act f, i.e., $\mathbb{E}_{\mu_{A,f}}(u(f)) > \mathbb{E}_{\mu_{|A}}(u(f))$. Similarly, when A is bad news the agent still overvalues f, and hence under-reacts to the negative information in A.

Additionally, Agrawal and Chen [1] provide evidence that analysts are more optimistic about firms that have relationships with their employer. This suggests that reference points may have an effect on how people interpret information. The differential treatment of affiliated and non-affiliated firms is not consistent with any type of non-Bayesian model without reference points, while it is consistent with the model presented here.[11]

1.9.3 Polarization

While the previous two applications are concerned with the implications of cognitive dissonance for a single individual, this section studies the effect of cognitive dissonance on the distribution of beliefs within a population. In particular, this section shows that whenever two agent take different actions, then even when they observe the same information and have identical prior beliefs they will have different posterior beliefs.

[10] $\mathcal{N} = \{(A, f) | f(s) \succ f(\tilde{s}) \text{ for some } s, \tilde{s} \in A\}$

[11] It should be acknowledged that both of these explanations require the joint assumption of *belief narrow framing*, where information is incorporated into an asset specific belief on a case-by-case basis. Narrow framing for risks has been argued for by Barberis et al. [7].

The seminal experiment on polarization comes from the psychology literature. Lord et al. [45] recruited subjects based on their differing views on the death penalty and presented them with identical essays. Afterwards their views were further apart, even though Bayesian updating predicts they should move closer together. For other explanations of polarization, see [55], [3].

Theorem 1.11. *Suppose* $\succsim^1 = \succsim^2$ *and for all* $A \in \Sigma$ *and* $f \in \mathscr{F}$, $\succsim^1_{A,f} = \succsim^2_{A,f}$, *and* v *is strictly increasing. For all* $A \in \Sigma$ *and* $f, g \in \mathscr{F}$, *if* (A, f) *and* (A, g) *are such that for some* $s, s' \in A$, $f(s) \nsim g(s)$ *and* $f(s') \sim g(s')$, *then* $\succsim^1_{A,f} \neq \succsim^2_{A,g}$, *hence* $\mu^1_{A,f} \neq \mu^2_{A,g}$.

That is, consider two individuals, 1 and 2, and suppose they begin with the same initial beliefs μ. For simplicity, I suppose both agents satisfy the conditions of the proportional distortion for some strictly increasing v. Then whenever the two individuals are in different, non-constant scenarios they will have different posterior beliefs.

Example 1.7. Consider the setup from Example 1.6 and two acts, $f = (y, x, z)$ and $g = (x, y, z)$ where $x \succ y \succ z$. The corresponding indifference curves for $\succsim^1_{A,f}$ and $\succsim^2_{A,g}$, in utility space, are illustrated in Figure 1.5. The solid purple line purple corresponds to the Bayesian posterior ($\delta = 0$). The horizontal (vertical) dotted line denotes the indifference curve of an agent that has taken action $f(g)$ with $\delta = 1$, whereas the intermediate dashed lines represents the corresponding indifference curve for $\delta = \frac{1}{2}$. For clarity, the curves all intersect at the constant utility line $y = x$.

Thus two agents that are identical in their prior beliefs and how they update their beliefs conditional on a given scenario can observe the same information and will typically have differing posterior beliefs when they have taken different actions. One point to note is that if both agents are in *constant scenarios* then their posterior beliefs should be identical. In the context of voting, this suggests that those who

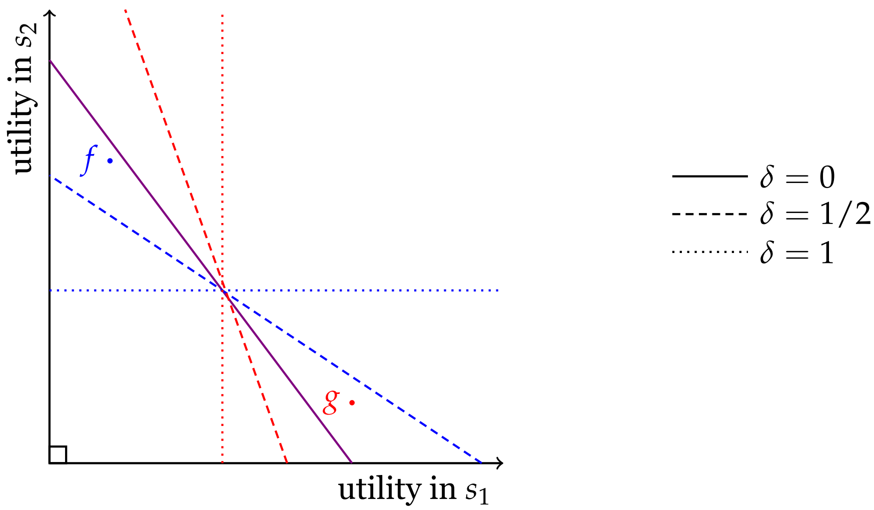

Figure 1.5: Indifference curves for $\succsim^1_{A,f}$ and $\succsim^2_{A,g}$.

abstain from voting or have not taken an initial stance on an issue will not exhibit polarization, while partisans will exhibit polarization.

However, polarization of beliefs is not simply restricted to the case when agents take different actions. If two agents have differing distortion functions then they may exhibit polarization even after both take the same initial action and observe the same information. Given the abundance of experimental evidence suggesting that an agent's actions influence how they update their beliefs, the phenomenon of belief polarization should be observed more often than not.

1.9.4 Purchase of Safety Equipment

The model of Akerlof and Dickens (1982) [2] showed that workers subject to cognitive dissonance may forgo the purchase of efficient safety equipment. The key variable in their model was the *cost of fear*, which was determined by the true probability of an accident and the agent's chosen perception. I now show that their model can be adapted into my framework and that the adapted model yields similar conclusions.

Consider two industries, one safe and one hazardous. However, the true risk of the hazardous industry is not known to the agent. This could be because the industry is relatively new or due to societal mis-perception. The state space is $S = \{s, m, h\}$, where a state corresponds to the *true nature* (risk level) of the hazardous industry. Here, s means that the industry is actually safe, m means it is moderately risky, and h means it is highly risky. Further, suppose that once the job is accepted, the agent learns whether or not the true state is h, so that in period 2 he knows either $\{h\}$ or $\{s, m\} \equiv A$. Let q_j denote the probability of accident in state j, and suppose that $q_s = 0 < q_m < q_h$.

Following Akerlof and Dickens, in period 1 the agent chooses an industry and in period 2 he is given the option to purchase safety equipment. The cost of an accident is c_a and the cost of the safety equipment is c_s. The agent has a prior over the states, denoted μ, and suppose that $\mu(m|A)q_m c_a > c_s$, so that the safety equipment is efficient.

Suppose the worker originally chose the hazardous industry and learns $\{s, m\}$. Then the agent's beliefs on $\{s, m\}$ are $\mu_A(s) = (1-\delta)\mu(s|A) + \delta$ and $\mu_A(m) = (1-\delta)\mu(m|A)$. Hence the agent purchases safety equipment if and only if $(1-\delta)\mu(m|A)q_m c_a \geq c_s$, or equivalently, $\delta \leq 1 - \frac{c_s}{\mu(m|A)q_m c_a}$. Behaviorally, this means that if agent is *not too* sensitive to dissonance (or only engages in a small amount of wishful thinking) he will make the *correct* decision, otherwise he forgoes the purchase of safety equipment.

However, if the true state were h, then the agent would correctly evaluate the risk as q_h and purchase safety equipment, regardless of his sensitivity to dissonance. That is, the agent's behavior only deviates from rationality when there is a *plausible* alternative state, s, that can justify past choice.

1.10 Conclusion

In this paper I developed, axiomatically, a model of an agent who is systemically biased in favor of a past choice. He behaves as if his subjective beliefs consistently increase the likelihood of states in which f yields better outcomes. This model connects models of non-Bayesian updating with cognitive dissonance and shows how data about previous actions may be necessary to understand and predict future behavior. This paper also serves as a first step toward developing a fully dynamic model of cognitive dissonance, which would allow us to understand the implications of cognitive dissonance in more complex environments, such as financial markets.

There are many possible extensions for this paper, two of which are outlined below. First, one could consider an extended version of this model in which sophistication is explicitly assumed. This would clarify how the anticipation of dissonance affects initial choice, which could sharpen our ability to identify dissonance sensitivity in choice data. Lastly, the model here can be seen as a specific type of reference-dependent updating. It may be fruitful to consider other types of reference-dependent behavior and consider how they might impact updating.

Chapter 2

Sticky Beliefs: A Characterization of Conservative Updating

2.1 Introduction

There has been much empirical and experimental work demonstrating biases in belief updating (see see Camerer [11], Kahneman and Tversky [36], El-Gamal and Grether [14]). In particular, many papers using models of non-Bayesian updating specifically model beliefs as conservative, in that they only partially incorporate the new information. That is, the agent puts too much weight on his prior beliefs. For example, Palfrey and Wang [53] consider agents that *subjectively update* and may under (or over) weight the informativeness of signals, referring to those that under-react as *skeptical* types.

Mobius et al. [48] find in a laboratory experiment that when interpreting information about one's own abilities agents tend to exhibit two regular biases—*asymmetric* updating bias and *conservative* updating bias. The asymmetric updating bias occurs when agents overweight positive signals relative to negative signals. Kovach [41] characterizes belief updating when an agent experiences cognitive dissonance and shows that such an agent's beliefs exhibit asymmetric updating. While it is possible to incorporate conservative bias into the model of cognitive dissonance, this paper studies conservative bias in isolation.

In notation, if prior beliefs are given by a probability distribution P, then the agent's posterior beliefs are

$$Q_A = \delta P + (1 - \delta) BU(P; A) \tag{2.1}$$

for some $\delta \in [0, 1]$, where $BU(P, A)$ denotes the Bayesian update of P given A.[1] In this case beliefs are sticky in the sense that the agent is reluctant to move away from his initial beliefs. The parameter δ can be interpreted as a measure of the agent's conservatism or skepticism about information, or rather $(1 - \delta)$ is a measure of his confidence in the new information. Using a framework of preferences over acts (Savage [60], Anscombe and Aumann [4]), this paper provide preference axioms for an agent who displays conservative updating.

In the context of preferences over menus, an axiomatization of non-Bayesian updating was provided in a three period model by Epstein [16] and extended to an infinite horizon model by Epstein et al. [21]. Further, each paper studies the specific instance of prior-bias as defined in this paper, referred to as *positive prior-bias*. Both [16, 21] utilize a setup of preferences over menus as in [32]. Implicit in this model setup is the assumption that the agent is aware of his non-Bayesian updating. This paper does not require preferences over menus nor make any assumptions about the agent's level of sophistication. Instead this paper assumes as a primitive preferences over acts conditional on the agent's information. That, I assume the agent's information is observable to the analyst and both the agent's preferences before and after information are known.

[1] $BU(P; A)(B) = \frac{P(B \cap A)}{P(A)}$

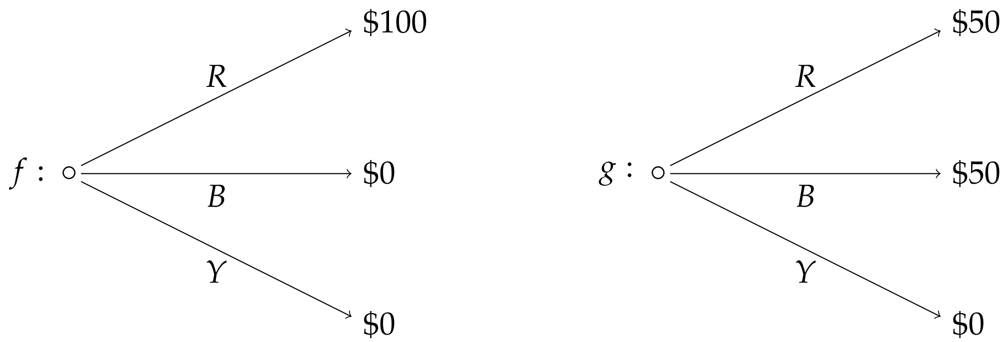

Figure 2.1: Two acts,f and g.

The characterization is achieved by weakening both *dynamic consistency* and *consequentialism*. The weak version of dynamic consistency, which I term dynamic dominance, has similarity to the dominance axiom of Saito [59]. The axiom states that for any two acts f and g and information A, if the agent prefers f to g *ex-ante* and a Bayesian analyst would prefer f to g conditional on A, then certainly the agent prefers f to g *ex-post* (after observing A).

Consider an urn with colored balls, red, blue, and yellow. An experimenter will draw a ball from the urn and the agent's payoff depends on the realization of the draw. Suppose the agent has the choice between the following two acts, f and g:

and the agent prefers g to f. Consider now an alternative experiment, the experiment may also give the agent information about the draw. Suppose the experimenter tells the agent that after drawing the ball, he will observe the draw and notify the agent if the draw is yellow or not. Dynamic consistency asserts that if g is preferred to f, then *conditional* on observing A (being told *not yellow*), the agent prefers $\hat{g}$ to $\hat{f}$.

However, if an agent is uncertain about the quality of information, or is concerned about being deceived or tricked, then concern for unrealized alternatives seems quite plausible. Consequently, the agent reporting a preference of $\hat{f}$ over $\hat{g}$ is reasonable. While consequentialism has a strong normative appeal, the ap-

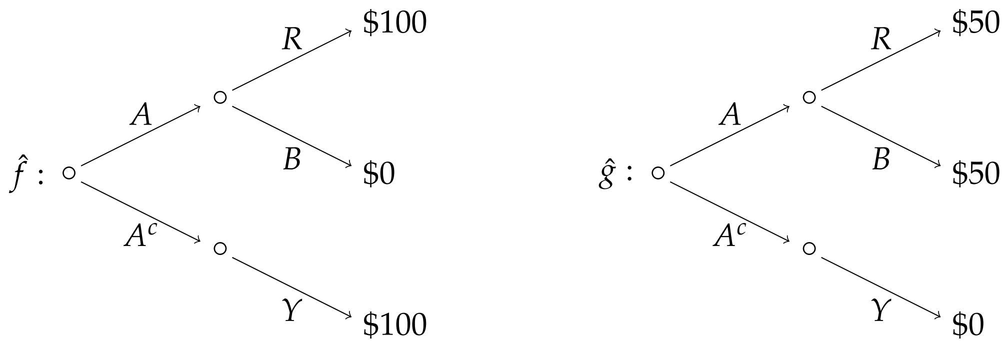

Figure 2.2: Two acts, $\hat{f}$ and $\hat{g}$, incorporating information.

peal is due to the assumption that the agent perceives the information correctly and believes it to be perfectly reliable. Without this assumption the violation of consequentialism above is sensible. The axiom of weak consequentialism that I impose then does not rule out the agent's concern for unrealized alternatives, but only imposes consistency of his evaluation of unrealized alternatives between different observations.

Another interpretation of violations of consequentialism is that while the information may seem objective and precise to an observer or econometrician, the agent subjectively evaluates the information quality and views it as noisy. Under this view we can think of an alternative situation where we believe the agent is Bayesian but with a subjective evaluation of the accuracy of information, where δ measures the agent's perception of accuracy. Under this interpretation an δ of 1 corresponds to the agent believing the information is pure nonsense. In line with this interpretation, Hilbert [34] has proposed that imperfect information processing attributes to conservatism bias, so that objective evidence (observations) appears noisy upon recollection.

2.2 Model

2.2.1 Setup

There is a (finite) set Ω of states of the world, an algebra Σ[2] of subsets of Ω, and a set of consequences, X. Let $\mathscr{F}$ denote the set of finite-valued Σ-measurable functions $f : \Omega \to X$. Each function is referred to as a simple act. Following a standard abuse of notation, for any $x \in X$, I mean by $x \in \mathscr{F}$ the constant act that returns x in every state. Lastly, for any $f, g \in \mathscr{F}$ and for any $A \in \Sigma$, let fAg denote the act that returns $f(\omega)$ when $\omega \in A$ and returns $g(\omega)$ when $\omega \in A^c \equiv \Omega \backslash A$.

Following the literature, I assume that X is a convex subset of a vector space[3]. Thus, mixed acts can be defined point-wise, so that for every $f, g \in \mathscr{F}$ and $\lambda \in [0,1]$, by $\lambda f + (1-\lambda)g$ I mean the act that returns $\lambda f(\omega) + (1-\lambda)g(\omega)$ for each $\omega \in \Omega$.

I assume that the agent has preferences over $\mathscr{F}$ conditional on the agent's information. That is, the agent has a collection of preference relations, $\{\succsim_A\}_{A\in\Sigma}$ over the acts in $\mathscr{F}$, where $\succsim_A$ are the agent's preferences after observing A. Let $\succ_A$ and $\sim_A$ represent the asymmetric and symmetric parts of $\succsim_A$. The case when the agent has no information is represented by $\succsim_\Omega$, or simply $\succsim$.

For a given (S, Σ), a probability charge is a finitely additive set-function $\mu : \Sigma \to [0,1]$. When Σ is a σ-algebra and μ is σ-additive, then μ is a probability measure. When S is finite the set of probability charges and probability measures coincide. Say that a probability charge is convex ranged if for any $A \in \Sigma$ and every $\alpha \in [0, \mu(A)]$, there exists some $B \subset A$ such that $\mu(B) = \alpha$. For any probability charge μ and event $A \in \Sigma$, define the Bayesian update of μ given A by $BU(\mu, A)(B) = \frac{\mu(B\cap A)}{\mu(A)}$ for $B \in \Sigma$.

[2] This need not be a sigma algebra, since I do not require countable additivity to hold.
[3] X may be an interval of monetary prizes or a set of lotteries of some set of prizes

2.2.2 Axioms

The first axiom is a collection of standard conditions. Collectively they are necessary and sufficient for an expected utility representation after the agent observes A, a result which is well-established in the literature.

Axiom 2.1 (Conditional Expected Utility)**.** For each $A \in \Sigma$, the following hold:

Weak order For all $f, g, h \in \mathscr{F}$: (i) either $f \succsim_A g$ or $g \succsim_A f$ and (ii) if $f \succsim_A g$ and $g \succsim_A h$, then $f \succsim_A h$.

Independence: For all $f, g, h \in \mathscr{F}$, and $\lambda \in (0,1]$, $f \succsim_A g \Leftrightarrow \lambda f + (1-\lambda)h \succsim_A \lambda g + (1-\lambda)h$.

Continuity: For all $f, g, h \in \mathscr{F}$, if $f \succ_A g$ and $g \succ_A h$, then there exist weights $\lambda, \gamma \in (0,1)$ such that $\lambda f + (1-\lambda)h \succ_A g$ and $g \succ_A \gamma f + (1-\gamma)h$

Monotonicity: If $f, g \in \mathscr{F}$ and $f(\omega) \succsim g(\omega)$ for all $\omega \in \Omega$, then $f \succsim_A g$.

Nondegeneracy: There are $x, y \in X$ such that $x \succ y$.

Notice that non-degeneracy is only required of the unconditional preference relation. Before introducing the new axioms, for comparison I state classic axioms of dynamic consistency and consequentialism.

Axiom 2.2 (Dynamic Consistency)**.** For any $A \in \Sigma$ and for all $f, g \in \mathscr{F}$

$$fAg \succsim g \Longleftrightarrow f \succsim_A g$$

Dynamic consistency states that if f is preferred to g when they are identical outside of A, then after learning A, f is preferred to g.

Axiom 2.3 (Consequentialism)**.** For any $A \in \Sigma$ and for all $f, g \in \mathscr{F}$,

$$f(\omega) = g(\omega) \text{ for all } \omega \in A \implies f \sim_A g$$

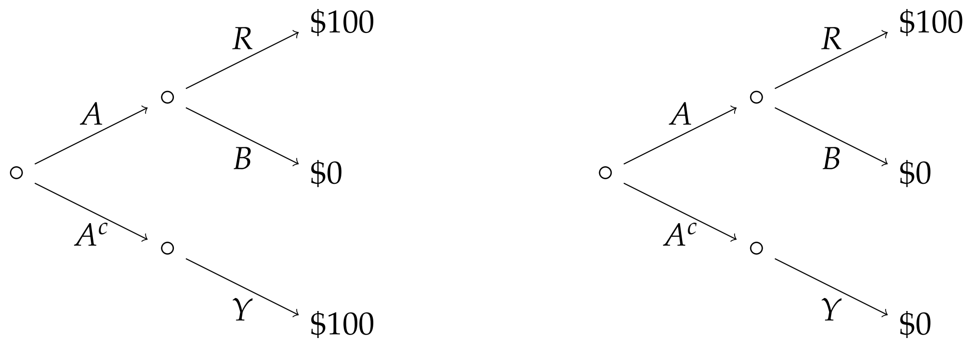

Figure 2.3: Dynamic extension of f and g

Consequentialism states that whenever two acts are identical within A, then after A the agent is indifferent. To gain some intuition as to why an agent may violate the above axioms, consider the following example.

Example 2.1. Consider an urn of red, blue, and yellow balls, $\Omega = \{R, B, Y\}$. An act is a bet on the color of the ball drawn from the urn. Suppose the experimenter informs the agent that a ball will be drawn and the agent will be informed whether the draw was yellow or not. If not, the agent may change his choice of acts. Consider the following two acts:

	R	B	Y
f	\$100	\$0	\$100
g	\$100	\$0	\$0

Ex-ante the agent strictly prefers f to g. Incorporating the information structure results in the following trees:

Consequentialism requires that after $A = \{R, B\}$, the agent is indifferent between the two acts. However, an agent may still assert a strict preference for f (the left tree) even conditional upon being told they are at A. For example, this could be due to concern with being tricked by the experimenter or concern that the information source is unreliable—the experimenter misperceives the color.

2.2.3 Novel Axioms

The first novel axiom is a weak version of dynamic consistency.

Axiom 2.4 (Dynamic Dominance)**.** For any $A \in \Sigma$ and for all $f, g \in \mathscr{F}$

$$\left.\begin{array}{l}\text{(i) } f \succsim g \\ \text{(ii) } fAg \succsim g\end{array}\right\} \implies f \succsim_A g$$

Further, if both (i) and (ii) are strict, then $f \succ_A g$.

First, while Dynamic Consistency is an if and only if condition, Dynamic Dominance is only an if-then condition. Second, while dynamic consistency requires that if fAg then $f \succsim_A g$, I generally allow for violations except when the agent also prefers f to g in the *ex-ante* preference—before any information. For intuition, consider an agent that is uncertain about the quality of his information. If the agent weakly prefers f to g both (i) in A and also (ii) before any information, then regardless of the quality of information he should prefer f to g, hence $f \succsim_A g$ holds. This axiom is satisfied by example 1. The next axiom is the relevant weakening of consequentialism, which regulates preference across information sets.

Axiom 2.5 (Weak Consequentialism)**.** For any $A, B, C \in \Sigma$ with $C \cap (A \cup B) = \varnothing$ and for all $f, g, h \in \mathscr{F}$,

$$fCh \sim_A gCh \Leftrightarrow fCh \sim_B gCh$$

To see how this is a weak form of consequentialism, suppose $C \cap (A \cup B) = \varnothing$ and consequentialism holds. Then for any pair f, g, consider acts of the form fCh and gCh. Then since for all $s \in A \cup B$, $fCh(s) = gCh(s)$, it follows that both $fCh \sim_A gCh$ and $fCh \sim_B gCh$. Thus while consequentialism imposes that the acts are always indifferent, weak consequentialism only states that if they are indifferent after A, then they are also indifferent after B, while allowing for the possibility of a strict ranking.

2.3 Main Results

In this section I state the main results of the paper.

Theorem 2.1 (Representation)**.** *The following are equivalent*

(i) $\{\succsim_A\}_{A\in\Sigma}$ *satisfy Conditional Expected Utility, Dynamic Dominance, and Weak Continuity;*

(ii) There is a non-constant utility function $u : X \to \mathbb{R}$*, a probability charge* μ*, and a unique* $\delta \in [0,1]$ *such that:*

$$f \succsim_A g \Longleftrightarrow \int_\Omega u(f(\omega))\mu_A(d\omega) \geq \int_\Omega u(f(\omega))\mu_A(d\omega)$$

and

$$\mu_A(B) = \delta\mu(B) + (1-\delta)BU(\mu, A)(B)$$

Theorem 1 shows the equivalence of the axioms with the general phenomenon of prior-bias. The general result holds regardless of the specific assumptions on the state space or the specific properties of the probability distribution, i.e., convex range or countable additivity (see Ghirardato [25], Kopylov [40] for additional axioms). Theorem 2 shows the uniqueness properties. The uniqueness of u and μ is standard and uniqueness of δ comes from the uniqueness of μ

Theorem 2.2 (Uniqueness)**.** *If* (u, μ, δ) *and* (u', μ', δ') *represent* $\{\succsim_A\}_{A\in\Sigma}$*, then*

(i) u' *is a positive affine transformation of* u*.*

(ii) $\mu' = \mu$*.*

(iii) $\delta' = \delta$.

Next, I introduce one final axiom. This axiom is a strengthening of the monotonicity condition, such that whenever f is point-wise better than g for each $\omega \in A$, then f is preferred to g after A.

Axiom 2.6 (A-Monotonicity)**.** For any $A \in \Sigma$ and for all $f, g \in \mathscr{F}$,

$$f(\omega) \succsim_A g(\omega) \text{ for all } \omega \in A \implies f \succsim_A g$$

It turns out that this strengthening of monotonicity, in the presence of the other axioms, is equivalent to both dynamic consistency and consequentialism.

Theorem 2.3. *Assume Axiom 2.1, Axiom 2.4, Axiom A.4. The following are equivalent*

(i) $\{\succsim_A\}_{A\in\Sigma}$ *satisfy Dynamic Consistency.*

(ii) $\{\succsim_A\}_{A\in\Sigma}$ *satisfy Consequentialism.*

(iii) $\{\succsim_A\}_{A\in\Sigma}$ *satisfy A-Monotonicity.*

(iv) $\delta = 0$

While generally dynamic consistency and consequentialism do not imply one another, in the presence of Axiom 2.4, Axiom A.4 they are equivalent. Theorem 3 may be understood as follows: if we require that the agent maintain constant relative probabilities between states within A after being told A and that the agent be minimally responsive to information, then consequentialism and dynamic consistency are equivalent and imply Bayesian beliefs. A-Monotonicity is equivalent in this context because it has a flavor on consequentialism. That is, it strengthens monotonicity in a way that implicitly requires that only states within A are relevant to the agent. Of course, the requirement that only states in A are relevant is precisely the spirit of consequentialism.

2.4 Conclusion

I have provided a preference-based characterization of sticky beliefs using preferences conditional on an information set as a primitive. The main innovation is a dual weakening of dynamic consistency and consequentialism and to that, when information is represented as an event tree, imposing the full version of one implies the other. In this way the paper illustrates a connection between the two conditions that has not previously been discussed. This representation can capture certain results in the experimental literature. Further, can potentially be combined with other models, such as the one from chapter 1, to jointly capture under-reaction and directional incorporation.

Chapter 3

Partial Bayesian Updating Under Ambiguity

3.1 Introduction

Suppose Mary is trying to save for her retirement. However, she is concerned that she may not know all the risks involved in the economy; hence she decides to consult a panel of economists. Each economist gives her a different picture of the economy, which is represented by a probability distribution over the possible states of the world, or a prior. Mary's world is in fact quite simple; the economy can grow, shrink, or remain constant. Therefore, the states of the world are $\{G, S, C\}$, and each economist has provided Mary with a probability distribution over these three states.[1] For simplicity, assume Mary solicits advice from three economists. Her set of priors is $\{\pi_1, \pi_2, \pi_3\}$, where $\pi_1 = (\frac{7}{12}, \frac{1}{12}, \frac{1}{3}), \pi_2 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $\pi_3 = (\frac{1}{12}, \frac{7}{12}, \frac{1}{3})$. Suppose Mary learns that there has been an increase in unemployment[2]. How should Mary incorporate this new information into her beliefs? Mary's problem is the topic of this paper.

[1] Ambiguity models explicitly assume the decision maker (DM) does not reduce all the priors into a single prior. Otherwise the DM would be indistinguishable from a standard Bayesian.

[2] Assume that the unemployment rate can fluctuate regardless of the state of the aggregate economy, but that increases in unemployment are much more likely if the economy is shrinking than if the economy is growing.

Since Ellsberg's seminal paper [15] on the distinction between ambiguity and risk,[3] numerous models of ambiguity sensitive agents have been proposed. The earliest and most well known are Schmeidler's Choqeut expected utility (CEU) model [61] and Gilboa and Schmeidler's "maxmin expected utility with non-unique prior" [29]. Both of these models are linked in the sense that agent's beliefs cannot be expressed as a single probability. In the first model beliefs are represented by a non-additive probability, or capacity; in the second model beliefs are represented by a convex set of probabilities, or multiple priors. This paper focuses on modeling beliefs of the second form.

In order for models of ambiguous beliefs to be useful in many areas of economics however, these models must be extended to an inter-temporal framework. Previous work on updating, when beliefs are represented by a set of priors, has focused on two procedures. For the first procedure, known as generalized Bayesian updating (GBU) or full Bayesian updating [28], the agent applies Bayes' rule to each prior. The second procedure, known as maximum likelihood updating (MLU) [28], recommends that the agent only retain priors that assigned the greatest probability to the observed event. In the case of Mary, if she practices GBU she believes that π_1 is just as good as π_3, even after observing the increase in unemployment. If she practices MLU, then she only retains π_3.

Both methods of updating beliefs are unsatisfactory. The problem with generalized Bayesian updating is that the DM treats all priors as equally good and is not able to use the new information to make inference about which priors to believe. Maximum likelihood updating does not suffer from this problem, but it seems to throw out too many priors. Returning to Mary's problem, under GBU she treats experts who were good predictors the same as those who were very bad, while under maximum likelihood updating she only believes those experts that gave the highest

[3]A distinction also noted by Knight [39] and Keynes [38].

probability to the outcome observed. That is, π_1 is clearly less likely to be true than π_3, but it is not clear that π_2 should be rejected.

My paper proposes an alternative updating procedure that is simultaneously more general than and a compromise between the above two procedures. I provide axioms on preferences that are equivalent to an agent engaging in my proposed updating procedure. Returning to my initial example, if Mary satisfies my axioms, then there is a unique parameter $\alpha \in [0,1]$ that describes her *willingness to infer*.[4] Mary uses this parameter to determine a threshold value for each event. After she observes a given event, she retains only those priors that perform well relative to this threshold value and then applies Bayes' rule to the retained priors. This procedure is a generalization of the other two, since a parameter value of zero or one corresponds to the agent performing generalized Bayesian updating or maximum likelihood updating, respectively. It can also be viewed as a compromise between the two, since Mary is only retaining "good" priors, but she is concerned about rejecting too many. That is, I allow for Mary to retain π_2 and π_3.

3.1.1 Dynamic Consistency and Conditional Preferences

My model is not truly inter-temporal, as I do not model time. Instead I model an agent with a collection of preference relations, where each relation is conditional on some event A. When the only information the agent has is the entire state space (i.e., the agent has no information), I refer to this relation as the unconditional preference relation. The interpretation is that after observing some event A, the agent updates her preferences from $\succsim = \succsim_\Omega$ to $\succsim_A$. Imposing conditions on how preferences can change after receiving information allows us to understand how the agent updates her beliefs. This framework is not purely for convenience, as it

[4] By this I mean the agent's willingness to discard priors, or equivalently the magnitude of her threshold.

allows for the conditions on preferences that generate a specific type of updating to be clearly understood. The *functional representation* of the behavior modeled in this paper can be adapted to a dynamic framework.

A concept closely linked to updating beliefs is dynamic consistency (DC). DC links unconditional and conditional preferences by requiring that an act f is preferred to g with the unconditional relation and f and g are the same outside of some event A, if and only if f is preferred to g conditional on A occurring. A famous result of Epstein and Breton [17] is that preferences are dynamically consistent if and only if the agent behaves as a Bayesian with beliefs represented by a single probability. That is, the agent satisfies the axioms of Savage [60].

Sacrificing dynamic consistency is something many economists are not willing to do. Indeed, DC has a very strong normative appeal; so strong that many consider it a property of rationality. However, it is not so clear that it is necessary for rational behavior. The willingness to revise one's beliefs, especially if they were formed without compelling evidence or copious data, in the face of new information seems much more in line with rational belief evolution. DC only allows for *revision* after zero-probability events, which is actually quite restrictive. Ortoleva [52] axiomatizes a model where a decision maker performs Bayesian updating if, according to prior beliefs, the observed event has a probability above some threshold. Otherwise the agent revises beliefs by updating a *prior over priors* and choosing a new prior according to a maximum likelihood rule.

Ghirardato et al. [27](GMM) impose DC on their derived relation of *unambiguous preference*. In their notation, f is *unambiguously preferred* to g ($f \succsim^* g$) if and only if for any act h and any $\lambda \in [0,1]$, $\lambda f + (1-\lambda)h \succsim \lambda g + (1-\lambda)h$. That is, f and g are such that hedging is not useful because f is *always* better than g. They find that dynamic consistency of $\succsim^*$ is equivalent to GBU. This derived relation is also essential to the analysis in this paper.

I impose a weakened version of DC on $\succsim^*$ because there are simple scenarios where it makes sense to violate DC (in a regulated fashion). First note that $\succsim^*$ admits a representation à la Bewley, where

$$f \succsim^* g \text{ if and only if } \int_\Omega u(f(\omega))d\pi(\omega) \geq \int_\Omega u(g(\omega))d\pi(\omega) \text{ for all } \pi \in \mathscr{C} \quad (3.1)$$

Suppose there are two acts that are not comparable with $\succsim^*$, but act f is *usually* better than g (in A) [5], and the agent is told that event A has occurred. An agent uses the information that A occurred in two ways. First, the agent determines which priors are *likely* to be true, given that A happened, and discards the rest as too implausible[6]. Second, the agent updates her beliefs using Bayes' rule. Since her posterior set is strictly smaller than the set obtained throughout GBU with $\mathscr{C}$, the agent might find that f and g are now comparable, with $f \succsim^*_A g$. This cannot occur when DC is imposed, but that does not mean that this is unreasonable behavior; I find it more reasonable than trying to satisfy DC in this context.

3.2 Related Literature

How people should update their beliefs in the presence of ambiguity has been studied for some time, though no clear solution has been discovered. Gilboa and Schmeidler [30] axiomatized MLU, which is shown to be equivalent to Dempster-Shafer updating when the set of priors is determined by a convex capacity. Generalized Bayesian updating, sometimes called full Bayesian updating, was first proposed by Jaffray [35], and was axiomatized by Pires [54]. It was later shown by Ghirardato et al. [27] that GBU is equivalent to imposing dynamic consistency only on a derived relation, $\succsim^*$, called the agent's *unambiguous preference*.

[5] By this I mean $\mathbb{E}_\pi[u(g)] > \mathbb{E}_\pi[u(f)]$ only when $\pi(A)$ is *small*.

[6] That is, the agent perceives less ambiguity by ignoring priors that assigned low probability to A.

The relationship between GBU, MLU and dynamic consistency has been studied by [19]. They propose an axiom that imposes a condition of rectangularity on the set $\mathscr{C}$ of priors. Such set of priors can be formed *as if* the agent constructs successively larger sets of priors via backward induction along an event tree. Under rectangularity, GBU and MLU are equivalent, and they show that rectangularity allows for behavior to remain dynamically consistent. While dynamic consistency may be normatively desirable, the condition of rectangularity seems unnatural in many environments for two reasons. First, while it is motivated by appealing to backwards induction there is experimental evidence suggesting that people are not very good at backward induction [24]. Second, in experimental settings when the decision maker is given an objective set of priors that does not conform to rectangularity, it is sensible to assume the agent uses the set of priors rather than transform them to satisfy rectangularity.

Epstein and Schneider [20] utilize a similar functional representation to study long-run learning under ambiguity and apply the model to dynamic portfolio choice. However, they do no introduce axioms that identify α as their focus is on studying when ambiguity is resolved over time. In contrast, this paper instead focuses on what preferences are consistent with the updating procedure under consideration and how α can be identified from preference data.

My paper takes a complementary approach to Epstein and Schneider, while building on the *unambiguous preference* studied in GMM. Rather than impose full dynamic consistency, I impose a weakened form only on $\succsim^*$. Thus I allow for a more natural approach to updating. I say that the approach is more natural in the sense that I believe it is closer to how a reasonable person would behave. Returning to my earlier example, Mary choosing to only retain π_2 and π_3 is a rational decision.

3.3 Preliminaries and Notation

There is a finite set Ω of states of the world[7], an algebra Σ of subsets of Ω, and a set of consequences, X. Let $\mathscr{F}$ denote the set of finite-valued Σ-measurable functions $f : \Omega \to X$. Each function is referred to as a simple act. Following a standard abuse of notation, for any $x \in X$, I mean by $x \in \mathscr{F}$ the constant act that returns x in every state. Lastly, for any $f, g \in \mathscr{F}$ and for any $A \in \Sigma$, let fAg denote the act that returns $f(\omega)$ when $\omega \in A$ and returns $g(\omega)$ when $\omega \in A^c \equiv \Omega \backslash A$.

Following the literature, I assume that X is a convex subset of a vector space. Thus, mixed acts can be defined point-wise, so that for every $f, g \in \mathscr{F}$ and $\lambda \in [0,1]$, by $\lambda f + (1-\lambda)g$ I mean the act that returns $\lambda f(\omega) + (1-\lambda)g(\omega)$ for each $\omega \in \Omega$.

I assume that the agent has preferences over $\mathscr{F}$ conditional on the agent's information. That is, the agent has a collection of preference relations, $\{\succsim_A\}_{A\in\Sigma}$ over the acts in $\mathscr{F}$. For each $A \in \Sigma$, $\succ_A$ and $\sim_A$ represent the asymmetric and symmetric parts of $\succsim_A$. The case when the agent has no information is represented by $\succsim_\Omega$, or simply $\succsim$.

Let $BU(\mathscr{C}, A)$ denote the set of prior-by-prior Bayesian updates conditional on A and let $BU(\pi, A)$ denote the Bayesian update of π conditional on A.

3.3.1 The *Unambiguously Preferred* Relation

The derived *unambiguously preferred* relation is essential to the analysis in this paper. The following definition is due to Ghirardato, Maccheroni, and Marinacci [26]:

Definition 3.1. Say that f is *unambiguously preferred* to g, denoted $f \succsim^* g$, if and only if for any act h and any $\lambda \in [0,1]$, $\lambda f + (1-\lambda)h \succsim \lambda g + (1-\lambda)h$.

[7]Finiteness is assumed for notational convenience. All results are unchanged if the axioms are applied to non-null events.

It is clear that $\succsim^*$ is complete only when the agent admits a subjective expected utility representation. Let $\bowtie^*$ represent the *incomplete* part of $\succsim^*$. The $\succsim^*$ relation admits a representation à la Bewley, as shown in Equation 3.1, for some closed, convex set $\mathscr{C}$. The cases where $\succsim$ satisfies independence, $\succsim^* = \succsim$, and $\mathscr{C}$ is a singleton set are all equivalent.

3.4 Model

This section presents the basic model of preferences used in the paper.

Axiom 3.1 (Conditional MMEU)**.** For each $A \in \Sigma$, the preference relation $\succsim_A$ satisfies the Gilboa-Schmeidler axioms:

Weak order: For all $f, g, h \in \mathscr{F}$: (i) either $f \succsim_A g$ or $g \succsim_A f$ and (ii) if $f \succsim_A g$ and $g \succsim_A h$, then $f \succsim_A h$.

Certainty independence: For all $f, g \in \mathscr{F}, x \in X$, and $\lambda \in (0, 1]$, $f \succsim_A g \Leftrightarrow \lambda f + (1-\lambda)x \succsim_A \lambda g + (1-\lambda)x$.

Continuity: For all $f, g, h \in \mathscr{F}$, if $f \succ_A g$ and $g \succ_A h$, then there exist weights $\lambda, \gamma \in (0, 1)$ such that $\lambda f + (1-\lambda)h \succ_A g$ and $g \succ_A \gamma f + (1-\gamma)h$

Strict Monotonicity: If $f, g \in \mathscr{F}$ and $f(\omega) \succsim_A g(\omega)$ for all $\omega \in \Omega$, then $f \succsim_A g$. If in addition there is some $\omega \in A$ such that $f(\omega) \succ_A g(\omega)$, then $f \succ_A g$.

Ambiguity Aversion: For all $f, g \in \mathscr{F}$, if $f \sim_A g$ then for all $\lambda \in [0, 1]$, $\lambda f + (1-\lambda)g \succsim_A f$.

Non-degeneracy: There are $f, g \in \mathscr{F}$ such that $f \succ_A g$.

Ordinal Preference Consistency: For all $x, y \in \mathscr{F}$, $x \succsim y$ if and only if $x \succsim_A y$.

Axiom 3.1 is comprised of standard conditions known to be equivalent to the MMEU representation ([29]), plus ordinal preference consistency (OPC). OPC is the requirement that tastes remain unchanged after information, requiring that preference changes are due to the agent's response to information. The conditions are applied to each preference relation in the collection $\{\succsim_A\}$ so that *ex-ante* and *ex-post* preferences both have the same structure ([33]).

Axiom 3.2 (Consequentialism)**.** For each $A \in \Sigma$ and for all $f, g \in \mathscr{F}$,

$$\text{if } f(\omega) = g(\omega) \text{ for all } \omega \in A, \text{ then } f \sim_A g.$$

Axiom 3.2 is a classic condition stating that preferences conditional on A only depend on how acts perform within A.

Axiom 3.3 (Weak Unambiguous Dynamic Consistency)**.** For each $A \in \Sigma$ and for all $f, g \in \mathscr{F}$,

$$\text{if } fAg \succsim^* g, \text{ then } f \succsim_A^* g.$$

Axiom 3.3 is a weakening of the Unambiguous Dynamic Consistency condition from Ghirardato et al. [27]. To motivate Axiom 3.3, consider the following example.

Example 3.1. There are three states of the world, $\Omega = \{R, B, Y\}$, and the agent believes that the true distribution over the states of the the world belongs to $\mathscr{C} = \{(\frac{1}{3}, \beta, \frac{2}{3} - \beta) | \beta \in [\frac{1}{12}, \frac{7}{12}]\}$. Consider the following acts,

	R	B	Y
f	\$90	\$100	\$50
g	\$100	70\$	\$50

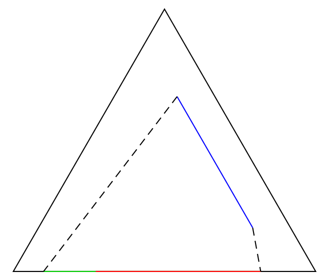

Figure 3.1: Sets of beliefs in the simplex.

Given the set $\mathscr{C}$, the agent prefers g to f because there is a lot of uncertainty about the true probability of sate B. However, there are distributions such that the agent prefers f to g, hence $f \bowtie^* g$. Suppose the agent learns that Y was not realized. That is, the state is in $A = \{R, B\}$. Since $fAg = f$, $fAg \bowtie^* g$. The agent conceives two explanations for A's occurrence. One is that that Y occurs with high probability but A occurred as a low probability event. The other is that A is a high probability event. Given that A did occur, it is reasonable to believe that *it is more likely that* A was a high probability event than not, hence the true probability of B is in fact not *that small*. Suppose the agent then uses this information (that A occurred) to adjust her *ex ante* beliefs to $\mathscr{C}' = \{((\frac{1}{3}, \beta, \frac{2}{3} - \beta)|\beta \in [\frac{3}{24}, \frac{7}{12}]\}$. After making this inference, she proceeds to apply Bayes' rule to all priors, yielding $\mathscr{C}'_A = \{\frac{1/3}{1/3+\beta}, \frac{\beta}{1/3+\beta}|\beta \in [\frac{3}{24}, \frac{7}{12}]\}$. Now, it can be easily verified that for all $\pi \in \mathscr{C}'_A$, $\mathbb{E}_\pi(f) \geq \mathbb{E}_\pi(g)$. But this is equivalent to $f \succsim^*_A g$. Hence the agent violates dynamic consistency of $\succsim^*$, but she does so because it is the most reasonable way to behave *ex post*. In the following figure, $\mathscr{C}$ is given by the blue line, $\mathscr{C}'_A$ is given by the red line, and $BU(\mathscr{C}, A)$ is given by the the union of the green and red lines.

Definition 3.2. Say that an agent performs α-Bayesian Updating (α-BU) if there exists an $\alpha \in [0, 1]$ so that for each $A \in \Sigma$, $A \neq S$,

$$\mathscr{C}_A = \{BU(\pi, A)|\pi \in \mathscr{C}, \pi(A) \geq \alpha \max_{\pi' \in \mathscr{C}} \pi'(A)\}. \tag{3.2}$$

It is clear that in the case $\alpha = 0$, the agent applies Bayes' rule to every prior in $\mathscr{C}$, and in the case $\alpha = 1$, the agent updates only those priors that gave the greatest likelihood to A. Thus, this representation encompasses both generalized Bayesian updating and maximum likelihood updating as special cases. We can think of this agent as one who simultaneously engages in Bayesian learning and makes inferences about which priors are most likely to be true. So while this model is more general, it perhaps more naturally describes the type of thinking a sophisticated agent performs. We can interpret the α parameter as the agent's *willingness to infer*.

3.4.1 Objective Randomizations

While Axiom 3.3 ensures that the agent's posterior beliefs are a subset of the GBU set of beliefs, it imposes no restrictions on which priors are updated or regularity across events. In order to characterize the right restriction I utilize the notion of an objective randomization. Suppose the agent were offered the opportunity to apply an objective randomization to an act. That is, the agent is given the opportunity to reduce an act to a lottery with specified odds.

Definition 3.3. Let $\rho \in \Delta(\Omega)$ denote an objective randomization. Then for each f and each $\rho \in \Delta(\Omega)$, we denote by f^ρ a lottery on X that returns $f(\omega)$ with probability $\rho(\omega)$ for each state $\omega \in \Omega$;

$$f^\rho := \left(\sum_{\omega} \rho(\omega) f(\omega) \right) \mathbf{1}_\Omega.$$

A similar notion of reducing subjective uncertainty to objective uncertainty via mixing was use in Ok et al. [51] to study incomplete preferences under uncertainty.

Then for any objective randomization ρ, and any event $A \in \Sigma$, let $\rho(A) = \sum_{\omega \in A} \rho(\omega)$. This is thus the objective probability given to A under the objective randomization. Given any objective randomization and event A, we can then define the derived conditional randomization that only gives weight to states in A.

Definition 3.4. For every $A \in \Sigma$ and $\rho \in \Delta(\Omega)$, let ρ_A denote the objective randomization such that $\rho_A(\omega) = \frac{\rho(\omega)}{\rho(A)}$ for $\omega \in A$ and 0 otherwise.

I also introduce a further restriction on the types of objective randomizations I will consider. Given some $A \in \Sigma$, it will be technically convenient to consider the collection of A-maximal randomization such that the agent always prefers to reduce subjective uncertainty to objective uncertainty. That is, given some distribution over A, we consider the objective randomization that is identical within A and puts maximal likelihood on A.

Definition 3.5. For every $A \in \Sigma$, say that $\rho \in \Delta(\Omega)$ is A-maximal if $f^\rho \succsim f$ for all $f \in \mathscr{F}$ and for any other ρ' such that $\rho_A = \rho'_A$ and $f^{\rho'} \succsim f$ for all $f \in \mathscr{F}$, $\rho(A) \geq \rho'(A)$.

It should be noted that the definition does not preclude the existence of two A-maximal randomizations such that $\rho(A) > \rho'(A)$. This is possible so long as $\rho_A \neq \rho'_A$. Before stating the final axiom, I require one more definition.

Definition 3.6. For every $A \in \Sigma$ for some $x, y \in X$, with $x \succ y$, define $\bar{m}_A \in [0,1]$ by the equation $\bar{m}_A y + (1 - \bar{m}_A)x \sim yAx$.

The value $\bar{m}_A$ captures the agent's most subjective, maximal probability of A. With all these concepts established, I now introduce the fourth and final axiom.

Axiom 3.4 (Dynamic Reduction Consistency)**.** For all $A, B \in \Sigma$, consider any $\rho, \rho' \in \Delta(\Omega)$ such that they are A and B-maximal, respectively. If $\frac{\rho(A)}{\bar{m}_A} \geq \frac{\rho'(B)}{\bar{m}_B}$ and $f^{\rho'_B} \succsim_B f$ for all $f \in \mathscr{F}$, then $f^{\rho_A} \succsim_A f$ for all $f \in \mathscr{F}$.

Axiom 4 states that if an agent prefers to reduce an act to a lottery when given ρ' (ρ'_B), before (after) learning some event B and also prefers to reduce acts to lotteries via ρ, then if ρ puts objectively higher weight on A than ρ' does on B (when normalized by $\bar{m}$) then the agent prefers reduction by ρ_A after A. This axiom provides both within event restrictions on updating behavior (if $A = B$) and between event regularity (when $A \neq B$). I now state the main result of the paper.

Theorem 3.1. *The following are equivalent:*

(i) The collection of preferences $\{\succsim_A\}_{A\in\Sigma}$ satisfy axioms 1-4.

(ii) The agent performs α-Bayesian Updating.

The next theorem characterizes the uniqueness properties of the representation.

Theorem 3.2. *Suppose that $(u, \mathscr{C}, \alpha)$ and $(u', \mathscr{C}', \alpha')$ both represent the same preferences. Then $u = u'$, $\mathscr{C} = \mathscr{C}'$, and*

(i) if $\succsim^*_A$ satisfies dynamic consistency for every $A \in \Sigma$, then

$$\alpha, \alpha' \in [0, \min_{A\in\Sigma}(\frac{\pi(A)}{\max_{\mu\in\mathscr{C}} \mu(A)})].$$

(ii) if $\succsim^*_A$ violates dynamic consistency at some $A \in \Sigma$, then

$$\alpha = \alpha'.$$

To gain a better intuition behind the result, consider the following example.

3.4.2 Example 2

Let $\Omega = \{1, 2, 3, 4\}, A = \{1, 2\}, B = \{3, 4\}, C = \{1, 2, 3\}$ and suppose $\mathscr{C} = \{(\frac{1}{4}, \gamma, \beta, \frac{3}{4} - \gamma - \beta) | \gamma, \beta \geq \frac{1}{24}, \gamma + \beta \leq \frac{17}{24}\}$. Then $\bar{m}_A = \frac{11}{12}, \bar{m}_B = \frac{17}{24}$, and $\bar{m}_C = \frac{23}{24}$. Let $\alpha = \frac{2}{5}$. The GBU sets of posterior beliefs are:

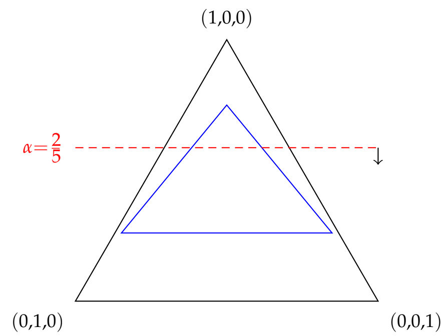

Figure 3.2: The set of posterior beliefs, $\mathscr{C}_C^{\alpha}$, as α varies.

(i) $BU(\mathscr{C}, A) = co\left\{(\frac{6}{7}, \frac{1}{7}, 0, 0), (\frac{3}{11}, \frac{8}{11}, 0, 0)\right\}$,

(ii) $BU(\mathscr{C}, B) = co\left\{(0, 0\frac{1}{17}, \frac{16}{17}), (0, 0, \frac{16}{17}, \frac{1}{17})\right\}$,

(iii) $BU(\mathscr{C}, C) = co\left\{(\frac{3}{4}, \frac{1}{8}, \frac{1}{8}, 0), (\frac{6}{23}, \frac{16}{23}, \frac{1}{23}, 0), (\frac{6}{23}, \frac{1}{23}, \frac{16}{23}, 0)\right\}$.

Under the $\frac{2}{5}$-Bayesian representation, the agents posterior sets are:

(i) $\mathscr{C}_A^{\frac{2}{5}} = co\left\{(\frac{15}{22}, \frac{7}{22}, 0, 0), (\frac{3}{11}, \frac{8}{11}, 0, 0)\right\}$,

(ii) $\mathscr{C}_B^{\frac{2}{5}} = BU(\mathscr{C}, B)$,

(iii) $\mathscr{C}_C^{\frac{2}{5}} = \left\{\left(\frac{12}{46} + \frac{18}{46}\beta, \frac{34}{46} - \frac{18}{46}\beta - \gamma, \gamma\right) : \beta \in [0, 1], \gamma \in \left[\frac{2}{46} + \frac{3}{46}\gamma, \frac{32}{46} - \frac{21}{46}\beta\right]\right\}$.

Figure 3.2 illustrates how her set of posterior beliefs change as α changes. The outer triangle is $\Delta(C)$ and the inner, blue triangle represents $BU(\mathscr{C}, C)$, while the portion below the dashed, red line shows $\mathscr{C}_C^{\alpha=\frac{2}{5}}$. As α increases towards 1, the red line falls to the bottom line of the blue triangle.

It is worth noting that for any $\alpha \in [0, 1], \mathscr{C}_B^{\gamma} = BU(\mathscr{C}, B)$. This has the nice interpretation that, due to the high degree of uncertainty about the states in B, the agent is unwilling to reject any priors since she is unable to make a "reasonable" inference.

3.5 Additional Properties of α-BU

In this section I consider an additional structural property linking an agent's belief sets. I then establish an method for comparing changes in belief sets between individuals and show that inference has a sharp relation to completeness.

3.5.1 Informational Path Independence

An interesting and desirable property in Example 1 is that $\left(\mathscr{C}_C^{\frac{2}{5}}\right)_A^{\frac{2}{5}} = \mathscr{C}_A^{\frac{2}{5}}$. In the example, the posterior belief set is independent of the sequence of information and only depends on the final information; the agent's beliefs are *informationally path independent*. Even though the environment does not consider multiple rounds of information revelation, a notion of *informational path independence* can be formalized as follows:

Definition 3.7 (Informational Path Independence)**.** For any $B \subset A \in \Sigma$,

$$(\mathscr{C}_A^\alpha)_B^\alpha = \mathscr{C}_B^\alpha$$

This property will often not be satisfied by the representation. It clearly holds whenever $\succsim^*$ is dynamically consistent. It also trivially holds when $|\omega| = 3$, since we would arrive at full information revelation. There may be a general result characterizing the compatibility of path independence and α-partial Bayesian updating rules in which $\succsim^*$ violates dynamic consistency for at least some event, but this is left for future work.

3.5.2 Comparative Inference and Completeness

Also, we might want to compare agents' *willingness to infer*. That is, the case when there are two agents such that $\alpha_1 \geq \alpha_2$ is worth studying. This type of comparison is partially confounded by the fact that we could initially have $\mathscr{C}_{2,\Omega} \subset \mathscr{C}_{1,\Omega}$, so we might observe that $\mathscr{C}_{2,A}^{\alpha_2} \subset \mathscr{C}_{1,A}^{\alpha_1}$. However, whenever $\mathscr{C}_{2,\Omega} = \mathscr{C}_{1,\Omega}$, it follows that $\mathscr{C}_{1,A}^{\alpha_1} \subset \mathscr{C}_{2,A}^{\alpha_2}$ if $\alpha_1 \geq \alpha_2$.

An agent that makes *more inferences* (has a larger α) will typically reveal less ambiguity *ex-post* since he eliminates more priors before updating. Thus such an agent should be able to make more unambiguous comparisons. It turns out that this intuition characterizes being *more willing to infer*. First, I introduce a definition of one binary relation being more complete than another.

Definition 3.8. Say that $\succsim^1$ is *more complete* than $\succsim^2$ if for all $f, g \in \mathscr{F}$,

$$f \succsim^2 g \Rightarrow f \succsim^1 g \text{ or } g \succsim^1 f.$$

That is, whenever agent 2 is able to make a comparison, then so is agent 1. The following theorem then shows that, under certain regularity conditions, being more complete characterizes a greater willingness to infer.

Theorem 3.3. *Consider two agents that satisfy axioms* $1-4$ *such that* $\succsim^1 = \succsim^2$ *and for some* $A \in \Sigma$, ${\succsim^1}^*_A$ *and* ${\succsim^1}^*_A$ *violates dynamic consistency. Then the following are equivalent,*

(i) For every $A \in \Sigma$, ${\succsim^1}^*_A$ *is more complete than* ${\succsim^2}^*_A$

(ii) $\alpha_1 \geq \alpha_2$.

This result also suggests an alternative axiomatization for the updating procedure presented that does not make use of reduction to objective lotteries. Since the agent's posterior set is typically smaller than the one obtained via updating all priors according to Bayes' rule, the agents conditional, unambiguous preference, $\succsim^*_A$, will typically be more complete. Thus a condition that regulates *increased completeness* across events A and B could potentially be developed to replace Axiom 3.4.

3.6 Conclusion

This paper axiomatizes a model of updating in multiple priors models that generalizes both GBU and MLU. Upon receiving information, the agent makes an inference about her priors and applies Bayes' rule to the *good* priors. Whenever the agent's *unambiguous preference* violates dynamic consistency, her *willingness to infer* is captured by a single parameter, $\alpha \in [0, 1]$. I also characterize a way to compare willingness to infer between individuals that can be elicited from preference data.

This model is flexible enough to capture a variety of phenomena. Since a prior updating procedure functionally similar to the one proposed in this paper has been utilized by Epstein and Schneider [20], it is know that this model can capture shrinking ambiguity and, in the financial markets, increased stock market participation. This model may also capture apparent overreaction to news. Before information, and agent evaluates an uncertain prospect according to the worst-case prior. After information, however, the agent's posterior set may be a strict subset of the one obtained by applying Bayes' rule to every prior. Hence the agent's posterior evaluation may never be less than and can be strictly larger than the evaluation obtained be keeping every prior. Thus while it may seem that decision makers are overreacting, they may be rationally using the information to distinguish between possible priors.

Finally, the model provides a rationalization for violations of dynamic consistency under ambiguity. In this framework violations of dynamic consistency arise because the agent uses the information to distinguish between priors, rather than treating all priors as equal. Thus type of inferential behavior requires belief revision, which creates dynamic preference reversals that a decision maker may rationally defend.

Appendix A

Appendix to Chapter 1

A.1 Best-Case Binary Distortion: $\delta(A, f)$ Comparative Statics

This section further examines the binary distortion from section 7 and provides comparative statics for the level of dissonance between scenarios. First I introduce two independence type axioms.

Axiom A.1 (Reference C-Independence)**.** For all $fg, h \in \mathscr{F}$, $x \in X$, and $\alpha \in (0, 1]$

$$g \succsim_{A,f} h \Leftrightarrow g \succsim_{A,\alpha f+(1-\alpha)x} h$$

Axiom A.2 (Reference A-Comonotonic Independence)**.** For all $f \asymp_A f'$, and all $g, h \in \mathscr{F}$, and $\alpha \in (0, 1]$

$$g \succsim_{A,f} h \Leftrightarrow g \succsim_{A,\alpha f+(1-\alpha)f'} h$$

The first axiom states that mixing the initially chosen action and a constant action does not change the evaluation of other actions. The second requires this property to hold when mixing between A-comonotonic actions. It turns out that both axioms are equivalent, as is shown in the following theorem.

Theorem A.1. *Suppose* $\{\succsim, \succsim_{A,f}\}$ *satisfy Axiom 1.1, Axiom 3.2, Axiom 1.4, Axiom 1.7 . Then the following are equivalent:*

(i) $\succsim_{A,f}$ *satisfies Axiom A.1*

(ii) $\succsim_{A,f}$ *satisfies Axiom A.2*

(iii) $f \asymp_A f' \implies \delta(A,f) = \delta(A,f')$. *That is,* $\delta(A,f)$ *is constant on* $\asymp_A$ *equivalence classes.*

That is, whenever two scenarios are *ordinally* equivalent, the agent has the same beliefs within each scenario. This eliminates any sensitivity of the distortion magnitude to payoff variation *within* a scenario. However, it may seem natural to think that scenarios in which there is a large difference between the best and worst possible payoffs induce more intense feelings of dissonance. The following definition captures one notion of what it means for one act to have greater payoff variance.

Definition A.1. An act f is *at least as dispersed as* g , denoted $f \ltimes g$, if there is some constant act x and an $\alpha \in [0,1]$ such that $g = \alpha f + (1-\alpha)x$

The next axiom imposes precisely the intuition discussed above, that scenarios with greater payoff variance create more intense feelings of dissonance and hence result in a greater need for justification. That is, scenarios that are more dispersed induce a greater level of dissonance.

Axiom A.3 (Increasing Sensitivity to Dissonance)**.** For all (A,f) and (A,g), if $f \ltimes g$, then for all $h \asymp_A f$,

$$h \succsim_{A,g} x \implies h \succsim_{A,f} x$$

While the binary distortions are not continuous, this is primarily due to the fact that small variation within an act may result in scenario with a dramatically different classification of good or bad states. However, by considering a sequence of strongly A-comonotone acts this problem is eliminated. This leads to the following weak continuity property.

Axiom A.4 (Weak Continuity)**.** For $f_n \to f$ such that $f_n \asymp_A f_m$ for all n, m, and any $g, h \in \mathscr{F}$,

$$\text{if } g \succsim_{A,f^n} h \text{ for all } n, \text{ then } g \succsim_{A,f} h$$

Before stating the next theorem, I require a bit more notation. Let $\max_{\succsim,A}(f)$ denote the preference maximal consequence under f for states in A. The notation is similar for preference minimal consequences.

Theorem A.2. *Suppose the collection of preferences, $\{\succsim, \succsim_{A,f}\}$ satisfy Axiom 1.1, Axiom 3.2, Axiom 1.4, Axiom 1.7, then preferences satisfy Axiom A.3 and Axiom A.4 if and only if there exists a monotone, continuous $\Gamma_{A,f} : X \times X \to [0,1]$, with $\Gamma_{A,f}(x, x) = 0$, such that*

$$\delta(A, f) := \Gamma_{A,f}\left(\max_{\succsim,A}(f), \min_{\succsim,A}(f)\right).$$

The next axiom imposes even more structure between different scenarios. In particular, it regulates the agent's reaction across information and between different actions by imposing a type of consistency between two scenarios having a common non-justifying state.

Axiom A.5 (Consistent Sensitivity to Dissonance)**.** For any (A, f) and (B, g) and $C \subset A \cap B$, if there exists $s \in A, s' \in B$ such that for all $\tilde{s} \in C$, $f(s) \succ f(\tilde{s})$, and $g(s') \succ g(\tilde{s})$, then for all $x, y, z, w \in X$ such that $yBz \sim xAz$,

$$yCz \succsim_{A,f} w \Leftrightarrow xCz \succsim_{B,g} w$$

Theorem A.3. *Suppose the collection of preferences, $\{\succsim, \succsim_{A,f}\}$ satisfy Axiom 1.1, Axiom 3.2, Axiom 1.7, Axiom 1.4, then preferences satisfy Axiom A.5 if and only if* *there exists a real number $\delta \in [0,1]$ such that for all (A, f),*

$$\delta(A, f) = \delta$$

A.2 Preliminary Results

Before proving the main theorems I establish some notation and lemmas. For any $u : X \to \mathbb{R}$, and $f \in \mathscr{F}$, then let $u(f)$ denote the vector in $\mathbb{R}^{|S|}$ given by $u(f)(s) := u(f(s))$. For any $A \subset S$, let $\Delta(A) := \{\mu \in \mathbb{R}^{|A|} | \sum_{s \in A} \mu(s) = 1, \mu(s) \geq 0\}$, the set of probability distributions on A. There is a natural bijection between $\{\mu \in \Delta(S) : \sum_{s \in A} \mu(s) = 1\}$ and $\Delta(A)$.

Say two linear U, V are functionals on $\mathscr{F}$ are normalized by u if for all $x \in \mathscr{F}$, $V(x) = U(x) = u(x)$. If V is a linear functional on $\mathscr{F}$, then there exists some affine $u : X \to \mathbb{R}$ and unique $\mu \in \Delta(S)$ such that $V(f) = \sum_{s \in S} u(f)\mu(s)$.

Lemma A.1. *If V, U, W are normalized linear functionals and there exist $f, g \in \mathscr{F}$ such that $V(f) > V(g)$, then for any $\delta \in [0,1]$ the following are equivalent:*

- $V = (1 - \delta)U + \delta W$
- $\mu_V = (1 - \delta)\mu_U + \delta\mu_W$

Proof.

$$V(f) = (1 - \delta)U(f) + \delta W(f) \Leftrightarrow$$

$$\sum_{s \in S} u(f)\mu_V(s) = (1 - \delta) \sum_{s \in S} u(f)\mu_U(s) + \delta \sum_{s \in S} u(f)\mu_W(s) \Leftrightarrow$$

$$\sum_{s \in S} u(f)\mu_V(s) = \sum_{s \in S} u(f)(1 - \delta)\mu_U(s) + \sum_{s \in S} u(f)\delta\mu_W(s) \Leftrightarrow$$

$$\sum_{s \in S} u(f)\mu_V(s) = \sum_{s \in S} u(f)[(1 - \delta)\mu_U(s) + \delta\mu_W(s)] \Leftrightarrow$$

$$\mu_V(s) = (1 - \delta)\mu_U(s) + \delta\mu_W(s) \Leftrightarrow$$

where the last line follows since μ_V is unique, f is arbitrary and the equation must hold for any f . The non-degeneracy assumption is required to give uniqueness of μ_V. □

Let $\mathcal{C} = \{(A,f) \in \Sigma \times \mathscr{F} | f(s) \sim f(s') \text{ for all } s, s' \in A\}$. We refer to a scenario $(A,f) \in \mathcal{C}$ as a constant scenario, since the action f returns constant payoff within A, and hence is now equivalent to a constant action. Let $\mathcal{N} = \mathcal{C}^c$ denote the set of non-constant scenarios.

Lemma A.2. *There exists a non-constant, affine utility function* $u : X \to \mathbb{R}$ *and a collection of probability distributions* $\{\mu, \mu_{A,f} | (A,f) \in \Sigma \times \mathscr{F}\}$ *such that* $\succsim$ $(\succsim_{A,f})$ *has an expected utility representation. Additionally,* $\mu(s) > 0$ *for every* $s \in S$*, and for each* (A,f), $\mu_{A,f}(s) > 0$ *for all* $s \in A$.

Proof. By Axiom 1 it is standard to show the existence of (u, μ) and $(u_{A,f}, \mu_{A,f})$ that represent $\succsim$ and $\succsim_{A,f}$, respectively. By ordinal preference consistency we know that $u(x) \geq u(y)$ if and only if $u_{A,f}(x) \geq u_{A,f}(y)$, hence it follows that $u_{A,f}$ is a positive affine transformation of u, so we simply apply the normalization that $u_{A,f} := u$. It also follows from monotonicity that $\mu(s) > 0$ for all s and $\mu_{A,f}(s) > 0$ for all $s \in A$. □

Let $V, V_{A,f}$ denote the linear functionals generated by $(u, \mu), (u, \mu_{A,f})$, that represent $\succsim, \succsim_{A,f}$, respectively. Further, let V_A and $V_{\mathscr{D}(A,f)}$ denote the linear functionals determined by $(u, \mu_{|A})$ and $(u, \mu_{|\mathscr{D}(A,f)})$, respectively. Note that these functionals are normalized by u so that $V(x) = V_{A,f}(x) = V_A(x) = V_{\mathscr{D}(A,f)}(x) = u(x)$.

Lemma A.3. *The relation of strong* A*-comonotonicity* ($\asymp_A$) *is an equivalence relation. Further, the set* $C_A(f) = \{h \in \mathscr{F} | h \asymp_A f\}$ *is convex.*

Proof. Reflexivity and symmetry are trivial. For transitivity, say that $h \asymp_A g$ and $g \asymp_A f$. Then $h(s) \succsim h(s') \Leftrightarrow g(s) \succsim g(s') \Leftrightarrow f(s) \succsim f(s')$, hence $h \asymp_A f$.

To prove convexity of $C_A(f)$, suppose $h, g \in C_A(f)$, and let $\alpha \in (0,1)$. Fix any $s, s' \in A$ and suppose $f(s) \succsim f(s')$. Then by $h(s) \succsim h(s')$ and $g(s) \succsim g(s')$), and the fact that u represents $\succsim$, $u(h(s)) \geq u(h(s'))$ and $u(g(s)) \geq u(g(s'))$. By the affinity of u it follows that $\alpha u(h(s)) + (1-\alpha)u(g(s)) \geq \alpha u(h(s')) + (1-\alpha)u(g(s')) \Leftrightarrow u(\alpha h(s) + (1-\alpha)g(s)) \geq u(\alpha h(s') + (1-\alpha)g(s')) \Leftrightarrow (\alpha h + (1-\alpha)g)(s) \succsim (\alpha h + (1-\alpha)g)(s')$. If we replace $f(s) \succsim f(s')$ with $f(s) \succ f(s')$, all inequalities become strict. To see the other direction, suppose for contradiction that $(\alpha h + (1-\alpha)g)(s) \succsim (\alpha h + (1-\alpha)g)(s')$ but $f(s') \succ f(s)$. But from the above argument, if $f(s') \succ f(s)$, then $(\alpha h + (1-\alpha)g)(s') \succ (\alpha h + (1-\alpha)g)(s)$, a contradiction. Hence it must be $f(s) \succsim f(s')$. Since s, s' were arbitrary, $\alpha h + (1-\alpha)g \in C_A(f)$. □

Let cl denote the closure operator and let $K_A(f) = \{h \in \mathscr{F} | f(s) \sim f(s') \implies h(s) \sim h(s')$ and $f(s) \succ f(s') \implies h(s) \succsim h(s')$ for all $s, s' \in A\}$. Notice that $C_A(f) \subset K_A(f)$ and that $K_A(f)$ contains all constant acts. Notice that if $(A, f) \in \mathcal{C}$, then $C_A(f) = K_A(f)$.

Lemma A.4. *Then* $cl(C_A(f)) = K_A(f)$.

Proof. First we show $K(f) \subset cl(C_A(f))$. Fix any $h \in K(f)$, then for each $n \in \mathbb{N}$, let $f_n = \frac{1}{n}f + (1-\frac{1}{n})h$. Clearly $f_n \to h$. For every s, s' and any n, $f(s) \sim f(s') \implies h(s) \sim h(s')$, hence $u(f(s)) = u(f(s')) \Leftrightarrow \frac{1}{n}u(f(s)) + (1-\frac{1}{n})u(h(s)) = \frac{1}{n}u(f(s')) + (1-\frac{1}{n})u(h(s')) \Leftrightarrow u(\frac{1}{n}f(s) + (1-\frac{1}{n})h(s)) = u(\frac{1}{n}f(s') + (1-\frac{1}{n})h(s')) \Leftrightarrow f_n(s) \sim f_n(s')$. Similarly, if $f(s) \succ f(s')$, then $h(s) \succsim h(s')$ and thus $\frac{1}{n}u(f(s)) + (1-\frac{1}{n})u(h(s)) > \frac{1}{n}u(f(s')) + (1-\frac{1}{n})u(h(s')) \Leftrightarrow u(\frac{1}{n}f(s) + (1-\frac{1}{n})h(s)) > u(\frac{1}{n}f(s') + (1-\frac{1}{n})h(s')) \Leftrightarrow f_n(s) \succ f_n(s')$. Hence $f_n \in C_A(f)$ for all n, and $h \in cl(C_A(f))$.

Next we must show that $cl(C_A(f)) \subset K_A(f)$. Towards a contradiction, suppose $g \in cl(C_A(f))$ but $g \notin K_A(f)$. If $g \notin K_A(f)$, then it must be that there is some pair of states s, s' such that $f(s) \sim f(s')$ and $g(s) \nsim g(s')$ or $f(s) \succ f(s')$ and $g(s') \succ g(s)$. In the first case, without loss suppose $g(s) \succ g(s')$, and $f_n \to g$ for some sequence $\{f_n\} \subset C_A(f)$. Then there is some n^* such that for any $n \geq n^*$, $f_n(s) \succ f_n(s')$. But for all $h \in C_A(f)$, $f(s) \sim f(s') \Leftrightarrow h(s) \sim h(s')$. Hence for any sequence $f_n \in C_A(f)$, $f_n(s) \sim f_n(s')$ for all n, contradiction. In the second case, again suppose for contradiction that $f(s) \succ f(s')$, $g(s') \succ g(s)$, and $f_n \to g$ for some sequence $\{f_n\} \subset C_A(f)$. Then there is some n^* such that for any $n \geq n^*$, $f_n(s') \succ f_n(s)$, but $f_n \in C_A(f)$ implies $f_n(s) \succ f_n(s')$, a contradiction. □

Corollary A.1. *Since the closure of a convex set is convex, if follows that $K_A(f)$ is convex.*

Lemma A.5. *For all $(A, f) \in \mathcal{N}$, (i) $\mathscr{D}(A, f) \subset A$ is non-empty, (ii) $f(s) \sim f(s')$ for all $s, s' \in \mathscr{D}(A, f)$, (iii) if $s \in \mathscr{D}(A, f)$ and $s' \notin \mathscr{D}(A, f)$, then $f(s) \succ f(s')$, (iv) if $h \asymp_A f$, then $\mathscr{D}(A, h) = \mathscr{D}(A, f)$.*

Proof. These follow directly from the definition. □

A.3 Proofs

A.3.1 Proof of Theorem 1.1

Lemma A.6. *If the collection of preferences $\{\succsim, \succsim_{A,f}\}$ satisfy axioms Axiom 1.1, Axiom 3.2, Axiom 1.4, then for all (A, f) and $B \subset A$ with $f(s) \sim f(\tilde{s})$ for all $s, \tilde{s} \in B$, then for any $h, g, z \in \mathscr{F}$,*

$$hBz \succsim gBz \Leftrightarrow hBz \succsim_{A,f} gBz.$$

Proof. Fix any $s, \tilde{s}$ such that $f(s) \sim f(\tilde{s})$. Then fix x, y, z such that $x, y \succ z$ and $x\{s\}z \sim y\{\tilde{s}\}z$. Since $f(s) \succsim f(\tilde{s})$ and $x\{s\}z \succsim y\{\tilde{s}\}z$, then by Axiom 1.4 it follows that $x\{s\}z \succsim_{A,f} y\{\tilde{s}\}z$. However, by symmetry it also follows from Axiom 1.4 that $y\{\tilde{s}\}z \succsim_{A,f} x\{s\}z$ and hence $x\{s\}z \sim_{A,f} y\{\tilde{s}\}z$. From this it follows that $u(x)\mu(s) +$

$u(z)(1-\mu(s)) = u(y)\mu(\tilde{s}) + u(z)(1-\mu(\tilde{s}))$ and $u(x)\mu_{A,f}(s) + u(z)(1-\mu_{A,f}(s)) = u(y)\mu_{A,f}(\tilde{s}) + u(z)(1-\mu_{A,f}(\tilde{s}))$. After algebra we conclude that $\frac{\mu(s)}{\mu(\tilde{s})} = \frac{\mu_{A,f}(s)}{\mu_{A,f}(\tilde{s})}$.

Now fix $B \subset A$ such that $f(s) \sim f(\tilde{s})$ for all $s, \tilde{s} \in B$. Since the above holds for all $s, \tilde{s} \in B$, we have

$$\begin{aligned}
\mu_{A,f}(\tilde{s})\mu(s) = \mu_{A,f}(s)\mu(\tilde{s}) &\Leftrightarrow \\
\sum_{\tilde{s}\in B} \mu_{A,f}(\tilde{s})\mu(s) = \sum_{\tilde{s}\in B} \mu_{A,f}(s)\mu(\tilde{s}) &\Leftrightarrow \\
\mu_{A,f}(B)\mu(s) = \mu_{A,f}(s)\mu(B) &\Leftrightarrow \\
\mu_{A,f}(s) = \mu(s)\frac{\mu_{A,f}(B)}{\mu(B)}
\end{aligned}$$

Now for any $h, g, z \in \mathscr{F}$,

$$\begin{aligned}
hBz \succsim gBz &\Leftrightarrow \\
\sum_{s\in B} u(h(s))\mu(s) + (1-\mu(B))u(z) \geq \sum_{s\in B} u(g(s))\mu(s) + (1-\mu(B))u(z) &\Leftrightarrow \\
\sum_{s\in B} u(h(s))\mu(s) \geq \sum_{s\in B} u(g(s))\mu(s) &\Leftrightarrow \\
\left(\frac{\mu_{A,f}(B)}{\mu(B)}\right) \sum_{s\in B} u(h(s))\mu(s) \geq \left(\frac{\mu_{A,f}(B)}{\mu(B)}\right) \sum_{s\in B} u(g(s))\mu(s) &\Leftrightarrow \\
\sum_{s\in B} u(h(s))\mu_{A,f}(s) \geq \sum_{s\in B} u(g(s))\mu_{A,f}(s) &\Leftrightarrow \\
\sum_{s\in B} u(h(s))\mu_{A,f}(s) + (1-\mu_{A,f}(B))u(z) \geq \sum_{s\in B} u(g(s))\mu_{A,f}(s) + (1-\mu_{A,f}(B))u(z) &\Leftrightarrow \\
hBz \succsim_{A,f} gBz
\end{aligned}$$

Since B was arbitrary, the result holds.

□

Lemma A.7. *For every* $(A, f) \in \mathcal{C}$, $\mu_{A,f} = \mu_{|A}$.

Proof. First, by Axiom 3.2 it follows that $\mu_{A,f}(A) = 1$, or equivalently, $h \sim_{A,f} hAz$

and $g \sim_{A,f} gAz$. The proof of this is standard. Next, since $f(s) \sim f(s')$ for all $s, s' \in A$, then from Lemma A.6 and taking $B = A$

$$\begin{aligned} hAz \succsim_{A,f} gAz \Leftrightarrow \Leftrightarrow \\ hAz \succsim gAz \Leftrightarrow \\ \sum_{s\in A} u(h(s))\mu(s) + (1-\mu(A))u(z) \geq \sum_{s\in A} u(g(s))\mu(s) + (1-\mu(A))u(z) \Leftrightarrow \\ \frac{1}{\mu(A)} \sum_{s\in A} u(h(s))\mu(s) \geq \frac{1}{\mu(A)} \sum_{s\in A} u(g(s))\mu(s) \Leftrightarrow \\ \sum_{s\in A} u(h(s))\mu_{|A}(s) \geq \sum_{s\in A} u(g(s))\mu_{|A}(s) \end{aligned}$$

Since $\mu_{A,f}$ is unique, it follows that for any $(A, f) \in \mathcal{C}$, $\mu_{A,f}(s) = \mu_{|A}(s)$. Since for all $x \in X$, $(A, x) \in \mathcal{C}$, $\mu_{A,x} = \mu_{|A}$. □

For all (A, f), define $\psi(A, f, s) := \frac{\mu_{A,f}(s)}{\mu_{A,x}(s)}$ for some x. By Lemma A.7 $\mu_{A,x} = \mu_{A,y}$ for all x, y, hence the function is well defined, and $\frac{\mu_{A,f}(s)}{\mu_{A,x}(s)} = \frac{\mu_{A,f}(s)}{\mu(s)}\mu(A) > 0$. By definition it is clear that

$$\sum_{s\in A} u(g(s))\psi(A,f,s)\mu(s|A) = \sum_{s\in A} u(g(s)) \left(\frac{\mu_{A,f}(s)}{\mu(s)}\mu(A) \right) \frac{\mu(s)}{\mu(A)} = \sum_{s\in A} u(h(s))\mu_{A,f}(s)$$

Suppose that $f(s) \sim f(\tilde{s})$. Then let $B = \{s, \tilde{s}\}$, and let $h = x\{s\}z, g = y\{\tilde{s}\}z$. Then by Lemma A.6, $hBz \succsim gBz \Leftrightarrow hBz \succsim_{A,f} gBz$. Without loss choose $x, y \succ z$ such that $h \sim g$, hence it follows that

$$\begin{aligned} \frac{\mu_{A,f}(s)}{\mu_{A,f}(\tilde{s})} = \frac{\mu(s)}{\mu(\tilde{s})} \implies \\ \frac{\psi(A,f,s)}{\mu(A)} = \frac{\mu_{A,f}(s)}{\mu(s)} = \frac{\mu_{A,f}(\tilde{s})}{\mu(\tilde{s}} = \frac{\psi(A,f,\tilde{s})}{\mu(A)} \implies \\ \psi(A,f,s) = \psi(A,f,\tilde{s}) \end{aligned}$$

Let $\{E_1, \ldots, E_n\}$ be a partition of A such that for all $s, \tilde{s} \in E_i$, $f(s) \sim f(\tilde{s})$ and for

$i < j$ and any $s \in E_i$ and $\tilde{s} \in E_j$, $f(s) \succ f(\tilde{s})$. Then define $\psi_{A,f}(E_i) = \psi(A,f,s)$ for some $s \in E_i$. By the result above this is well defined. Next, E_k, E_{k+1} satisfy the conditions of Axiom 1.4. Thus let x, y, z satisfy $xE_kz \sim yE_{k+1}z$ with $x, y \succ z$. Then it follows that $xE_kz \succsim_{A,f} yE_{k+1}z$ and $u(x)\mu(E_k) + u(z)(1 - \mu(E_k)) = u(y)\mu(E_{k+1}) + u(z)(1 - \mu(E_{k+1}))$, hence $(u(x) - u(z))\mu(E_k) = (u(y) - u(z))\mu(E_{k+1})$.

$$
\begin{aligned}
xE_kz \succsim_{A,f} yE_{k+1}z \Leftrightarrow \\
u(x)\frac{\mu(E_k)}{\mu(A)}\psi_{A,f}(E_k) + u(z)[1 - \frac{\mu(E_k)}{\mu(A)}\psi_{A,f}(E_k)] \geq \\
u(y)\frac{\mu(E_{k+1})}{\mu(A)}\psi_{A,f}(E_{k+1}) + u(z)[1 - \frac{\mu(E_{k+1})}{\mu(A)}\psi_{A,f}(E_{k+1})] \Leftrightarrow \\
(u(x) - u(z))\mu(E_k)\psi_{A,f}(E_k) \geq \\
(u(y) - u(z))\mu(E_{k+1})\psi_{A,f}(E_{k+1}) \Leftrightarrow \\
\psi_{A,f}(E_k) \geq \psi_{A,f}(E_{k+1})
\end{aligned}
$$

Next, define $\delta_{A,f} : X \to (0, \infty)$ by $\delta_{A,f}(f(s)) = \psi(A,f,s)$ if $x = f(s)$. We can define $\delta_{A,f}$ outside of $f(A)$ so that it is non-decreasing, by making it constant for $x \prec \inf\{f(s)|s \in A\}$ and $x \succ \sup\{f(s)|s \in A\}$, and extending it piecewise linearly otherwise. ■

A.3.2 Proof of Theorem 1.2

The uniqueness properties are simple to show. The uniqueness properties of u, μ, and $\mu_{A,f}$ all follow from standard results. Given $\mu, \mu_{A,f}$, there is a unique value for $\delta(A,f,f(s))$ given by $\delta(A,f,f(s)) = \frac{\mu_{A,f}(s)}{\mu(s)}\mu(A)$. We can define $\delta(A,f,x)$ arbitrarily outside of $f(A)$.

A.3.3 Proof of Theorem 1.3

The proof proceeds by constructing a value function v such that $\delta_{A,f}$ is a proportional distortion relative to v.

Lemma A.8. *For all* $(A, f), (B, g)$ *and any* $s_1, s_2 \in A$, $s_3, s_4 \in B$, *if* $f(s_1) \sim g(s_3)$ *and* $f(s_2) \sim g(s_4)$, *then* $\frac{\psi(A,f,s_1)}{\psi(A,f,s_2)} = \frac{\psi(B,g,s_3)}{\psi(B,g,s_4)}$.

Proof. **Case 1:** Suppose $s_1 = s_3 = s$ and $s_2 = s_4 = \tilde{s}$, then for $C = \{s, \tilde{s}\}$, $f(s) \sim g(s)$ for all $s \in C$. For all h, j, z

$$\begin{aligned} hCz \succsim_{A,f} jCz \Leftrightarrow \\ u(h(s))\mu_{A,f}(s) + u(h(\tilde{s}))\mu_{A,f}(\tilde{s}) + u(z)(1 - \mu_{A,f}(C)) \geq \\ u(j(s))\mu_{A,f}(s) + u(j(\tilde{s}))\mu_{A,f}(\tilde{s}) + u(z)(1 - \mu_{A,f}(C)) \Leftrightarrow \\ [u(h(s)) - u(j(s))]\mu_{A,f}(s) \geq [u(j(\tilde{s})) - u(h(\tilde{s}))]\mu_{A,f}(\tilde{s}) \end{aligned}$$

Suppose h, j are such that $h(s) \succ j(s)$ and $j(\tilde{s}) \succ h(\tilde{s})$. Then by Axiom 1.5, $hCz \sim_{A,f} jCz \Leftrightarrow hCz \sim_{B,g} jCz$, and hence

$$\frac{\mu_{A,f}(s)}{\mu_{A,f}(\tilde{s})} = \frac{u(j(\tilde{s})) - u(h(\tilde{s}))}{u(h(s)) - u(j(s))} = \frac{\mu_{B,g}(s)}{\mu_{B,g}(\tilde{s})}.$$

Since $\psi(A, f, s) := \frac{\mu_{A,f}(s)}{\mu_{A,x}(s)}$, it follows that

$$\frac{\psi(A,f,s)}{\psi(A,f,\tilde{s})} = \frac{\mu_{A,f}(s)}{\mu_{A,x}(s)} / \frac{\mu_{A,f}(\tilde{s})}{\mu_{A,x}(\tilde{s})} = \frac{\mu_{A,f}(s)}{\mu_{A,f}(\tilde{s})} \times \frac{\mu(\tilde{s})}{\mu(s)} = \frac{\mu_{B,g}(s)}{\mu_{B,g}(\tilde{s})} \times \frac{\mu(\tilde{s})}{\mu(s)} = \frac{\psi(B,g,s)}{\psi(B,g,\tilde{s})}.$$

Case 2: $C \equiv \{s_1, s_2, s_3, s_4\} \subset A \cap B$. Then define h such that $h(s_1) = h(s_3)$ and $h(s_2) = h(s_4)$. Then

$$\frac{\psi(A,f,s_1)}{\psi(A,f,s_2)} = \frac{\psi(B,h,s_1)}{\psi(B,h,s_2)} = \frac{\psi(B,h,s_3)}{\psi(B,h,s_4)} = \frac{\psi(B,g,s_3)}{\psi(B,g,s_4)}.$$

Where the first and third equalities follows by case 1, while the second equality holds since from Theorem 1, $h(s_1) \sim h(s_3) \implies \psi(B,h,s_1) = \psi(B,h,s_3)$ and $h(s_2) \sim h(s_4) \implies \psi(B,h,s_2) = \psi(B,h,s_4)$

Case 3: Suppose $\{s_3, s_4\} \cap A = \varnothing$ or $\{s_1, s_2\} \cap B = \varnothing$[1]. Then let $D = \{s_1, s_2, s_3, s_4\}$. As before, define h such that $h(s_1) = h(s_3)$ and $h(s_2) = h(s_4)$. It then follows that

$$\frac{\psi(A,f,s_1)}{\psi(A,f,s_2)} = \frac{\psi(D,h,s_1)}{\psi(D,h,s_2)} = \frac{\psi(D,h,s_3)}{\psi(D,h,s_4)} = \frac{\psi(B,g,s_3)}{\psi(B,g,s_4)}$$

Where the equalities follow by repeated applications of case 1 or case 2. □

Define the function $\phi : X \times X \to \mathbb{R}_+$ by

$$\phi(x,y) := \frac{\psi(A,f,s)}{\psi(A,f,\tilde{s})} \tag{A.1}$$

for some (A, f) where $f(s) = x$ and $f(\tilde{s}) = y$. By the previous lemma, for all (A, f) and (B, g) such that $f(s_1) = x = g(s_3)$ and $f(s_2) = y = g(s_4)$, $\frac{\psi(A,f,s_1)}{\psi(A,f,s_2)} = \frac{\psi(B,g,s_3)}{\psi(B,g,s_4)}$, hence ϕ is well defined.

Lemma A.9. *ϕ satisfies the following properties: (i) $x \succsim y \implies \phi(x,y) \geq 1$, (ii) $\phi(x,y)\phi(y,z) = \phi(x,z)$, (iii) $\frac{1}{\phi(x,y)} = \phi(y,x)$, and (iv) $\phi(x,x) = 1$*

Proof. *(i)* Fix $s, \tilde{s}$ such that $f(s) = x \succsim y = f(\tilde{s})$. By the previous theorem $\psi(A,f,s) \geq \psi(A,f\tilde{s})$, hence $\phi(x,y) = \frac{\psi(A,f,s)}{\psi(A,f,\tilde{s})} \geq 1$. *(ii)* Fix three states s_x, s_y, s_z, where $f(s_i) = i$, $i \in \{x,y,z\}$. Then $\phi(x,y)\phi(y,z) = \frac{\psi(A,f,s_x)}{\psi(A,f,s_y)}\frac{\psi(A,f,s_y)}{\psi(A,f,s_z)} = \phi(x,z)$. *(iii)* For any $s, \tilde{s}$ with $f(s) = x, f(\tilde{s}) = y$, $\frac{1}{\phi(x,y)} = \frac{1}{\frac{\psi(A,f,s)}{\psi(A,f,\tilde{s})}} = \frac{\psi(A,f,\tilde{s})}{\psi(A,f,s)} = \phi(y,x)$. *(iv)* It follows from *(iii)* $\phi(x,x) = \frac{1}{\phi(x,x)}$, hence $\phi(x,x)\phi(x,x) = \phi(x,x) = 1$. □

[1]If only a single state is missing from A, the steps are similar

Fix some $x_* \in X$[2] and define $v : X \to \mathbb{R}_+$ by

$$v(x) := \phi(x, x_*). \tag{A.2}$$

Then for any (A, f) such that $x = f(s)$ and $y = f(\tilde{s})$ for some $s, \tilde{s} \in A$,

$$\frac{v(x)}{v(y)} = \frac{\phi(x,x_*)}{\phi(y,x_*)} = \phi(x,y) = \frac{\psi(A,f,s)}{\psi(A,f,\tilde{s})} = \frac{\mu_{A,f}(s)}{\mu_{A,f}(\tilde{s})}\frac{\mu(\tilde{s})}{\mu(s)} \Leftrightarrow$$

$$\frac{v(x)}{v(y)}\frac{\mu(s|A)}{\mu(\tilde{s}|A)} = \frac{\mu_{A,f}(s)}{\mu_{A,f}(\tilde{s})} \Leftrightarrow \tag{A.3}$$

$$\mu_{A,f}(\tilde{s}) = \frac{\mu_{A,f}(s)}{v(f(s))\mu(s|A)} v(f(\tilde{s}))\mu(\tilde{s}|A)$$

Thus, summing over $\tilde{s}$ yields

$$1 = \sum_{\tilde{s}} \mu_{A,f}(\tilde{s}) = \left(\sum_{\tilde{s}} v(f(\tilde{s}))\mu(\tilde{s}|A)\right) \frac{\mu_{A,f}(s)}{v(f(s))\mu(s|A)},$$

hence

$$\mu_{A,f}(s) = \frac{v(f(s))}{\sum_{\tilde{s}} v(f(\tilde{s}))\mu(\tilde{s}|A)} \mu(s|A). \tag{A.4}$$

Lemma A.10. *v is $\succsim$-increasing.*

Proof. Suppose $x \succsim y$. Then $\frac{v(x)}{v(y)} = \frac{\phi(x,x_*)}{\phi(y,x_*)} = \phi(x,x_*)\phi(x_*,y) = \phi(x,y) \geq 1$, hence $v(x) \geq v(y)$.

□

[2]For example, suppose $X = \Delta(Y)$ for a finite prize set Y as in Anscombe and Aumann [4]. Then a natural choice for x_* is the $\succsim$-worst prize in Y. Alternatively, If X is some interval of $\mathbb{R}$ containing 0, then we might take $x_* = 0$.

A.3.4 Proof of Theorem 1.4

The representation is given by the triple (u, v, μ). Given the standard uniqueness for μ and $\mu_{A,f}$, then v' and v represent the same preferences if and only if for all x, y, $\frac{v'(x)}{v'(y)}\frac{\mu(s)}{\mu(s)} = \frac{v(x)}{v(y)}\frac{\mu(s)}{\mu(s)}$, if and only if for some $k > 0$, $v'(x) = kv(x)$. ■

A.3.5 Proof of Theorem 1.5

Since for all $(A, f) \in \mathcal{C}$, $\mathscr{D}(A, f) = A$, hence for all $\delta \in [0, 1]$,

$$\mu_{A,f} = \mu_{|A} = (1-\delta)\mu_{|A} + \delta\mu_{|\mathscr{D}(A,f)}.$$

For the remainder of the proof, suppose $(A, f) \in \mathcal{N}$. Let $H(A, f) := C_A(f) \cup X$, where X is understood to mean the set of constant acts. Note that $H(A, f) \subset cl(C_A(f))$.

Step 2: There exists a function $\delta : \mathcal{N} \to [0, 1]$ such that for all $h, g \in cl(C_A(f))$,

$$h \succsim_{A,f} g \Leftrightarrow$$

$$(1-\delta(A, f))V_A(h) + \delta(A, f))V_{\mathscr{D}(A,f)}(h) \geq$$

$$(1-\delta(A, f))V_A(g) + \delta(A, f)V_{\mathscr{D}(A,f)}(g)$$

Proof. Fix some $(A, f) \in \mathcal{N}$. Define the relation $\geqslant_A$ by:

$$h \geqslant_A g \text{ if and only if } hAg \succsim g,$$

and let $\gg_A$ and $\cong_A$ denote the strict and symmetric parts of $\geqslant_A$. It is simple to show that $hAg \succsim g$ if and only if $hAg' \succsim gAg'$ for all $g' \in \mathscr{F}$. Then for any h, g,

$$\begin{aligned} h \geqslant_A g &\Leftrightarrow \\ hAg \succsim g &\Leftrightarrow \\ \sum_{s \in A} u(h(s))\mu(s) + \sum_{s \in A^c} u(g(s))\mu(s) \geq \sum_{s \in A} u(g(s))\mu(s) + \sum_{s \in A^c} u(g(s))\mu(s) &\Leftrightarrow \\ \frac{1}{\mu(A)} \sum_{s \in A} u(h(s))\mu(s) \geq \frac{1}{\mu(A)} \sum_{s \in A} u(g(s))\mu(s) &\Leftrightarrow \\ \sum_{s \in A} u(h(s))\mu_{|A}(s) \geq \sum_{s \in A} u(g(s))\mu_{|A}(s) & \end{aligned}$$

Hence $h \geqslant_A g$ has a subjective expected utility representation $(u, \mu_{|A})$.

Next, define $\trianglerighteq_A$ by $h \trianglerighteq_A g$ if and only if for some $s \in A$, $h(s) \succsim g(s')$ for all $s' \in A$. let $\triangleright_A$ and $\approx_A$ denote the strict and symmetric parts of $\trianglerighteq_A$. Further, $\trianglerighteq_A$ is represented by $M_A(h) = \max\{u(h(s)) | s \in A\}$. Suppose $h \trianglerighteq_A g$. Then for some $\hat{s} \in A$, $h(\hat{s}) \succsim g(s')$ for all $s' \in A$. Hence $\max\{u(h(s)) | s \in A\} \geq u(h(\hat{s})) \geq \max\{u(g(s)) | s \in A\}$. Next, suppose $\max\{u(h(s)) | s \in A\} \geq \max\{u(g(s)) | s \in A\}$. Then let s^* solve $u(h(s^*)) = \max\{u(h(s)) | s \in A\}$. Then clearly $h(s^*) \succsim g(s)$ for all $s \in A$.

If $h, g \in C_A(f)$, it follows that $\mathscr{D}(A, h) = \mathscr{D}(A, g) = \mathscr{D}(A, f)$, hence $h \trianglerighteq_A g$ is equivalent to $h(s) \succsim g(s)$ for all $s \in \mathscr{D}(A, f)$. By the previous lemma, for $h, g \in C_A(f)$ there is some x_h, x_g such that $h(s) \sim x_h$ and $g(s) \sim x_g$ for all $s \in \mathscr{D}(A, f)$. It is then clear that $h \trianglerighteq_A g \Leftrightarrow x_h \succsim x_g$ for $h, g \in C_A(f)$, or equivalently, $u(x_h) = \max\{u(h(s)) | s \in A\}$. Then for any $\rho \in \Delta(S)$ satisfying $\rho(\mathscr{D}(A, f)) = 1$,

$$\sum_{s \in \mathscr{D}(A,f)} u(h(s))\rho(s) = u(x_h) = M_A(h).$$

Further, for any $x \in \mathscr{F}$,

$$\sum_{s \in \mathscr{D}(A,f)} u(x(s))\rho(s) = u(x) = M_A(x),$$

hence

$$U_\rho(h) := \sum_{s \in \mathscr{D}(A,f)} u(h(s))\rho(s)$$

represents $\trianglerighteq_A$ on $H(A,f)$, and U_ρ is a normalized linear functional on $\mathscr{F}$.

For $\rho \in \Delta(S)$ such that $\rho(\mathscr{D}(A,f)) = 1$, define the set

$$\mathscr{U}_\rho := \{(U_\rho(h), V_A(h)) \in \mathbb{R}^2 | h \in H(A,f)\}.$$

For $(v_1, v_2), (v_1', v_2') \in \mathscr{U}$, define $\succsim^*$ by

$$(v_1, v_2) \succsim^* (v_1', v_2') \Leftrightarrow h \succsim_{A,f} g$$

for some $h, g \in H(A,f)$ such that $(U_\rho(h), V_A(h)) = (v_1, v_2)$ and $(U_\rho(g), V_A(g)) = (v_1', v_2')$. This relation is well defined, since $(U_\rho(h), V_A(h)) = (U_\rho(g), V_A(g))$ implies $h \approx_A g$ and $h \cong_A g$, hence $h \sim_{A,f} g$. Further, for all $h \in H(A,f)$, $U_\rho(h) \geq V_A(h)$, hence $v_1 \geq v_2$ for all $(v_1, v_2) \in \mathscr{U}$. This holds because U_ρ coincides with the maximal payoff of h in A. Further, it is obvious that $\succsim^*$ is complete, transitive, monotonic, and satisfies independence and continuity. Let $\bar{s} \in \mathscr{D}(A,f)$ and $\underline{s} \in \arg\min\{u(f(s)) | s \in A\}$. Then since $(A,f) \in \mathcal{N}$, the constant acts $f(\bar{s})$ and $f(\underline{s})$ satisfy $f(\bar{s}) \succ f(\underline{s})$. Further, for $\alpha \in (0,1)$, $h := \alpha f + (1-\alpha) f(\bar{s})$ satisfies $u(f(\bar{s})) > U_\rho(h) > V_A(h) > u(f(\underline{s}))$, hence there are $(v_1^*, v_2^*), (\bar{v}, \bar{v}), (\underline{v}, \underline{v}) \in \mathscr{U}$ such that $v_1^* > v_2^*$ and $\bar{v} > v_1^* > \underline{v}$, where this follows due to the convexity of $H(A,f)$ and the fact that U_ρ and V_A are normalized. Hence by lemma 2 of Saito

[58], there exists some $\delta \in [0,1]$ such that

$$\delta U_\rho(h) + (1-\delta)V_A(h) \geq \delta U_\rho(g) + (1-\delta)V_A(g) \Leftrightarrow \delta v_1 + (1-\delta)v_2 \geq \delta v_1' + (1-\delta)v_2'$$

$$\Leftrightarrow (v_1, v_2) \succsim^* (v_1', v_2') \Leftrightarrow h \succsim_{A,f} g.$$

By lemma 2, for every ρ, $W_\rho := (1-\delta)V_A(h) + \delta U_\rho(h)$ is a normalized linear functional given by $\mu_W = (1-\delta)\mu_{|A} + \delta\rho$. Then by continuity of W_ρ we extend it to $cl(C_A(f))$. If $\mathscr{D}(A,f)$ is a singleton then ρ is uniquely given. Suppose then without loss that $|\mathscr{D}(A,f)| \geq 2$ and fix $s, s' \in \mathscr{D}(A,f)$. Consider x, y, z such that $x, y \succ z$ and $x\{s\}z \sim y\{s'\}z$. Then define $h := x\{s\}z$, $g := y\{s'\}z$ and $B := \{s, s'\}$. Then by $hBz \sim gBz$, and by Lemma A.6, $hBz \sim_{A,f} gBz \Leftrightarrow x\{s\}z \sim_{A,f} y\{s'\}z$. W_ρ satisfies the equation $W_\rho(x\{s\}z) = W_\rho(y\{s'\}z)$ if and only if

$$\begin{aligned}
\mu_W(s)u(x) + (1-\mu_W(s))u(z) = \mu_W(s')u(y) + (1-\mu_W(s'))u(z) \Leftrightarrow \\
[u(x) - u(z)]\mu_W(s) = [u(y) - u(z)]\mu_W(s') \Leftrightarrow \\
\frac{\mu_W(s)}{\mu_W(s')} = \frac{u(y) - u(z)}{u(x) - u(z)}
\end{aligned}$$

However, from $x\{s\}z \sim y\{s'\}z$ we also know that $\frac{\mu(s)}{\mu(s')} = \frac{u(y)-u(z)}{u(x)-u(z)}$, hence

$$\frac{(1-\delta)\mu_{|A}(s) + \delta\rho(s)}{(1-\delta)\mu_{|A}(s') + \delta\rho(s')} = \frac{\mu_W(s)}{\mu_W(s')} = \frac{\mu(s)}{\mu(s')}.$$

Algebra yields $\frac{\rho(s)}{\rho(s')} = \frac{\mu(s)}{\mu(s')}$, which when combined with $\rho(\mathscr{D}(A,f)) = 1$, implies $\rho = \mu_{|\mathscr{D}(A,f)}$ is the unique ρ such that W_ρ represents $\succsim_{A,f}$ on $cl(C_A(f))$, hence $W_{\mu_{\mathscr{D}(A,f)}} = (1-\delta)V_A(h) + \delta V_{\mathscr{D}(A,f)}(h)$.

Since for each (A,f) $\delta_{(A,f)}$ is unique, we simply define the function $\delta : \mathcal{N} \to [0,1]$ by $\delta(A,f) = \delta_{(A,f)}$. □

Step 3: For all $h, g \in \mathscr{F}$,

$$h \succsim_{A,f} g \Leftrightarrow (1-\delta(A,f))V_A(h) + \delta(A,f)V_{\mathscr{D}(A,f)}(h) \geq (1-\delta(A,f))V_A(g) + \delta(A,f)V_{\mathscr{D}(A,f)}(g).$$

Proof. Again by lemma 2, linearity of $(1-\delta(A,f))V_A + \delta(A,f)V_{\mathscr{D}(A,f)}$ implies we can extend it to all of $\mathscr{F}$ by the equation

$$(1-\delta(A,f))V_A(h) + \delta(A,f)V_{\mathscr{D}(A,f)}(h) = \sum_{s \in A} u(h(s))[(1-\delta(A,f))\mu_{|A} + \delta(A,f)\mu_{|\mathscr{D}(A,f)}].$$

That is, we can define a linear functional on $\mathscr{F}$ by

$$U_{A,f}(h) = \sum_{s \in A} u(h(s))[(1-\delta(A,f))\mu_{|A} + \delta(A,f)\mu_{|\mathscr{D}(A,f)}].$$

Then $U_{A,f}$ and $V_{A,f}$ are both normalized linear functionals that agree on $cl(C_A(f))$, and hence by uniqueness of subjective probabilities, it follows that $\mu_{A,f} = (1 - \delta(A,f))\mu_{|A} + \delta(A,f)\mu_{|\mathscr{D}(A,f)}$.

That is, I claim that agreement on $cl(C_A(f))$ is sufficient for uniqueness, or in other words, $cl(C_A(f))$ is large enough to identify beliefs uniquely. Consider any two states $s, s' \in A$. Then without loss $f(s) \succ f(s')$ or $f(s) \sim f(s')$. Suppose the first case holds, and for convenience, ignore the dependence of δ on (A,f). Then for any $x \succ y$, $x\{s\}y \in cl(C_A(f))$. Say for some $w \in X$, $x\{s\}y \sim_{A,f} w$, then it follows that $U_{A,f}(x\{s\}y) = u(w) = V_{A,f}(x\{s\}y)$, hence

$$u(x)\mu_{A,f}(s) + u(y)(1-\mu_{A,f}(s)) = u(x)[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s) + u(y)(1 - [(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,}$$

Since $u(x) > u(y)$, it immediately follows that $\mu_{A,f}(s) = [(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s)$

Next, it is immediately apparent that for $f(s) \sim f(s')$,

$$\frac{\mu_{A,f}(s)}{\mu_{A,f}(s')} = \frac{[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s)}{[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s')},$$

since either $\{s, s'\} \subseteq \mathscr{D}(A, f)$ or $\{s, s'\} \nsubseteq \mathscr{D}(A, f)$. Hence in any case, for any $s, s' \in A$

$$\frac{\mu_{A,f}(s)}{\mu_{A,f}(s')} = \frac{[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s)}{[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s')} \implies$$
$$\mu_{A,f}(s)[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s') = \mu_{A,f}(s')[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s) \implies$$
$$\sum_{s'\in A} \mu_{A,f}(s)[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s') = \sum_{s'\in A} \mu_{A,f}(s')[(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s) \implies$$
$$\mu_{A,f}(s) = [(1-\delta)\mu_A + \delta\mu_{|\mathscr{D}(A,f)}](s)$$

□

The previous steps conclude the proof of Theorem 4. The following step proves Theorem 8, the case where $\delta(A, f)$ is a constant function.

Step 4: For all $(A, f), (B, g) \in \mathcal{N}$, $\delta(A, f) = \delta(B, g)$.

Proof. **(A)** First, note that $\tilde{s} \in A \cap B$ and $s \in A, s' \in B$ such that $f(s) \succ f(\tilde{s}), g(s') \succ g(\tilde{s})$ is equivalent to $\tilde{s} \in (A \backslash \mathscr{D}(A, f)) \cap (B \backslash \mathscr{D}(B, g))$. Consider some $s \in (A \backslash \mathscr{D}(A, f)) \cap (B \backslash \mathscr{D}(B, g))$. Choose $x, y, z \in X$ and so that $yAz \sim xBz$, and without loss suppose $u(z) = 0$. By Axiom A.5, this is equivalent to $[w \sim_{A,f} x\{s\}z \Leftrightarrow w \sim_{B,g} y\{s\}z]$, whence, $V_{A,f}(x\{s\}z) = u(x)(1-\delta(A, f))\mu(s|A) = u(w) = u(y)(1-\delta(B, g))\mu(s|B) = V_{B,g}(y\{s\}z)$. From $yAz \sim xBz$ it follows that $u(y)\mu(A) = u(x)\mu(B) \Leftrightarrow \frac{u(y)}{u(x)} = \frac{\mu(B)}{\mu(A)}$, which when combined with $u(x)(1-\delta(A, f))\frac{\mu(s)}{\mu(A)} = u(y)(1-\delta(B, g))\frac{\mu(s)}{\mu(B)}$, immediately implies that $\delta(A, f) = \delta(B, g)$.

(B) Now suppose $(A\backslash\mathscr{D}(A,f)\cap(B\backslash\mathscr{D}(B,g))=\varnothing$.

There are two mutually exclusive possibilities:

(I) $\mathscr{D}(A,f)\cup\mathscr{D}(B,g)=S$

(II) There is some state $s\in S\backslash(\mathscr{D}(A,f)\cup\mathscr{D}(B,g))$

Case (I): Let $s^f\in\mathscr{D}(A,f)\backslash\mathscr{D}(B,g)$ and let $s^h\in\mathscr{D}(B,g)\backslash\mathscr{D}(A,f)$. Such states must exist because $\mathscr{D}(A,f)\subsetneq A$, $\mathscr{D}(B,g)\subsetneq B$, and $\mathscr{D}(A,f)\cup\mathscr{D}(B,g)=S$. Then for $x\succ y$, consider the act $x\{s^f,s^h\}y$. Then it plainly follows that $(A\backslash\mathscr{D}(A,f))\cap(A\backslash\mathscr{D}(A,x\{s^f,s^h\}y))\neq\varnothing$ and that $(B\backslash\mathscr{D}(B,g))\cap(B\backslash\mathscr{D}(B,x\{s^f,s^h\}y))\neq\varnothing$. Additionally, consider the event $C=\{s^f,s^h,s,s'\}$ where $s\in A\backslash\mathscr{D}(A,f)$ and $s'\in B\backslash\mathscr{D}(B,g)$. Then it is also the case that $(C\backslash\mathscr{D}(C,x\{s^f,s^h\}y))\cap(B\backslash\mathscr{D}(B,x\{s^f,s^h\}y))\neq\varnothing\neq(C\backslash\mathscr{D}(C,x\{s^f,s^h\}y))\cap(A\backslash\mathscr{D}(A,x\{s^f,s^h\}y))$. But by the conditions established we can apply the same argument in (A) and conclude that $\delta(A,f)=\delta(A,x\{s^f,s^h\}y)=\delta(C,x\{s^f,s^h\}y)=\delta(B,x\{s^f,s^h\}y)=\delta(B,g)$.

Case (II): Suppose $|A|=|S|-1=|B|$. Choose any $\tilde{s}\in(A\backslash\mathscr{D}(A,f))$, and define the act g' by $g'(s)=g(s)$ for $s\in B\backslash\{\tilde{s}\}$ and $g'(\tilde{s})=z$ for some $z\prec g(s)$, for $s\in\mathscr{D}(B,g)$. Then $(A\backslash\mathscr{D}(A,f))\cap(B\backslash\mathscr{D}(B,g'))\neq\varnothing\neq(B\backslash\mathscr{D}(B,g'))\cap(B\backslash\mathscr{D}(B,g'))$. Hence by the argument in (A), $\delta(A,f)=\delta(B,g')=\delta(B,g)$.

If $|A|<|S|-1=|B|$, pick some $s'\in B\backslash\mathscr{D}(B,g)$ and let $\hat{A}=A\cup\{s'\}$. Let $\hat{f}(s)=f(s)$ for all $s\neq s'$, and let $\hat{f}(s')=z$, where $z\prec f(s)$ for some $s\in A$. Now it follows that $(A\backslash\mathscr{D}(E,f))\cap(\hat{A}\backslash\mathscr{D}(\hat{A},\hat{f}))$ and $(\hat{A}\backslash\mathscr{D}(\hat{A},\hat{f}))\cap(B\backslash\mathscr{D}(B,g))$ are both non-empty, hence $\delta(A,f)=\delta(\hat{A},\hat{f})=\delta(B,h)$. The case when $|B|<|S|-1=|A|$ is nearly identical.

Lastly, suppose $|A|\leq|S|-2$ and $|B|\leq|S|-2$. By the assumption that $(A,f),(B,g)\in\mathcal{N}$, there is some $z\in X$ so that for $s\in\mathscr{D}(A,f)$, $f(s)\succ z$, and for $s\in\mathscr{D}(B,g)$, $h(s)\succ z$. Next, define $\hat{f}$ so that for all $s\in A$, $\hat{f}(s)=f(s)$ and for all $s\notin A$,

$\hat{f}(s) = z$. Similarly, define $\hat{g}$ so that for all $s \in B$, $\hat{g}(s) = g(s)$, and for all $s \notin B$, $\hat{g}(s) = z$. Pick some $\hat{s} \in S\backslash(\mathscr{D}(A,f) \cup \mathscr{D}(B,g))$, which exists by assumption. Define $\hat{A} = A \cup \{\hat{s}\}$ and $\hat{B} = B \cup \{\hat{s}\}$. By construction, $(A\backslash\mathscr{D}_{A,f}) \cap (\hat{B}\backslash\mathscr{D}(\hat{E},\hat{f}))$, $(\hat{A}\backslash\mathscr{D}(\hat{A},\hat{f})) \cap (\hat{B}\backslash\mathscr{D}(\hat{B},\hat{g})$, and $(\hat{B}\backslash\mathscr{D}(\hat{B},\hat{h})) \cap (B\backslash\mathscr{D}(A,f))$ are non-empty. Hence we conclude that $\delta(A,f) = \delta(\hat{A},\hat{f}) = \delta(\hat{B},\hat{g}) = \delta(B,g)$.

Thus we have shown that for arbitrary $(A,f),(B,g) \in \mathcal{N}$, $\delta(A,f) = \delta(B,g)$ and the proof of claim 4 is complete. □

A.3.6 Proof of Theorem 1.7

Proof. It is clear from classic results that that $\delta(A,f) = 0$ if and only if dynamic consistency holds. Otherwise $\delta \in (0,1]$ and is unique.

Lemma A.11. *If Axiom 1.8 holds, then for every* (A,f), $\mu_{A,f}(\mathscr{D}(A,f)) = 1$.

Proof. First, consider the case when (A,f) is constant. That is, $f(s) \sim f(s')$ for every $s,s' \in A$. Thus by Consequentialism it follows that $\mu_{A,f}(A) = 1$. Since $A = \mathscr{D}(A,f)$ the result holds. Suppose next that f is non-constant on A. Thus there exists some $s,s' \in A$ such that $f(s) \succ f(s')$. Hence $\mathscr{D}(A,f)$ is a strict, non-empty subset of A. Suppose $\mu_{A,f}(\mathscr{D}(A,f)) < 1$. Then choose $\bar{x} \in f(\mathscr{D}(A,f))$ and some $\underline{x} \in f(A\backslash\mathscr{D}(A,f))$, and define g^ϵ by $g^\epsilon(s) = (1-\epsilon)\bar{x} + \epsilon\underline{x}$ for all $s \in A$. Thus for every $\epsilon > 0$ and every $s \in A$, $\bar{x} \succ g^\epsilon(s)$, and hence by axiom 3, $f \succ_{A,f} g^\epsilon$. Since $\mu_{A,f}(\mathscr{D}(A,f)) < 1$ it follows that $u(\bar{x}) > V_{A,f}(f)$. By continuity there is some $\hat{\epsilon} > 0$ such that $u(\bar{x}) > V_{A,f}(g^{\hat{\epsilon}}) > V_{A,f}(f)$, a contradiction. Hence $\mu_{A,f}(\mathscr{D}(A,f)) = 1$. □

Now, from it follows that for every $s,s' \in \mathscr{D}(A,f)$, $\frac{\mu(s)}{\mu(s')} = \frac{\mu_{A,f}(s)}{\mu_{A,f}(s')}$, hence $\mu_{A,f}(s) = \mu(s|\mathscr{D}(A,f))$ for all $s \in \mathscr{D}(A,f)$ and the proof is complete.

□

A.3.7 Proof of Theorem 1.9

Case 2: (Best-case Binary Distortion)

Proof. Since $\succsim^1 = \succsim^2$, it follows that $u^1 = u^2 = u$ and $\mu^1 = \mu^2 = \mu$. Suppose for all (A, f), $f \succsim^1_{A,f} x \Rightarrow f \succsim^2_{A,f} x$, but $\delta^2 < \delta^1$. Let $(A, f) \in \mathcal{N}$ and pick $\bar{x} \sim^1_{A,f} f$. Hence $(1 - \delta^1) \sum_{s \in A} u(f(s))\mu_{|A}(s) + \delta^1 \sum_{s \in A} u(f(s))\mu_{|\mathscr{D}_{A,f}}(s) = u(\bar{x})$. Since $\delta^2 < \delta^1$ and $\sum_{s \in A} u(f(s))\mu_{|\mathscr{D}_{A,f}}(s) > \sum_{s \in A} u(f(s))\mu_{|A}(s)$, it follows that $u(\bar{x}) > (1 - \delta^2) \sum_{s \in A} u(f(s))\mu_{|A}(s) + \delta^2 \sum_{s \in A} u(f(s))\mu_{|\mathscr{D}_{A,f}}(s)$. But this contradicts $f \succsim^2_{A,f} \bar{x}$, thus $\delta^2(A, f) \geq \delta^1(A, f)$ for all (A, f). $\square$

Appendix B

Appendix to Chapter 2

Before proving Theorem 1 I state and prove some basic results.

Lemma B.1. *If* $\{\succsim_A\}_{A\in\Sigma}$ *satisfy Axiom 2.1, Axiom 2.4, Axiom A.4, then for all* $A \in \Sigma$ *such that* A *is non-null, and any* $x, y \in X$*,*

$$x \succsim y \Longleftrightarrow x \succsim_A y$$

Proof. First suppose $x \succsim y$. By monotonicity of $\succsim$ this is equivalent to $xAy \succsim y$ for all A, then by Axiom 2, $x \succsim_A y$. Suppose that $x \succsim_A y$ but $y \succ x$. Then it follows from monotonicity and the fact that A is non-null that $yAx \succ x$, but then again by axiom 2 $y \succ_A x$, which is a contradiction. Hence $x \succsim y$. □

Lemma B.2. *If* $\{\succsim_A\}_{A\in\Sigma}$ *satisfy Axiom 2.1, Axiom 2.4, Axiom A.4, then for all* $A \in \Sigma$ *such that* A *is non-null,* $\succsim_A$ *satisfies monotonicity. I.e.,* $f(\omega) \succsim_A g(\omega)$ *for all* $\omega \in \Omega$ *implies* $f \succsim_A g$*.*

Proof. Suppose $f(\omega) \succsim_A g(\omega)$ for all $\omega \in \Omega$. By lemma 1 we know that $f(\omega) \succsim g(\omega)$ for all $\omega \in \Omega$. Then from monotonicity of $\succsim$ it follows that $fAg \succsim g$, and hence by Axiom 2 it follows that $f \succsim_A g$. □

Lemma B.3. *For each $A \in \Sigma$, A non-null, it follows from Axiom 1 that there is a utility index $u_A : X \to \mathbb{R}$ and probability μ_A such that*

$$f \succsim_A g \Longleftrightarrow \int_\Omega u(f(\omega))\mu_A(d\omega) \geq \int_\Omega u(f(\omega))\mu_A(d\omega).$$

Proof. This follows from standard results. □

Lemma B.4. *For each $A \in \Sigma$, A non-null, and all $f, g \in \mathscr{F}$,*

$$fAg \succsim g \Longleftrightarrow fAh \succsim gAh \text{ for all } h \in \mathscr{F}.$$

Proof.

$$\begin{aligned}
fAg \succsim g \Leftrightarrow & \int_A u(f(\omega))\mu(d\omega) + \int_{\Omega \setminus A} u(g(\omega))\mu(d\omega) \geq \\
& \int_A u(g(\omega))\mu(d\omega) + \int_{\Omega \setminus A} u(g(\omega))\mu(d\omega) \\
\Leftrightarrow & \int_A u(f(\omega))\mu(d\omega) \geq \int_A u(g(\omega))\mu(d\omega) \\
\Leftrightarrow & \int_A u(f(\omega))\mu(d\omega) + \int_{\Omega \setminus A} u(h(\omega))\mu(d\omega) \geq \int_A u(g(\omega))\mu(d\omega) + \int_{\Omega \setminus A} u(h(\omega))\mu(d\omega) \\
\Leftrightarrow & fAh \succsim gAh
\end{aligned}$$

□

B.1 Proof of Theorem 2.1

I now prove Theorem 1. The proof does not rely on assumptions about finiteness of Ω, and hence holds for rather general state spaces. Through the addition of standard axioms we could impose countable additivity of beliefs or convex range of beliefs, though neither of which are needed for the proof. What is essential is the convex structure of X and the existence of at least three non-null events.

Proof. From lemma 1 we know that for every non-null A, u_A is a positive affine transformation of $u_\Omega \equiv u$. Without loss normalize all the utility functions to u. Since X is convex it follows that $u(X)$ is a convex subset of $\mathbb{R}$, so again without loss suppose $[-1,1] \subset u(X)$.

Step 1: Let $\mu \equiv \mu_\Omega$ and for any A such that $\mu(A) > 0$, let $BU(\mu, A)$ denote the Bayesian update of μ conditional on A: $BU(\mu, A)(B) = \frac{\mu(A \cap B)}{\mu(A)}$. Fix some non-null A and define the relation $\trianglerighteq_A$ on $\mathscr{F}$ by $f \trianglerighteq_A g$ if and only if $fAg \succsim g$. Let $\triangleright_A$ and $\approx_A$ be the strict and symmetric parts of $\trianglerighteq_A$. We then establish that $\trianglerighteq_A$ has an expected utility representation (v_A, π_A) where $v_A = u$ and $\pi_A = BU(\mu, A)$[1]. Let $V_{\mathscr{B}(A)}$ denote the functional that represents $\trianglerighteq_A$ and let V_A denote the functional that represents $\succsim_A$, with $V \equiv V_\Omega$.[2] It follows from Axiom 2 that

$$V_A(f) \in [\min\{V(f), V_{\mathscr{B}(A)}(f)\}, \max\{V(f), V_{\mathscr{B}(A)}(f)\}] \tag{B.1}$$

It is worth noting that unlike in Saito [59], here we do not have a convenient ordering between V and $V_{\mathscr{B}(A)}$. Suppose there exists some $\delta \in [0,1]$ such that $V_A(f) = \delta V(f) + (1-\delta)V_{\mathscr{B}(A)}(f)$ for every f. Since every f is finite-valued it is simple to show that since V_A represents $\succsim_A$ and $u = u_A$:

$$\int_\Omega u(f(\omega))\mu_A(d\omega) = \delta \int_\Omega u(f(\omega))\mu(d\omega) + (1-\delta) \int_\Omega u(f(\omega))BU(\mu, A)(d\omega) =$$

$$\int_\Omega u(f(\omega))(\delta\mu + (1-\delta)BU(\mu, A))(d\omega).$$

Thus $\succsim_A$ is represented by $(u, \delta\mu + (1-\delta)BU(\mu, A))$. In the following steps we construct such a δ.

Step 2:

[1] See proof of Theorem 4, Kovach [41]

[2] In every instance the functionals are normalized by the utility index u. This ensures that $V_A(x) = V(x) = V_{\mathscr{B}(A)}(x) = u(x)$ for all $x \in X$.

Note that if $V(f) = V_{\mathscr{B}(A)}(f)$ then by (1), $V_A(f) = \gamma V(f) + (1-\gamma)V_{\mathscr{B}(A)}(f)$ for every γ. Further, if $\succsim_A$ is identically $\succsim$ for A non-null then we may take $\gamma = 1$, while if $\succsim_A$ is identically $\trianglerighteq_A$ (i.e., dynamic consistency holds) then we may take $\delta = 0$. Suppose there is some non-null A and acts f, g such that $f \succsim_A g$ and $g \succ fAg$, or $f \succ_A g$ and $g \succsim fAg$.[3] Let $\mathcal{K}_{\succsim}$ denote the set of all events at which the agent violates dynamic consistency. Also, by (1) we can restrict attention to f such that $V(f) \neq V_{\mathscr{B}(A)}(f)$. For any $A \in \mathcal{K}_{\succsim}$, let

$$\mathscr{F}(A) = \{f \in \mathscr{F} | V(f) \neq V_{\mathscr{B}(A)}(f)\} = \{f \in \mathscr{F} | zAf \sim f \text{ and } z \nsim f \text{ or, } z \sim f \text{ and } zAf \nsim f\}$$

Since we have normalized u across representations, if $V_A(f) \neq V_{\mathscr{B}(A)}(f)$ then $f \in \mathscr{F}(A)$.

Step 3:

Define the function $\delta^A : \mathscr{F}(A) \to [0,1]$ by $\delta^A(f) = \frac{V_A(f) - V_{\mathscr{B}(A)}(f)}{V(f) - V_{\mathscr{B}(A)}(f)}$. Note that since $V_A(f)$ is always between $V(f)$ and $V_{\mathscr{B}(A)}(f)$ the numerator and denominator always have the same sign, hence the ratio is always (weakly) positive. Further, from step 1 it follows that $|V_A(f) - V_{\mathscr{B}(A)}(f)| \leq |V(f) - V_{\mathscr{B}(A)}(f)|$, hence the ratio is always less than 1. Since there is some act f such that $V_A(f) \neq V_{\mathscr{B}(A)}(f)$, it follows that for some f, $\delta^A(f) > 0$.

In the following steps, fix $f, g \in \mathscr{F}(A)$

Step 4:

Suppose $f \sim g$ and $fAg \sim g$. From dominance it follows that $f \sim_A g$ and thus $V(f) = V(g)$ and $V_A(f) = V_A(g)$. Further $fAg \sim g$ implies $f \approx_A g$ and thus $V_{\mathscr{B}(A)}(f) = V_{\mathscr{B}(A)}(g)$. It follows directly that $\delta^A(f) = \delta^A(g)$.

Step 5:

For any $\gamma \in (0,1]$, $g \in \mathscr{F}(A)$ and $x \in X$, it follows immediately from the

[3] Note that this implies that for some f, $V_A(f) \neq V_{\mathscr{B}(A)}(f)$.

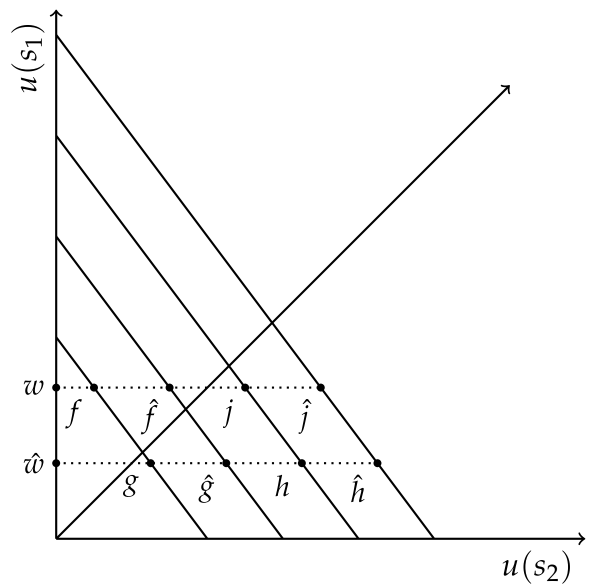

Figure B.1: Acts illustrating various cases, where $A = \{s_1\}$.

definition of δ^A and the linearity of the functionals V, V_A and $V_{\mathscr{B}(A)}$ that $\delta^A(g) = \delta^A(\gamma g + (1-\gamma)x)$

Step 6:

Suppose $f \succ g$ and $fAg \sim g$. Let $w \in X$ satisfy $wAf \sim f$. From lemma 4 $fAg \sim g$ is equivalent to $fAh \sim gAh$ for any h, so we can take $h = f$ and hence $f \sim gAf$, hence $wAf \sim f \sim gAf$. Again by lemma 4 it follows that $wAg \sim g$. Suppose $w \succ f \succ g$. By continuity we can take $\alpha \in (0,1)$ such that $\alpha w + (1-\alpha)g \sim f$. Since $wAf \sim f \sim gAf$ it follows that $[\alpha w + (1-\alpha)g]Af \sim f$, and hence by step 4 and step 5, we conclude that $\delta^A(f) = \delta^A(\alpha w + (1-\alpha)g) = \delta^A(g)$. If $f \succ g \succ w$ the same argument holds by replacing f and g in the previous steps. Suppose $f \succ w \succ g$. Then it must follow that there is some $\omega \in A^c$ such that $f(\omega) = z_f \succ w$, otherwise we violate monotonicity. Then define $h = wAz_f$. It then follows that $h \succ w \succ f \succ g$, and since $wAh \sim h$ by construction, the previous arguments apply, and hence $\delta^A(f) = \delta^A(h) = \delta^A(g)$.

Step 7:

Suppose $f \sim g$ and $fAg \succ g$. Such a case is illustrated by f, g in Figure B.1. We will use a technique similar to the one in step 6. Let w satisfy $wAf \sim f$ and let $\hat{w}$ satisfy $\hat{w}Ag \sim g$. From $f \sim g$ it follows that there is some z_g such that $gAz_g \sim g$ and $z_g \succ w$. We can then define $\hat{g} = \hat{w}Az_g$ and $\hat{f} = wAz_g$. Clearly $z_g \succ \hat{f} \succ \hat{g}$. By the argument in step 6 we can conclude that $\delta^A(\hat{g}) = \delta^A(\hat{f})$. From $\hat{f} \succ f$ and $\hat{f}Af \sim f$, it also follows from 6 that $\delta^A(\hat{f}) = \delta^A(f)$, and hence $\delta^A(g) = \delta^A(f)$.

Step 8:

Suppose $f \succ g$ and $fAg \succ g$. Let $w \in X$ satisfy $w \sim fAw$, it then follows that $wAg \sim fAg$. Without loss suppose there is some z such that $z \succ fAz$.[4] Then by continuity there is some $\alpha \in (0,1)$ such that $\hat{g} = \alpha g + (1-\alpha)z \sim f$. Now, by completeness either $\hat{g}Af \succ f$, $fA\hat{g} \succ \hat{g}$, or $\hat{g}Af \sim f$. In either the first or second case, this is now equivalent to step 7, whereas in the third case we are in step 4, and hence $\delta^A(g) = \delta^A(\hat{g}) = \delta^A(f)$.

Step 9:

Suppose $f \succ g$ and $fAg \prec g$. In this case, as before, it is without loss to suppose there is some $z \in X$ such that $z \succ f$. Hence by continuity we have $\hat{g} = \alpha g + (1-\alpha)z \sim f$ for some $\alpha \in (0,1)$. Hence by the same argument from step 8, $\delta^A(g) = \delta^A(f)$.

Step 10:

Suppose $f \prec g$ and $fAg \succ g$. This case is identical to step 9, by simply relabeling the acts.

By combining all the above steps, and since f and g were arbitrary, we conclude that for all $f, g \in \mathscr{F}(A)$, $\delta^A(f) = \delta^A(g) = \delta^A > 0$. Also, if for all f, g $f \sim g \implies f \sim_A g$, then it must be that $\delta^A = 1$. Otherwise $\delta^A \in (0,1)$. Now while we have only shown that there is some A at which $\delta^A \in (0,1)$. The next steps show that for any other non-null information set $B \in \Sigma$, that $\delta^B = \delta^A$, and hence if the agent violates dynamic consistency at some information set, he must violate it at every

[4]Suppose not, then by non-degeneracy there is some y such that $fAy \succ y$. Then we can take $\hat{f} = \alpha f + (1-\alpha)y$, and by step 5 we may without loss proceed with $\hat{f}$ in place of f.

information set.

Step 11:

Case 1: $A \cup B \neq \Omega$

For any $C \cap (A \cup B) = \varnothing$ choose x, y, z such that $xCy \sim_A z$. By axiom 3 it follows that $xCy \sim_B z$. Hence $\mu_A(C)u(x) + (1-\mu_A(C))u(y) = \delta^A \mu(C)u(x) + (1-\mu_A(C))u(y) = u(z) = \delta^B \mu(C)u(x) + (1-\mu_B(C))u(y) = \mu_B(C)u(x) + (1-\mu_B(C))u(y)$, from which it follows that

$$(\delta^A - \delta^B)\mu(C)u(x) = (\delta^A - \delta^B)\mu(C)u(y).$$

Since x, y are arbitrary, it is without loss to suppose that $u(x) > u(y)$, hence equality is true if an only if $\delta^A = \delta^B$.

Case 2: $A \cup B = \Omega$

The argument is not that different. We simply find $A' \subset A$ and $B' \subset B$ such that A' and B' fall under case 1.[5] We then show that as long as $\mu(A) < 1$ and $\mu(B) < 1$, then A, A' and B, B' also fall under case 1. Hence $\delta^A = \delta^{A'} = \delta^{B'} = \delta^B$. For further details when S is finite, see the proof of theorem 8 in Kovach [41].

□

B.2 Proof of Theorem 2.2

Proof. It is standard that u is unique up to positive affine transformations and μ is unique. Given uniqueness of μ and μ_A, it is trivial that there is a unique δ that satisfies $\mu_A = \delta\mu + (1-\delta)BU(\mu, A)$. □

[5] That we can do this is trivial if μ has convex range. see [25].

B.3 Proof of Theorem 2.3

Proof. $(i) \iff (iv)$: It is well know that (i) and (iv) are equivalent.

$(iv) \iff (ii)$: By non-degeneracy pick three constant acts $x \succ y \succ z$ and some event A with $\mu(A) \in (0,1)$. Normalize $u(x) = 1, u(y) = 0, u(z) = -1$. Define $f = yAz$ and $g = yAx$. Since $f(\omega) = y = g(\omega)$ for all $\omega \in A$, consequentialism holds if and only if $f \sim_A g$. Since $\mu(A^c) > 0$, if $\delta > 0$ it follows that $V_A(g) = \delta\mu(A^c) > 0 > -\delta\mu(A^c) = V_A(f)$. Hence consequentialism holds if and only if $\delta = 0$.

$(iv) \iff (iii)$: This proof is similar to the previous one. Let $x \succ y \succ z$ and suppose $u(x) = 1$ and $u(z) = 0$. Then define $f = xAz$ and $g = y$. Then by Axiom 2.6 it must be that $f \succsim_A g$ for any y. However, $V_A(f) = \delta\mu(A) + (1-\delta) < 1$ for all $\delta > 0$. Since $V_A(g) = u(y)$, we can choose y such that $u(y) > \delta\mu(A) + (1-\delta)$, hence $g \succ_A f$, a contradiction for $\delta > 0$.

□

Appendix C

Appendix to Chapter 3

C.1 Proof of Theorem 3.1

Proof. Necessity is obvious, so only sufficiency is proved. For any $u : X \to \mathbb{R}$, let $K = u(X) \subset \mathbb{R}$. For any $a \in K^{|\Omega|} \subset \mathbb{R}^{|\Omega|}$, it is clear that there exists some $f \in \mathscr{F}$ such that $(u \circ f)(\omega) = a(\omega)$.

Step 1: Axiom 3.1 implies for each $A \in \Sigma$, there exists a closed, convex set of priors $\mathscr{C}_A$ and a non-constant affine function $u_A : X \to \mathbb{R}$ so that $f \succsim_A g$ if and only if $\min_{\pi \in \mathscr{C}_A} \sum_\Omega u_A(f(\omega))d\pi(\omega) \geq \sum_{\pi \in \mathscr{C}_A} \int_\Omega u_A(g(\omega))d\pi(\omega)$. By ordinal preference consistency, we can suppose without loss that for all A, $u_A = u_\Omega$. Further, by strict monotonicity it follows that for all $A \in \Sigma$ and for all $\pi \in \mathscr{C}_A$, $\pi(\omega) > 0$ for all $\omega \in A$.

If $\succsim^*$ satisfies dynamic consistency then the representation holds for $\alpha = 0$. In what follows suppose $\succsim^*$ violates dynamic consistency at some event. Then let $\mathcal{K}_{\succsim^*} = \{A \in \Sigma | f \succsim^*_A g \text{ and } g \succ^* fAg, \text{ or } f \succ^*_A g \text{ and } g \succsim^* fAg\}$ denote the set of events at which $\succsim^*$ violates dynamic consistency. Further, it is clear from [27] that Axiom 3.3 implies that for any $A \in \Sigma$, $\mathscr{C}_A \subset BU(\mathscr{C}, A)$. Hence for every $\mu \in \mathscr{C}_A$ there is some $\pi \in \mathscr{C}$ such that $BU(\pi, A) = \mu$

Step 2: For any objective randomization $\rho \in \Delta(\Omega)$, if $f^{\rho_A} \succsim_A f$ for all $f \in \mathscr{F}$, then there is some $\pi \in \mathscr{C}_A$ such that $\pi = \rho_A$. To see this, suppose $\rho_A \neq \pi$ for all $\pi \in \mathscr{C}_A$. Then since $\mathscr{C}_A$ is closed and convex, by the separating hyperplane theorem there is some $a \in \mathbb{R}^{|\Omega|}$ such that $\sum_{\omega \in A} a(\omega)\rho(\omega) < \min_{\pi \in \mathscr{C}_A} \sum_{\omega \in A} a(\omega)\pi(\omega)$. By certainty independence we can without loss assume that $a \in K^{|\Omega|}$, hence there exists $f \in \mathscr{F}$ so that $u \circ f = a$, hence $\sum_{\omega \in A} u(f(\omega))\rho(\omega) = u(\sum_{\omega \in A} f(\omega)\rho(\omega)) < \min_{\pi \in \mathscr{C}_A} \sum_{\omega \in A} u(f(\omega))\pi(\omega) \Leftrightarrow f^{\rho_A} \prec_A f$. Further, it is clear that if $\pi \in \mathscr{C}_A$, then for any randomization such that $\rho_A(\omega) = \pi(\omega)$ for all $\omega \in A$, $f^{\rho_A} \succsim_A f$ for all $f \in \mathscr{F}$.

Step 3: For every $A \in \mathcal{K}_{\succsim^*}$, there is some ϵ_A so that if $\pi \in \mathscr{C}$ and $\pi(A) \geq \epsilon_A$, then $BU(\pi, A) \in \mathscr{C}_A$. Let $\epsilon_A = \sup\{\pi(A) | \pi \in \mathscr{C} \text{ and } BU(\pi, A) \notin \mathscr{C}_A\}$. Since $A \in \mathcal{K}_{\succsim^*}$, it follows that

$$\min_{\pi \in \mathscr{C}} \pi(A) < \epsilon_A \leq \bar{m}_A.$$

This does not rule out some π so that $\pi(A) < \epsilon_A$ and $BU(\pi, A) \in \mathscr{C}$. However, if this is the case then it must the that there is some π' so that $\pi(A) > \epsilon_A$ and $BU(\pi', A) = BU(\pi, A)$. To see this, suppose to the contrary. Then there is some $\mu \in \mathscr{C}_A$ so that for every $\pi \in \mathscr{C}$ such that $BU(\pi, A) = \mu$, and $\pi(A) < \epsilon_A$. Then consider two A-maximal objective randomizations ρ, ρ' such that $\rho(A) \geq \rho'(A)$, $\rho'_A = \mu$ and for every $f \in \mathscr{F}$, the following hold: $f^{\rho} \succsim f$,$f^{\rho'} \succsim f$, $f^{\rho'_A} \succsim_A f$. By hypothesis, if $\rho'_A = \mu$, $\rho(A) < \epsilon_A$. Since $\bar{m}_A > 0$, $\rho(A) \geq \rho'(A)$ is equivalent to $\frac{\rho(A)}{\bar{m}_A} \geq \frac{\rho'(A)}{\bar{m}_A}$. Thus the requirements of Axiom 3.4, with $A = B$, are satisfied and $f^{\rho_A} \succsim f$, for every f. Hence from step 2, there is some $\pi \in \mathscr{C}_A$ and $\mu \in \mathscr{C}$ such that $\rho_A = \pi = BU(\mu, A)$ and $\rho = \mu$. But then $\epsilon_A \leq \rho'(A)$, a contradiction.

Step 4: We know that for each $A \in \mathcal{K}_{\succsim^*}$ there is an ϵ_A so that $\pi(A) \geq \epsilon_A$ implies $BU(\pi, A) \in \mathscr{C}_A$. Consider any two $A, B \in \mathcal{K}_{\succsim^*}$. The we have both

$$\min_{\pi \in \mathscr{C}} \pi(A) < \epsilon_A \leq \bar{m}_A$$

and

$$\min_{\pi \in \mathscr{C}} \pi(B) < \epsilon_B \leq \bar{m}_B.$$

Now, let ρ be A-maximal and ρ' be B-maximal and $f^{\rho'_B} \succsim_B f$ for every f. Then by Axiom 3.4, if $\rho(A) \geq \rho'(B)\frac{\bar{m}_A}{\bar{m}_B}$, then there is some $\pi \in \mathscr{C}$ such that $\pi = \rho$ and $BU(\pi, A) \in \mathscr{C}_A$, hence

$$\epsilon_A \geq \rho'(B)\frac{\bar{m}_A}{\bar{m}_B}$$

for all ρ' whereby $\rho'_B = \mu$ for some $\mu \in \mathscr{C}_B$ and ρ' is B-maximal. Then this must when $\rho'(B) = \epsilon_B$, and hence

$$\frac{\epsilon_A}{\bar{m}_A} \geq \frac{\epsilon_B}{\bar{m}_B}.$$

Then symmetry implies $\frac{\epsilon_A}{\bar{m}_A} = \frac{\epsilon_B}{\bar{m}_B}$ for all $A, B \in \mathcal{K}_{\succsim^*}$, hence we may define

$$\alpha := \frac{\epsilon_A}{\bar{m}_A}.$$

Step 5: Now, for any $A \in \Sigma$ let

$$\mathscr{C}_A^\alpha = \{BU(\pi, A) | \pi \in \mathscr{C} \text{ and } \pi(A) \geq \alpha \bar{m}_A\}.$$

The final step is to show that if $A \in \Sigma \backslash \mathcal{K}_{\succsim^*}$, then $\mathscr{C}_A = \mathscr{C}_A^\alpha$. Since $A \in \Sigma \backslash \mathcal{K}_{\succsim^*}$, it follows that $\mathscr{C}_A = BU(\mathscr{C}, A)$, and clearly $\mathscr{C}_A^\alpha \subset BU(\mathscr{C}, A)$. Suppose for contradiction that there is some $\mu \in BU(\mathscr{C}, A)$ such that $\mu \notin \mathscr{C}_A^\alpha$. Then let $\pi \in \mathscr{C}$ satisfy $BU(\pi, A) = \mu$. Next, let ρ be an A-maximal randomization so that $\rho_A(\omega) = \mu(\omega)$ for all $\omega \in A$. It then follows that $\rho(A) < \alpha \bar{m}_A$. Next, consider any $B \in \mathcal{K}_{\succsim^*}$. There there is some B-maximal $\hat{\rho}$ so that $\hat{\rho}(B) < \alpha \bar{m}_B$. Since $\mathscr{C}$ is closed and convex, we can without loss take $\hat{\rho}$ so that

$$|\hat{\rho}(B) - \bar{m}_B| < \frac{\bar{m}_A - \rho(A)}{\bar{m}_A} \bar{m}_B.$$

However after some algebra it follows that $\frac{\hat{\rho}(B)}{\bar{m}_B} > \frac{\rho(A)}{\bar{m}_A}$. But, by assumption $f^{\rho} \succsim f$ and $f^{\rho_A} \succsim f$ for every $f \in \mathscr{F}$, and therefore by Axiom 3.4 we require $f^{\hat{\rho}_B} \succsim f$ for every f, which implies that there is some $\hat{\mu} \in \mathscr{C}_B$ with $\hat{\rho}_B(\omega) = \hat{\mu}(\omega)$ for $\omega \in B$. But, since $\hat{\rho}$ is B-maximal and $B \in \mathcal{K}_{\succsim^*}$, it follows that $\hat{\rho}(B) \geq \alpha \bar{m}_B$. This contradicts our assumption that there is some $\mu \in BU(\mathscr{C}, A)$ such that $\mu \notin \mathscr{C}^{\alpha}_A$. Hence $\mathscr{C}_A = \mathscr{C}^{\alpha}_A$. Thus we have shown that for every $A \in \Sigma$,

$$\mathscr{C}^{\alpha}_A = \{BU(\pi, A) | \pi \in \mathscr{C} \text{ and } \pi(A) \geq \alpha \bar{m}_A\} = \mathscr{C}_A,$$

and the proof is complete.

□

C.2 Proof of Theorem 3.3

Proof. Since both agents satisfy the representation and $\succsim^1 = \succsim^2$, we can conclude that $(u_1, \mathscr{C}_1) = (u_2, \mathscr{C}_2) = (u, \mathscr{C})$.

Step 1 $(ii) \Rightarrow (i)$: First, it is trivial that if $\alpha_1 \geq \alpha_2$, then $\mathscr{C}^{\alpha_1}_A \subset \mathscr{C}^{\alpha_2}_A$ for any $A \in \Sigma$. Then suppose $f \succsim^{2*}_A g$. It follows that for every $\pi \in \mathscr{C}^{\alpha_2}_A$, $\sum_{\omega \in A} u(f(\omega))\pi(\omega) \geq \sum_{\omega \in A} u(f(\omega))\pi(\omega)$. Since $\mathscr{C}^{\alpha_1}_A \subset \mathscr{C}^{\alpha_2}_A$, it follows that

$$\sum_{\omega \in A} u(f(\omega))\pi(\omega) \geq \sum_{\omega \in A} u(f(\omega))\pi(\omega)$$

for every $\pi \in \mathscr{C}^{\alpha_1}_A$, hence $f \succsim^{1*}_A g$.

Step 2 $(i) \Rightarrow (ii)$: Since both agents violate $\succsim^{i*}_A$ - dynamic consistency at some A, α_i is unique. Let A be an event at which the both violate DC and suppose that $\alpha_1 < \alpha_2$. Then by the above result $\mathscr{C}^{\alpha_2}_A \subset \mathscr{C}^{\alpha_1}_A$, and the relation is strict. Choose $\mu \in \mathscr{C}^{\alpha_1}_A \backslash \mathscr{C}^{\alpha_2}_A$. Then since $\mathscr{C}^{\alpha_2}_A$ is closed and convex, we can use a separating hyperplane argument to show there is some f for which

$$\min_{\pi \in \mathscr{C}_A^{\alpha_1}} \sum_{\omega \in A} u(f(\omega))\pi(\omega) \leq \sum_{\omega \in A} u(f(\omega))\mu(\omega) < \min_{\pi \in \mathscr{C}_A^{\alpha_2}} \sum_{\omega \in A} u(f(\omega))\pi(\omega).$$

Then, since X is convex we can choose $x \in \mathscr{F}$ such that

$$\min_{\pi \in \mathscr{C}_A^{\alpha_1}} \sum_{\omega \in A} u(f(\omega))\pi(\omega) < u(x) < \min_{\pi \in \mathscr{C}_A^{\alpha_2}} \sum_{\omega \in A} u(f(\omega))\pi(\omega),$$

from which it follows that

$$f \succsim^{2^*} x \text{ but } f \bowtie^{1^*} x,$$

hence $\succsim^{1^*}_A$ is not more complete than $\succsim^{2^*}_A$. Therefore if $\alpha_1 \geq \alpha_2$ it follows that $\succsim^{1^*}_A$ is more complete than $\succsim^{2^*}_A$.

□

Bibliography

[1] Anup Agrawal and Mark a. Chen. Analyst conflicts and research quality. *Quarterly Journal of Finance*, 2012.

[2] George A. Akerlof and William T. Dickens. The economic consequences of cognitive dissonance. *The American Economic Review*, 1982.

[3] James Andreoni and Tymofiy Mylovanov. Diverging opinions. *American Economic Journal: Microeconomics*, 2012.

[4] F J Anscombe and Robert J Aumann. A definition of subjective probability. *The Annals of Mathematical Statistics*, 1963.

[5] Elliot Aronson. The theory of cognitive dissonance: A current perspective. In Leonard Berkowitz, editor, *Advances in Experimental Social Psychology*. Academic Press, 1969.

[6] Kenneth Arrow and Leonid Hurwicz. An optimality criterion for decision making under ignorance. In C.F. Carter and J.L. Ford, editors, *Undertainty and Expectations in Economics*. Basil Blackwell, 1972.

[7] Nicholas Barberis, Ming Huang, and Richard H. Thaler. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *American Economic Review*, 96(4):1069–1090, 2006. doi:

10.1257/aer.96.4.1069. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.96.4.1069.

[8] Roland Bénabou. Groupthink: Collective delusions in organizations and markets. *The Review of Economic Studies*, 80(2):429–462, 2013. doi: 10.1093/restud/rds030. URL http://restud.oxfordjournals.org/content/80/2/429.abstract.

[9] Roland Bénabou and Jean Tirole. Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics*, 126(2):805–855, 2011. doi: 10.1093/qje/qjr002. URL http://qje.oxfordjournals.org/content/126/2/805.abstract.

[10] Markus K. Brunnermeier and Jonathan A. Parker. Optimal expectations. *The American Economic Review*, 2005.

[11] Colin Camerer. Individual decision making. In Alvin Roth and John Kagel, editors, *Handbook of Experimental Economics*. Princeton University Press, 1995.

[12] Andrew Caplin and John Leahy. Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics*, 116(1):55–79, 2001. doi: 10.1162/003355301556347. URL http://qje.oxfordjournals.org/content/116/1/55.abstract.

[13] John C. Easterwood and Stacey R. Nutt. Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism? *The Journal of Finance*, 1999.

[14] Mahmoud A. El-Gamal and David M. Grether. Are people bayesian? uncovering behavioral strategies. *Journal of the American Statistical Association*,

90(432):pp. 1137–1145, 1995. ISSN 01621459. URL http://www.jstor.org/stable/2291506.

[15] Daniel Ellsberg. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 1961.

[16] Larry Epstein. An axiomatic model of non-bayesian updating. *Review of Economic Studies*, 2006.

[17] Larry G Epstein and Michel Le Breton. Dynamically consistent beliefs must be bayesian. *Journal of Economic Theory*, 1993.

[18] Larry G. Epstein and Igor Kopylov. Cold feet. *Theoretical Economics*, 2007.

[19] Larry G. Epstein and Martin Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113(1):1–31, November 2003.

[20] Larry G Epstein and Martin Schneider. Learning under ambiguity. *Review of Economic Studies*, 2007.

[21] Larry G. Epstein, Jawwad Noor, and Alvaro Sandroni. Non-bayesian updating: A theoretical framework. *Theoretical Economics*, 2008.

[22] Leon Festinger. *A Theory of Cognitive Dissonance*. Stanford University Press, 1957.

[23] Leon Festinger and James M. Carlsmith. Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 1959.

[24] Mark Fey, Richard D. McKelvey, and Thomas R. Palfrey. An experimental study of the constant-sum centipede games. *International Journal of Game Theory*, 1996.

[25] Paolo Ghirardato. Revisiting savage in a conditional world. *Economic Theory*, 2002.

[26] Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118(2):133 – 173, 2004.

[27] Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci. Revealed ambiguity and its consequences: Updating. 42:3–18, 2008.

[28] Itzhak Gilboa and Massimo Marinacci. Ambiguity and the bayesian paradigm. *Advances in Economics and Econometrics: Theory and Applications*, 2011.

[29] Itzhak Gilboa and David Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 1989.

[30] Itzhak Gilboa and David Schmeidler. Updating ambiguous beliefs. *Journal of Economic Theory*, 1993.

[31] Christian Gollier. Optimal illusions and the simplification of beliefs, 2011.

[32] Faruk Gul and Wolfgang Pesendorfer. Temptation and self-control. *Econometrica*, 2001.

[33] Anna Gumen and Andrei Savochkin. Dynamically stable preferences. *Journal of Economic Theory*, 148(4):1487 – 1508, 2013. ISSN 0022-0531. doi: http://dx.doi.org/10.1016/j.jet.2013.04.018. URL http://www.sciencedirect.com/science/article/pii/S0022053113000847.

[34] Martin Hilbert. Toward a synthesis of cognitive biases: How noisy information processing can bias human decision making. *Psychological Bulletin*, 2012.

[35] Jean-Yves Jaffray. Application of linear utility theory to belief functions. In *Uncertainty and Intelligent Systems*. Springer, 1988.

[36] Daniel Kahneman and Amos Tversky. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 1972.

[37] Ethan Kaplan and Sharun Mukand. The persistence of political partisanship: Evidence from 9/11. *Working*, 2014.

[38] John Maynard Keynes. *A Treatise on Probability*. MacMillan and Co, 1921.

[39] Frank Knight. *Risk, Uncertainty, and Profit*. Houghton Mifflin, 1921.

[40] Igor Kopylov. Simple axioms for countably additive subjective probability. *Journal of Mathematical Economics*, 2010.

[41] Matthew Kovach. A model of cognitive dissonance. *Working Paper*, 2014.

[42] Arie W. Kruglanski. Motivation, cognition, and reality: Three memos for the next generation of research. *Psychological Inquiry*, 1999.

[43] Camelia M. Kuhnen and Brian Knutson. The influence of affect on beliefs, preferences, and financial decisions. *Journal of Financial and Quantitative Analysis*, 2011.

[44] Ziva Kunda. The case for motivated reasoning. *Psychological Bulletin*, 1990.

[45] Charles G. Lord, Lee Ross, and Mark R. Lepper. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 1979.

[46] Guy Mayraz. Wishful thinking, 2011. URL http://dx.doi.org/10.2139/ssrn.1955644.

[47] Rajnish Mehra and Edward C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 1985.

[48] Markus M. Mobius, Muriel Niederle, Paul Niehaus, and Tanya S. Rosenblat. Managing self-confidence. *Working Paper*, 2014.

[49] Sendhil Mullainathan and Ebonya Washington. Sticking with your vote: Cognitive dissonance and political attitudes. *American Economic Journal: Applied Economics*, 2009.

[50] Sendhil Mullainathan, Joshua Schwartzstein, and Andrei Shleifer. Coarse thinking and persuasion. *Quarterly Journal of Economics*, 2008.

[51] Efe A. Ok, Pietro Ortoleva, and Gil Riella. Incomplete preferences under uncertainty: Indecisiveness in beliefs versus tastes. *Econometrica*, 2012.

[52] Pietro Ortoleva. Modeling the change of paradigm: Non-bayesian reactions to unexpected news. *The American Economic Review*, 2012.

[53] Thomas R. Palfrey and Stephanie W. Wang. Speculative overpricing in asset markets with information flows. *Econometrica*, 2012.

[54] Cesaltina P. Pires. A rule for updating ambiguous beliefs. *Theory and Decision*, 33:137–152, 2002.

[55] Matthew Rabin and Joel L. Schrag. First impressions matter: A model of confirmatory bias. *Quarterly Journal of Economics*, 1999.

[56] Matthew Rabin and Dimitri Vayanos. The gambler's and hot-hand fallacies: Theory and applications. *Review of Economic Studies*, 2010.

[57] David P. Redlawsk. Hot cognition or cool consideration? testing the effects of motivated reasoning on political decision making. *The Journal of Politics*, 2002.

[58] Kota Saito. Social preferences under uncertainty. *American Economic Review*, 2013.

[59] Kota Saito. Preference for flexibility and preference for randomization under ambiguity. *Working Paper*, 2014.

[60] Leonard J. Savage. *The Foundations of Statistics*. John Wiley & Sons, 1954.

[61] David Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 1989.

[62] Charles S. Taber and Milton Lodge. Motivated skepticism in the evaluation of political beliefs. *American Journal of Pollitical Science*, 2006.

[63] Leeat Yariv. Believe and let believe: Axiomatic foundations for belief dependent utility functionals, 2001.

[64] Leeat Yariv. I'll see it when i believe it - a simple model of cognitive consistency, January 2005.