Rochester Center for Economic Research

Cognitive Dissonance and Choice

Epstein, Larry G., and Igor Kopylov

Working Paper No. 525
March 2006

UNIVERSITY OF
ROCHESTER

COGNITIVE DISSONANCE AND CHOICE*

Larry G. Epstein   Igor Kopylov

February 2006

Abstract

People like to feel good about past decisions. This paper models self-justification of past decisions. The model is axiomatic: axioms are defined on preference over ex ante actions (modeled formally by menus). The representation of preference admits the interpretation that the agent adjusts beliefs after taking an action so as to be more optimistic about its possible consequences. In particular, the ex post choice of beliefs is part of the representation of preference and not a primitive assumption. Behavioral characterizations are given to the comparisons "1 exhibits more dissonance than 2" and "1 is more self-justifying than 2."

1. INTRODUCTION

1.1. Objective

There is considerable evidence in psychology that people like to view themselves as being smart, and in particular, as having made correct decisions in the past. Thus they may change beliefs *after* taking an action and become more optimistic about its possible consequences, in order to feel better about having chosen it. Such behavior is a special case of an affinity for cognitive consistency - for example, an affinity for consistency among beliefs or opinions (Festinger [13]). Here the two

*Epstein is at Department of Economics, University of Rochester, Rochester, NY 14627, lepn@troi.cc.rochester.edu; Kopylov is at Department of Economics, UC Irvine, Irvine, CA 92697, ikopylov@uci.edu. We would like to acknowledge comments by seminar participants at BU, and helpful conversations with Emanuela Cardia, Faruk Gul, Jawwad Noor, Wolfgang Pesendorfer and especially Massimo Marinacci.

cognitions are "I have taken an action that could lead to unfavorable outcomes" and "I am a smart person who would not make poor choices"; adopting a more optimistic belief about future outcomes serves to reduce this dissonance. Though the term cognitive dissonance is often used more broadly, we use it here to refer to ex post self-justification of past actions. Our objective is to model an agent who exhibits such cognitive dissonance.

Models of cognitive dissonance in economics treat beliefs as choice variables, on a par with other more standard choice variables, such as consumption and savings. Thus Akerlof and Dickens [1, p. 307] propose as basic propositions of their model of cognitive dissonance that preference is defined over beliefs and that beliefs are subject to choice. While a more optimistic outlook makes one feel better about the past decision, the agent recognizes that adopting more optimistic beliefs would take her further from the "truth" and thus would lead to suboptimal choices in decisions still to be made. The optimal belief is determined by making this trade-off. Similarly, in Rabin [28], utility depends directly on beliefs. This modeling approach is nonstandard in economics and may make one uncomfortable for a number of reasons. First, it begs the question "what is the feasible set from which beliefs are chosen?" Unlike other choice variables for which the market determines feasible (budget) sets, the feasible set of beliefs is presumably subjective (in the mind of the agent) and thus invariably requires an ad hoc specification. A possibly more important concern is the observability of chosen beliefs and hence testability of the model. While in psychology it is standard to take beliefs as observable through interviews or questionnaires, many economists adhere to the choice-theoretic approach to beliefs, initiated by Savage, whereby beliefs are observable only indirectly through choices among actions.

In this paper, we propose a choice-theoretic and axiomatic model of cognitive dissonance. Preferences are defined over actions (modeled formally by menus) and axioms are imposed on these preferences. Thus empirical testability relies only on the ranking of actions being observable. The functional form for utility admits an interpretation whereby the agent behaves *as if* she chooses beliefs ex post in the manner described above, but this is a *result* - part of the representation of preferences over actions. Finally, the above question about the feasible set of beliefs is answered automatically by the representation.

We emphasize that our agent is not boundedly rational or myopic. Rather she is sophisticated and forward-looking - when choosing an action ex ante she is fully aware that she will later experience cognitive dissonance and that this will affect her later decisions. She has this sophistication in common with agents in most economic models, but one may wonder whether individuals outside those models are typically self-aware to this degree. We are not familiar with definitive evidence on this question and in its absence, we are inclined to feel that full self-awareness is a plausible working hypothesis. Even where the opposite extreme of complete naivete seems descriptively more accurate, our model may help to clarify which economic consequences are due to cognitive dissonance per se and which are due to naivete. In addition, the assumption of sophistication is vital for a choice-theoretic approach: because she anticipates her cognitive dissonance, it affects her current choice of actions. This makes it possible to infer cognitive dissonance from her (in principle observable) choice of actions, consistent with the choice-theoretic tradition of Savage. Thus sophistication seems justifiable also on the methodological grounds of permitting the exploration of modest departures from standard models.

1.2. Model Outline

As described above, cognitive dissonance implies changing beliefs, hence changing preferences, which poses difficulties for modeling behavior. One possible modeling route is to specify dynamically inconsistent preferences and then to tackle the questions of to what degree the agent anticipates future changes in preference and how intrapersonal conflicts are resolved. These are the issues familiar from Strotz [33]. We follow instead the route advocated by Gul and Pesendorfer [16] (henceforth GP) whereby behavior that indicates changing preferences over underlying alternatives can alternatively be viewed as coming from stable preferences over menus of these alternatives.

A brief outline is as follows: uncertainty is represented by the (finite) state space S. Time varies over three periods. The true state is realized and payoffs are received at the terminal time. The intermediate time is called the *ex post* stage. Physical actions chosen then are identified with Anscombe-Aumann acts, maps from S into lotteries over consumption. A physical action is chosen also at the initial *ex ante* stage. Each such action is modeled by a menu of acts - the idea is that any action taken ex ante limits options ex post. The agent understands when choosing a menu that ex post she will choose an act from that menu. She also knows, when ranking menus, that her beliefs about S will change ex post so as to make the previously chosen menu seem more attractive. She will try to resist the temptation because she views her prior beliefs, formed with the detachment afforded by the ex ante stage, as being ‘correct.’ Whether or not she succeeds in exerting self-control, however, temptation is costly, and this affects her ranking of menus. Thus the latter reveals her expected change in beliefs, or her cognitive dissonance.

As a concrete illustration of the relevance of choice of menus and the behavioral manifestation of cognitive dissonance, consider a job choice model along the lines of Akerlof and Dickens [1]. Ex ante the worker chooses a job, either in a hazardous industry or in a safe one. If he chooses the hazardous industry, then ex post he can select between two kinds of safety equipment (high quality h and low quality ℓ). Each kind affects the likelihood of an accident but may not eliminate the risk entirely. Thus h and ℓ each imply a random payoff, net of cost of the equipment, that depends on the exogenous state of the world. In other words, each can be viewed as an act and the job corresponds to the menu $\{h, \ell\}$. For the safe industry, there are no choices to be made ex post and the ultimate payoff is certain and given by c (a constant act). Therefore, the safe industry corresponds to the singleton menu $\{c\}$ and the choice of job corresponds to the choice between the menus $\{h, \ell\}$ and $\{c\}$.

If the worker can commit to safety equipment at the same time that he chooses the job, then ex ante beliefs are such that he would prefer the high quality equipment, that is,

$$\{h\} \succ \{\ell\}. \tag{1.1}$$

In the standard model, menus are valued according to the best alternative that they contain, and thus the worker would also exhibit the indifference

$$\{h\} \sim \{h, \ell\}.$$

However, an agent who exhibits cognitive dissonance, and knows this ex ante, may exhibit the ranking

$$\{h\} \succ \{h, \ell\}.$$

The intuition is as follows: after accepting the job in the hazardous industry, the worker faces the two cognitions - "my job is dangerous" and "I am a smart person and would not choose a precarious job". He relieves this dissonance, and reduces doubts about his job choice, by changing his prior beliefs, as reflected in the ex ante ranking (1.1), and believing instead that the job is not so dangerous after all. This creates the temptation to choose ℓ rather than h. The worker anticipates this temptation. Accordingly, if he dislikes temptation, he would rank $\{h, \ell\}$ as strictly worse than $\{h\}$.

If commitment to high quality equipment can be made simultaneously with choice of the hazardous job, the worker would so commit. Because that would leave no decisions left to be made ex post, cognitive dissonance would not be behaviorally relevant. Assume that such commitment is not possible (Akerlof and Dickens give reasons why commitment may not be possible). Then there remains the question of whether given the menu $\{h, \ell\}$ ex post, he yields to the temptation and chooses ℓ. He feels that his prior beliefs were "correct" and thus "should" be used to guide decisions - in other words, h is the correct choice. The balance between what he ought to do and the tempting alternative depends on the worker's self-control. With high self-control, he may resist the temptation and choose h. Following GP, we suppose that this case (or rather the ex ante expectation thereof) is captured by the ranking

$$\{h, \ell\} \succ \{\ell\}.$$

The expectation of yielding and choosing ℓ is captured by the ranking

$$\{h, \ell\} \sim \{\ell\}. \tag{1.2}$$

Rational expectations about cognitive dissonance may lead to choice of the safe industry. However, if

$$\{h, \ell\} \succ \{c\},$$

then the worker chooses the hazardous industry and, assuming (1.2), later adopts the poor safety equipment corresponding to ℓ. To an outsider, or from the perspective of (1.1), the worker may appear careless or overconfident.

To this point, we have suggested that cognitive dissonance could explain the ranking

$$\{h\} \succ \{h, \ell\} \succeq \{\ell\}. \tag{1.3}$$

This ranking is a special case of GP's central axiom of Set-Betweenness. While GP argue that such rankings reveal the presence of temptation and self-control problems, the reason for temptation is unspecified. Put another way, the ranking under commitment may conflict with choice behavior out of the menu available ex post, but the reason for this difference is not clear given only (1.3). For example, (1.3) could be due to underlying preferences (taste or risk aversion) changing with the passage of time, rather than beliefs changing in order to justify the previous choice of menu. But there is other behavior, described via axioms in our formal model, that would exclude such interpretations and support an interpretation in terms of cognitive dissonance.

1.3. Related Literature

It has been argued that a moderate degree of (optimistic) illusion can be psychologically beneficial even net of the loss in efficacy of decisions; see Taylor and Brown [35], Taylor [34] and Baumeister [4], for example.

The psychological theory of cognitive dissonance is due to Festinger [13]. Dissonance originates with an action and the subsequent evaluation of that action. Where there exists dissonance between having taken that action and subsequent beliefs, the theory posits that those beliefs will be changed to match or justify the past action. Aronson [3] is an excellent textbook treatment and overview of the supporting evidence from psychology. Some of this evidence is strongly suggestive that cognitive dissonance has economic consequences; for example, the efficacy of the "foot-in-the-door-technique", whereby a small commitment by individuals makes it easier to persuade them later to commit further in that direction, suggests the efficacy of two-stage mechanisms, possibly including an entry fee at the first stage. Several other applications have been developed in formal economic models as we describe below.

Akerlof and Dickens suggest that cognitive dissonance can play a role in explaining some economic phenomena that are arguably puzzling from the perspective of more standard models. These include the existence of safety regulation (based on the job-choice model sketched above), why noninformational advertising is effective (it gives external justification for an individual to believe she is making a smart decision in buying the product), and why persons often fail to purchase actuarially favorable disaster (flood or earthquake) insurance. The story here is analogous to that concerning safety equipment and fits naturally into our modeling approach: after choosing a house (or menu), it reduces dissonance to believe that a flood is so unlikely as to not justify buying insurance (the choice of a particular act), even though she would have bought insurance simultaneously with the house purchase. Similarly, cognitive dissonance can explain why researchers may appear "overly optimistic" in their pursuit of a previously chosen project (a menu). It feels good to believe that the research project previously embarked on is a promising one and thus ongoing efforts may be guided by otherwise unwarranted optimism.

Rabin [28] models the choice of an enjoyable but immoral activity in light of dissonance between one's beliefs about what is moral and the chosen level of activity. Haagsma and Koning [17] show how cognitive dissonance can generate barriers to exiting an unproductive industry. Smith [32] shows how cognitive dissonance can explain why wages tend to rise faster than productivity. The worker justifies his job situation ex post by adjusting his beliefs about the cost of effort needed to fulfill his duties - the need for self-justification, and the adjustment in beliefs, are greater the lower is his past wage. The employer can exploit this by offering a contract with an increasing wage profile. Goetzmann and Peles [15] argue that cognitive dissonance leads investors to justify remaining in mutual funds that consistently perform poorly; and that such inertia can help to explain why money flows in more rapidly to mutual funds that have performed well than flows out from those that have performed poorly. See Dickens [8] and Oxoby [27] for further applications of cognitive dissonance.

With regard to modeling, we have already acknowledged our debt to GP. Their model does not apply directly, however. One difference is that while they study preferences over menus of lotteries, it is important for our story that menus consist of (Anscombe-Aumann) acts. Kopylov [19] has extended the GP theorem from (menus of) lotteries to abstract mixture spaces, including, in particular, the space of Anscombe-Aumann acts. A second and more important formal difference from GP, and also from Kopylov's extension, and the major source of technical difficulty in our model, is that we relax the Independence axiom - the latter is not intuitive given cognitive dissonance. Finally, we note that Dekel, Lipman and Rustichini [7] generalize GP's model of temptation. However, their motivation is much different than ours - in particular, they assume Independence.[1]

There are two final but important connections to the literature. The more optimistic beliefs held ex post by our agent come about in our model because she uses a (nonsingleton) *set* of probability measures, and when evaluating a prospect, she chooses the measure that maximizes its utility. This recalls Dreze's [9] model of choice between Anscombe-Aumann acts under moral hazard. It recalls also Gilboa and Schmeidler [14] - they model agents who are averse to ambiguity, in the sense illustrated by the Ellsberg Paradox, by assuming that they minimize (rather than maximize) over a set of priors, but their model has an obvious counterpart for ambiguity loving. Our model differs from both of these primarily through its focus on the time-varying nature of beliefs and the corresponding value of

[1] In their concluding remarks about possible directions for further research, they mention that accommodating guilt may be a reason for relaxing Independence when modeling temptation. This rationale is obviously much different than ours. There exist other representation results in the menus-of-lotteries/acts setting that do not rely on Independence. Epstein and Marinacci [11] study an agent who is not subject to temptation, but rather values flexibility because she is uncertain about the future; she violates Independence because her conception of the future is coarse. More recent results, with still different objectives, appear in Ergin and Sarver [12] and in Noor [26].

commitment. In principle, one could reinterpret our model in terms of a change from ex ante probabilistic beliefs to ex post multiple-priors reflecting ambiguity loving, but then there is no apparent reason for the agent to exert self-control as she does in our model. Thus we disregard this interpretation of the model.

2. UTILITY

The model has the following primitives:

- time $t = 0, 1, 2$
- finite state space S
- $\mathcal{C}$: set of (Borel) probability measures over a compact metric space
  refer to $c \in \mathcal{C}$ as a lottery over consumption, or more briefly as consumption
  $\mathcal{C}$ is compact metric under the weak convergence topology
- $\mathcal{H}$: set of acts $h : S \longrightarrow \mathcal{C}$, with the usual mixture operation
- compact sets of acts are called *menus* and denoted A, B, ...
  $\mathcal{K}(\mathcal{H})$ is the set of all menus
  it is compact metric under the Hausdorff metric[2]
- preference $\succeq$ defined on $\mathcal{K}(\mathcal{H})$

The interpretation is that a menu A is chosen *ex ante* (at time 0) according to $\succeq$. This choice is made with the understanding that at the unmodeled *ex post* stage (time 1), the agent will choose an act from A. Uncertainty is resolved and consumption is realized in the terminal period $t = 2$. Cognitive dissonance and choice behavior at time 1 are anticipated ex ante and underlie the ranking $\succeq$.

Menus are natural objects of choice.[3] The consequence of a physical action taken at time 0 is that it determines a feasible set of physical actions at time 1, and these actions can be modeled by acts in the usual way. Thus each physical action at time 0 corresponds to a menu of acts.

[2] See [2, Theorem 3.58], for example.

[3] Kreps [21, 23] was the first to propose menus as a way to model physical actions in an ex ante stage.

Our model of utility has the form[4]

$$\mathcal{U}(A) = \max_{h \in A} \left[(1-\kappa) U(h) + \kappa V(h)\right] - \kappa \max_{h' \in A} V(h'), \tag{2.1}$$

where

$$U(h) = \ p \cdot u(h), \text{ and} \tag{2.2}$$

$$V(h) = \max_{q \in Q} \ q \cdot u(h). \tag{2.3}$$

Here $0 \leq \kappa \leq 1$, p is a probability measure on S, Q is a closed and convex set of probability measures on S containing p, and $u : \mathcal{C} \longrightarrow \mathbb{R}^1$ is mixture linear and continuous.

The standard model of subjective expected utility maximization is the special case where $\kappa = 0$ or $Q = \{p\}$. More generally, the functional form can be interpreted along the lines suggested by GP. When restricted to singletons, $\mathcal{U}$ coincides (ordinally) with U; thus expected utility with prior p represents preference over consumption lotteries when the agent can commit ex ante. When she does not commit, then the new (temptation) utility function V over lotteries becomes relevant. Temptation utility is computed by maximizing over probability measures in the set Q. Since $p \in Q$, V imputes higher expected utility to the menu at hand than was the case ex ante using p, corresponding to cognitive dissonance. She is tempted to maximize V ex post. Though she views p as "correct", there is a self-control cost of resisting the temptation given by

$$\kappa \left(V(h) - \max_{h' \in A} V(h') \right) \leq 0.$$

Thus a compromise is struck between maximizing U and maximizing V - choice out of A is described by maximization of the weighted sum, or by solving

$$\max_{h \in A} \ \max_{q \in (1-\kappa)\{p\} + \kappa Q} q \cdot u(h). \tag{2.4}$$

which balances ex ante realism and ex post cognitive dissonance. The nature of the compromise is further illustrated by the fact that

$$p \in (1-\kappa)\{p\} + \kappa Q \subseteq Q,$$

[4]For any real-valued random variable x on S, and probability measure q, $q \cdot x$ is short-hand for the expected value $\int_S x dq$.

so that the set of beliefs underlying the choice of an act ex post lies "between" the prior view p and the optimistic view represented by Q.[5]

Since ex post choice out of the menu maximizes the utility function $\max_{q\in(1-\kappa)\{p\}+\kappa Q} q\cdot u(h)$, which does not depend on the menu, one may wonder whether the model captures beliefs that adjust to make the previously chosen menu attractive. To see a sense in which this is true, note that (by reversing the order of the maximizations),

$$\max_{h\in A}\max_{q\in(1-\kappa)\{p\}+\kappa Q} q\cdot u(h) = \max_{h\in A} q_A^*\cdot u(h),$$

for any q_A^* that solves $\max_{q\in(1-\kappa)\{p\}+\kappa Q}\max_{h\in A} q\cdot u(h)$. Thus ex post choice conforms with SEU and probabilistic beliefs given by q_A^*. Evidently, q_A^* depends on the menu A and is chosen to make the value of the menu, given by $\max_{h\in A} q\cdot u(h)$, as large as possible.

We can say something about the qualitative difference between ex post choice and the "correct" choice. Given any menu A ex post, the choice out of A is determined by maximizing $\max_{q\in(1-\kappa)\{p\}+\kappa Q} q\cdot u(h)$, while the "correct" choice would maximize $p\cdot u(h)$. Suppose for concreteness that consumption is real-valued ($\mathcal{C}$ consists of lotteries over $[a,b]\subset\mathbb{R}^1$) - typically, one assumes that $u(\cdot)$ is concave on $[a,b]$ corresponding to risk aversion. On the other hand, the maximization over q in the ex post utility function introduces some convexity. Thus it may not be concave in h and may even exhibit risk loving. (For example, if one restricts attention to Savage acts $h: S\to[a,b]$, then ex post utility is a convex function of h if u is linear.) Consequently, ex post choice may appear extreme - for example, it may correspond more to boundary optima.

A final comment is that both subjective and objective probabilities are present in the model - the latter underlie consumption lotteries - but they are treated differently: while the agent chooses new beliefs ex post about her subjective uncertainty (the state space S), she does not distort or modify objective probabilities. For example, both U and V agree about the ranking of lotteries in that, for every lottery c, $U(c)=V(c)=u(c)$, the vNM expected utility of c. Because an objective probability law is based on undeniable fact, distorting it to a more favorable one, is folly or ignorance that would not be undertaken by the sophisticated individuals that we model. But where facts alone do not pin down beliefs uniquely,

[5] As is familiar from GP-style models, this interpretation in terms of ex post choice is suggested by the functional form, and by intuition for the underlying axioms, but ex post choice lies outside the scope of our formal model. See Noor [25] for a model of temptation where ex post choice is part of the primitives.

an agent is free to choose beliefs and feeling good about oneself is one possible consideration in doing so. As an illustration of the difference, note that Knox and Inkster [18] report that persons leaving the betting window after placing bets at a race-track are more optimistic about "their horse" than persons about to place bets. On the other hand, it is more difficult to imagine someone being similarly optimistic about a coin, which is known to be unbiased, after choosing that coin for a game of chance.

3. AXIOMS

The first two axioms require no discussion.

Axiom 1 (Order). $\succeq$ *is complete and transitive.*

Axiom 2 (Continuity). $\succeq$ *is continuous.*

Menus can be mixed via

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

Formally, the indicated mixture of A and B is another menu and thus when the agent contemplates that menu ex ante, she anticipates choosing out of $\alpha A + (1-\alpha) B$ ex post. It follows that one should think of the randomization corresponding to the α and $(1-\alpha)$ weights as taking place at the end - after she has chosen some mixed act $\alpha f + (1-\alpha) g$ out of the menu. In fact, since the mixture of acts is defined by $(\alpha f + (1-\alpha) g)(s) = \alpha f(s) + (1-\alpha) g(s)$ for each s, the randomization occurs after realization of the state.

The above mixture operation permits one to state the Independence axiom, which is adopted by GP. However, Independence is *not* intuitive under cognitive dissonance.[6] To see this, suppose for concreteness that $A \sim B$ and consider whether the mixture $\alpha A + (1-\alpha) B$ should also be indifferent to A as required by Independence. Indifference between A and B is based on the anticipation that, in each case, the agent will choose beliefs ex post to make the menu in hand attractive, and that these beliefs will tempt her to choose out of the given

[6] The reason is essentially that because the agent anticipates that she will adjust her beliefs ex post to the menu at hand, the situation is analogous to that of choice between "temporal risks". As explained by Machina [24], for example, preferences over temporal risks typically violate Independence even at a normative level.

menu differently from what she would have prescribed ex ante. Evaluation of the mixture $\alpha A + (1-\alpha) B$ can be thought of similarly, but the important point is that beliefs for the mixed menu must be chosen before the randomization is played out. Since also *beliefs chosen given A generally differ from those chosen given B*, optimistic beliefs for the mixed menu bear no simple relation to those for A and B. A similar disconnect applies to anticipated temptation and ex post choices across the three menus. For example, it is possible that the acts f and g be chosen out of A and B respectively, while $\alpha f + (1-\alpha) g$ not be chosen out of $\alpha A + (1-\alpha) B$. As a result, the agent will generally not be indifferent between the mixed menu and A, violating Independence. (The deviation from indifference could go in either direction: $\alpha A + (1-\alpha) B \succ A$ and $\alpha A + (1-\alpha) B \prec A$ are both possible.)

However, suitable relaxations of Independence are intuitive. To proceed, for any act $f \in \mathcal{H}$, let

$$\mathcal{H}_f = \{tc + (1-t)f : t \in [0,1],\ c \in \mathcal{C}\}.$$

If $h = tc + (1-t)f$ is an act in $\mathcal{H}_f$, then for any mixture linear u and for all probability measures q,

$$q \cdot u(h) = tu(c) + (1-t)\, q \cdot u(f).$$

Because the first term on the right is independent of q, it follows that any menu A that is a subset of $\mathcal{H}_f$ is rendered attractive by beliefs that make f attractive. In particular, for any two menus A and B in $\mathcal{H}_f$, when the agent chooses beliefs to fit the menu, there is an optimistic measure that is common to both A and B. But this invalidates the reason given above for violating Independence. Thus we adopt:

Axiom 3 (Collinear Independence). *For all $\alpha \in (0,1)$, for all $f \in \mathcal{H}$, and for all menus $A', A, B \subseteq \mathcal{H}_f$,*

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

Acts h' and h in $\mathcal{H}_f$ are naturally called *collinear*, which explains the name of the axiom.[7]

[7] For any collinear acts h' and h, it is easy to see that for every s' and s, $u(h'(s')) > u(h'(s)) \Longrightarrow u(h(s')) \geq u(h(s))$, that is, the real-valued functions $u(h(\cdot))$ and $u(h'(\cdot))$ are comonotonic. Collinearity implies the stronger restriction $(1-t)(u(h'(s')) - u(h'(s))) = (1-t')(u(h(s')) - u(h(s)))$ for some t and t'. Thus collinearity can be viewed as a cardinal counterpart of comonotonicity.

When ranking singleton menus, there is no choice to be made ex post. Thus cognitive dissonance is not relevant and there is no reason for Independence to be violated. This motivates the following second relaxation of Independence:

Axiom 4 (Commitment Independence). *For all $f, g, h \in \mathcal{H}$ and $\alpha \in (0,1)$,*

$$\{f\} \succ \{g\} \Longrightarrow \{\alpha f + (1-\alpha) h\} \succ \{\alpha g + (1-\alpha) h\}.$$

In the standard model, a menu is as good as the best alternative that it contains. Then

$$A \succeq B \Longrightarrow A \sim A \cup B,$$

a property called *strategic rationality* by Kreps [22]. Such a model excludes temptation. Temptation is an integral part of cognitive dissonance because the agent changes beliefs to make the menu at hand look attractive and then is tempted to make subsequent choices accordingly (see the discussion of utility in Section 2). In seeking a suitable relaxation of strategic rationality, we begin with GP's central axiom Set-Betweenness.

Set-Betweenness (SB): For all menus A and A', if $A \succeq A'$, then $A \succeq A \cup A' \succeq A'$.

An equivalent and perhaps more revealing, though less compact, statement is that if $A \succeq A'$, then one of the following conditions holds: (i) $A \sim A \cup A' \sim A'$, or (ii) $A \succ A \cup A' \succ A'$, or (iii) $A \sim A \cup A' \succ A'$, or (iv) $A \succ A \cup A' \sim A'$.

Following GP (p. 1408), we may interpret these conditions intuitively. The underlying assumptions are that: unchosen acts can only reduce utility, acts can be ranked according to how tempting they are, and only the most tempting act affects utility. Consider an agent having the menu $A \cup A'$, and who expects to choose g though she finds f most tempting. (i) is the residual case. (ii) indicates that g is in A' (hence $A \succ A \cup A'$), and that f is in A (hence $A \cup A' \succ A'$). The next two cases are our main interest.

In (iii), she still plans to choose out of A which now also contains the most tempting act. Confront her next with the larger menu $A \cup A' \cup B$. The most tempting act lies in $A \cup B$. What about her choice out of $A \cup A' \cup B$? Suppose that her expected choices satisfy the Nash-Chernoff condition (or Sen's property α); defer discussion of possible objections to this assumption. Then having rejected acts in A' when facing $A \cup A'$, she would (expect to) reject them also when facing

$A \cup A' \cup B$. Thus $A \cup B$ contains both the act to be chosen and also the act in $A \cup A' \cup B$ that is most tempting. The indifference $A \cup B \sim A \cup A' \cup B$ follows.

Finally, consider (iv), which indicates that both f and g lie in A'. Again confront the agent with the larger menu $A \cup A' \cup B$. The most tempting act lies in $B \cup A'$ and, assuming the Nash-Chernoff condition, so does the act to be chosen. Deduce the indifference $A' \cup B \sim A \cup A' \cup B$.

The preceding provides intuition for the following axiom:

Axiom 5 (Strong Set-Betweenness (SSB)). *For all menus A and A', if $A \succeq A'$, then: (i) $A \sim A \cup A' \sim A'$, or (ii) $A \succ A \cup A' \succ A'$, or*
(iii) $A \sim A \cup A' \succ A'$ and $A \cup B \sim A \cup A' \cup B$ for all menus B, or
(iv) $A \succ A \cup A' \sim A'$ and $A' \cup B \sim A \cup A' \cup B$ for all menus B.

Obviously SSB implies Set-Betweenness. The two axioms are equivalent given Independence (and Order and Continuity) - this follows from counterparts of the representation results in GP [16] and Kopylov [19] - and thus SSB is not invoked explicitly in those papers. However, we do not adopt Independence, and we show below that SSB is strictly stronger than Set-Betweenness even given all our other axioms.

Our intuition for SSB assumed the Nash-Chernoff condition, which can be criticized in a model of temptation - the addition of the acts in B should not affect the normative appeal of A versus A', but it may change the self-control costs associated with various choices, and this may lead to the choice of an act in A' when facing $A \cup B \cup A'$ even where she chooses an act in A when facing $A \cup A'$. GP's rationale (Theorem 5) for Set-Betweenness also relies on Nash-Chernoff, at least implicitly; see specifically their Axiom T1, which states that choosing an alternative from a menu A is always at least as good as choosing it from a larger menu.

The Nash-Chernoff condition is weaker than the weak axiom of revealed preference (WARP), which requires also Sen's condition β (see [22]). Noor [25] provides an example to illustrate why WARP may be problematic in a model of temptation, and in [26] he develops a model of temptation and self-control that does not impose WARP for ex post choice. Such objections apply also to the GP model since SSB and WARP for ex post choice are implied when one assumes Independence. Moreover, while they may be important for guiding development of a general model of temptation, these concerns do not seem germane to temptation generated by cognitive dissonance.

One can raise other objections to Set-Betweenness, and hence *a fortiori* to our stronger axiom. Dekel, Lipman and Rustichini [7] argue that Set-Betweenness excludes some forms of temptation, for example, where the temptation generated by different alternatives is cumulative, or where there is uncertainty ex ante about which alternatives will be tempting. Once again, we do not view these concerns as especially important for a model of cognitive dissonance.

Say that $f \in \mathcal{H}$ *dominates* $g \in \mathcal{H}$ if $\{f(s)\} \succeq \{g(s)\}$ for every $s \in S$. If the evaluation of a lottery does not depend on the state, then a dominating act should be preferred under commitment. Similarly, if f dominates g, we would not expect f to be tempted by g. Thus we assume:

Axiom 6 (Monotonicity). *If f dominates g, then $\{f\} \sim \{f, g\} \succeq \{g\}$.*

Our axioms thus far have for the most part been concerned with modeling temptation in general, that is, not tied specifically to cognitive dissonance. A partial exception is Collinear Independence, the intuition for which did rely on the assumption that temptation arises because of an ex post choice of beliefs to "fit the menu" in hand. However, Collinear Independence is satisfied even if the agent becomes more pessimistic ex post and adopts beliefs that make the menu less attractive ex post. The final two axioms build in ex post optimism and hence cognitive dissonance.

Axiom 7 (Constants-Do-Not-Tempt). *For all $c \in \mathcal{C}$ and $f \in \mathcal{H}$,*

$$\{f\} \succ \{c\} \Longrightarrow \{f\} \sim \{c, f\}.$$

Temptation is due to a change in beliefs (as opposed to a change in risk aversion, for example), which leaves the evaluation of constant acts unaffected. In addition, the noted change is always to become more optimistic ex post about the available menu, rendering it even more attractive relative to any constant act c than it was ex ante. Therefore, constant acts cannot tempt. Note that, in contrast, $\{c\} \succ \{c, f\} \succeq \{f\}$ is both permitted by the axiom and intuitive given our story.

Axiom 8 (Convex Temptation). *The set $\{f \in \mathcal{H} : \{c\} \sim \{c, f\} \succ \{f\}\}$ is convex for every $c \in \mathcal{C}$.*

Suppose that f and g both lie in the indicated set, that is, each is worse than c under commitment and neither tempts c, and consider the mixture $\alpha f + (1-\alpha)\, g$. By Commitment Independence, $\{c\} \succ \{\alpha f + (1-\alpha)\, g\}$. We now argue that in addition, $\alpha f + (1-\alpha)\, g$ should not tempt c, thus completing intuition for the axiom. We are given that $\{c\} \sim \{c, f\}$. Because ex post beliefs are chosen to make the menu $\{c, f\}$ attractive, and because the expected utility of c does not depend on beliefs, we can interpret the indicated indifference as follows: the act f, when matched with the beliefs that make it attractive, does not tempt c. A similar statement applies for g. Consider now the menu $\{c, \alpha f + (1-\alpha) g\}$. Beliefs to render this menu attractive are chosen ex post (time 1), before the randomization is completed (which, as noted earlier, occurs only at the terminal time after the true state in S is realized). Since *the beliefs that make f attractive may differ from those that make g attractive*, matching beliefs with the mixed act is more difficult. Therefore, one would expect the mixed act not to tempt c.

4. REPRESENTATION RESULT

Our main result is that the preceding axioms characterize the functional form described in Section 2.

Theorem 4.1. *The binary relation $\succeq$ on $\mathcal{K}(\mathcal{H})$ may be represented as in (2.1)-(2.3) if and only if it satisfies Axioms 1-8. Moreover, u is unique up to a positive linear transformation, and if $\succeq$ is not strategically rational, then p, Q and κ are unique.*

Convex Temptation is used only at the very end of the sufficiency proof in order to prove that V has the form given in (2.3). If the axiom is deleted, then the remaining axioms characterize the representation (2.1)-(2.2), for some $V : \mathcal{H} \to \mathbb{R}^1$ that is continuous, monotone ($V(f) \geq V(g)$ if f dominates g), satisfies certainty additivity ($V(\alpha f + (1-\alpha)\, c) = \alpha V(f) + (1-\alpha)\, V(c)$ for all c in $\mathcal{C}$), and that satisfies $V(f) \geq p \cdot u(f)$ for all f with equality if f is constant. (See Example 3 below.)

We present some examples to demonstrate the tightness of the characterization in the theorem. Each of the first three examples satisfies Order, Continuity, Commitment Independence, Strong Set-Betweenness and Monotonicity, and violates precisely one of the axioms that relate more specifically to cognitive dissonance - Collinear Independence, Constants-Do-Not-Tempt and Convex Temptation. The final example violates only Strong Set-Betweenness, though it satisfies GP's Set-Betweenness, thus proving that our adoption of the stronger axiom is necessary.

Example 1: Let

$$\mathcal{U}(A) = \frac{w(A)}{v(A)} \equiv \frac{\max_{h \in A} [U(h) V(h)]}{\max_{h' \in A} V(h')},$$

where U and V are as in (2.2)-(2.3), and where $u > 0$. Then $\succeq$ violates only Collinear Independence. In particular, to verify SSB note that for any menus A and A', there are only four possible cases:

(i) $w(A) > w(A')$ and $v(A') > v(A)$; then $\mathcal{U}(A) > \mathcal{U}(A \cup A') > \mathcal{U}(A')$.

(ii) $w(A') > w(A)$ and $v(A) > v(A')$; then $\mathcal{U}(A') > \mathcal{U}(A \cup A') > \mathcal{U}(A)$.

(iii) $w(A) \geq w(A')$ and $v(A) \geq v(A')$; then for all menus B, $w(A \cup B) = w(A \cup A' \cup B)$, $v(A \cup B) = v(A \cup A' \cup B)$, and hence, $\mathcal{U}(A \cup B) = \mathcal{U}(A \cup A' \cup B)$.

(iv) $w(A') \geq w(A)$ and $v(A') \geq v(A)$; then analogously to (iii), $\mathcal{U}(A' \cup B) = \mathcal{U}(A \cup A' \cup B)$ for all menus B.

There exist simpler examples violating only Collinear Independence - these retain (2.1)-(2.2) but modify the specification of V. However, because the above ratio form deviates from the GP functional form, we find it more revealing about the power of Collinear Independence.[8]

Example 2: Assume (2.1)-(2.2), but take

$$V(h) = \ q \cdot u(h),$$

for some probability measure $q \neq p$. Then $\succeq$ violates only Constants-Do-Not-Tempt.

Example 3: Modify Example 2 by taking

$$V(h) = \max \left\{ p \cdot u(f), \ \int u(f) \, d\nu \right\},$$

[8] The example is inspired by weighted utility theory [5], a model of risk preference in which the utility function over lotteries equals a ratio of expected utility functions. Readers familiar with the 'non-expected utility' literature will not be surprised by the observation that $\succeq$ satisfies the following alternative relaxation of Independence: $A \sim B \Longrightarrow \ \alpha A + (1-\alpha) B \sim A$.

where ν is a capacity on S and the integral $\int u\,(f)\,d\nu$ is in the sense of Choquet (see Schmeidler [31]). Then $\succeq$ violates only Convex Temptation.

Example 4: This example violates Strong Set-Betwenness, but satisfies all other axioms adopted in Theorem 4.1 as well as GP's Set-Betweenness. Let $S = \{s_1, s_2\}$, and fix a vNM utility function u such that $u(\mathcal{C}) = [0, 1]$. For every $f \in \mathcal{H}$, let $u_1(f) = u(f(s_1))$, $u_2(f) = u(f(s_2))$, and

$$\gamma(f) = \max\{0, u_2(f) - u_1(f) - \tfrac{4}{5}\},$$
$$\nu(f) = \max\{0, u_1(f) - u_2(f) - \tfrac{4}{5}\}.$$

Let $\succeq$ be represented by $\mathcal{U}$, where, for all menus A,

$$\mathcal{U}(A) = \max_{f \in A}[u_2(f) - \gamma(f) \max_{g \in A} \nu(g)].$$

Then $\succeq$ satisfies Order and Continuity. By construction,

$$\gamma\,(f) \cdot \nu(g) > 0 \quad \Longrightarrow \quad u_1\,(f)\,,\ u_2\,(g) < \tfrac{1}{5} \quad \text{and} \quad u_2\,(f)\,,\ u_1\,(g) > \tfrac{4}{5}.$$

Therefore, $\gamma(f)\nu(f) = 0$ for all f, so that $\mathcal{U}(\{f\}) = u_2\,(f)$, implying Commitment Independence. In addition, $\gamma(f)\nu(g) = 0$ holds in each of the following cases (i) f or g is constant; (ii) f and g are collinear; (iii) f dominates g; (iv) g dominates f. Thus $\succeq$ satisfies Collinear Independence, Monotonicity, Constants-Do-Not-Tempt, and Convex Temptation. For Set-Betweenness, take any menus A and B and acts $f, g \in A \cup B$ that deliver the maxima in the definition of $\mathcal{U}(A \cup B)$, so that $\mathcal{U}(A \cup B) = u_2\,(f) - \gamma\,(f)\,\nu\,(g)$. Wlog $f \in A$. Then $\mathcal{U}(A) = \mathcal{U}(A \cup B)$ if $g \in A$, and $\mathcal{U}(A) \geq \mathcal{U}(A \cup B) \geq \mathcal{U}(B)$ if $g \in B$. However, $\succeq$ violates SSB: if $u_1(f) = 0$, $u_2(f) = \frac{1}{2}$, $u_1(f') = 1$, $u_2(f') = 0$, $u_1(g) = 0$, and $u_2(g) = 1$, then $\{f\} \sim \{f, f'\} \succ \{f'\}$, but $\{f, g\} \succ \{f, f', g\}$.

A tuple (u, p, Q, κ) as in the theorem is said to *represent* $\succeq$. The representing tuple is unique (up to cardinal equivalence for u) if the degenerate case of strategic rationality is excluded. Thus it is meaningful to ask about behavioral interpretations of its components. We have already noted those of u and p: u ranks lotteries (constant acts) and p is the "commitment prior" - it underlies the ranking of singleton menus. Turn to Q and κ. In what follows, we adopt variants of GP's comparative notions "greater preference for commitment" and "greater self-control", renamed so as to reflect better the psychological motives we have in mind.

Say that $\succeq^*$ *has greater dissonance than* $\succeq$ if for all acts f and g,

$$\{f\} \succ \{f,g\} \Longrightarrow \{f\} \succ^* \{f,g\}. \tag{4.1}$$

The ranking $\{f\} \succ \{f,g\}$ indicates that though f is better than g ex ante, g is better ex post when holding the menu $\{f,g\}$. Then there is dissonance for the agent with preference $\succeq$ between the ex ante ranking under commitment (or the underlying beliefs) and the distinct ex post ranking (or ex post beliefs). If $\succeq^*$ has greater dissonance, then she should strictly prefer $\{f\}$ to $\{f,g\}$.

Theorem 4.2. *Suppose that both $\succeq$ and $\succeq^*$ have utility representations (2.1)-(2.3), with components (u,p,Q,κ) and (u^*,p^*,Q^*,κ^*) respectively, and that neither is strategically rational. Then $\succeq^*$ has greater dissonance than $\succeq$ if and only if*

$$(u,p) = (au^* + b, p^*) \text{ for some } a > 0 \text{ and some } b, \text{ and} \tag{4.2}$$

$$Q = (1-\epsilon)\{p\} + \epsilon Q^*, \text{ for some } 0 < \epsilon \leq 1. \tag{4.3}$$

The characterizing conditions assert both that the commitment rankings induced by $\succeq$ and $\succeq^*$ coincide (this is (4.2)) and that Q is "closer to p" than is Q^* in the sense of an epsilon contamination (this is (4.3)). Since Q^* is convex and contains p, (4.3) implies in particular that $Q \subseteq Q^*$, but it implies more. Note that if $\succeq$ is strategically rational, then any $\succeq^*$ has greater dissonance—the defining condition is satisfied vacuously—and no restrictions on commitment preferences are implied. If $\succeq^*$ is strategically rational, then (4.1) is satisfied if and only if $\succeq$ is also strategically rational, and again, condition (4.2) is not implied.

We are interested not only in how much dissonance an agent experiences (or expects to experience), but also in what she does about it, or more precisely, in the extent to which ex post choices are distorted by dissonance. Say that $\succeq^*$ *is more self-justifying than* $\succeq$ if it has more dissonance than $\succeq$ and

$$\{f\} \succ \{f,g\} \sim \{g\} \Longrightarrow \{f\} \succ^* \{f,g\} \sim^* \{g\}.$$

The hypothesized rankings for $\succeq$ indicate not only that there is dissonance but also that given $\{f,g\}$ at the ex post stage, the agent succumbs and chooses g, even though f was optimal ex ante under commitment. She does this because the choice of g better justifies her previous choice of $\{f,g\}$. If $\succeq^*$ is more self-justifying, then she should also choose g out of $\{f,g\}$.

Theorem 4.3. *Suppose that both $\succeq$ and $\succeq^*$ have utility representations (2.1)-(2.3), with components (u, p, Q, κ) and $(u^*, p^*, Q^*, \kappa^*)$ respectively, and that neither is strategically rational. Then $\succeq^*$ is more self-justifying than $\succeq$ if and only if $(u^*, p^*, Q^*, \kappa^*)$ and (u, p, Q, κ) satisfy (4.2), (4.3) and $\kappa^* \geq \epsilon\kappa$.*

It follows that a change from κ to $\kappa^* > \kappa$, keeping other components of the functional form fixed, renders $\succeq^*$ more self-justifying than $\succeq$ but leaves the two preference orders equally dissonant (each has greater dissonance than the other).

5. EXTENSIONS

To conclude, we outline two generalizations of the above model.[9]

5.1. Effort and Dissonance

An intuitive prediction of dissonance theory is that cognitive dissonance is more pronounced when past actions are "difficult". As Aronson writes (p. 175), "if a person works hard to attain a goal, that goal will be more attractive to the individual than it will be to someone who achieves the same goal with little or no effort." See [3, pp. 175-8] and [1, p. 310] for discussion and references to supporting experimental evidence. Here we outline an extension of the model that can accommodate this prediction.

Modify the time line described in Section 2 only by supposing that the choice to made at the ex ante stage is of a pair (e_0, A), where $e_0 \in \mathcal{E}$ denotes effort in period 0 and A is, as before, a menu of Anscombe-Aumann acts one of which will be chosen in the following period. Ex ante choices are assumed to maximize preference $\succeq$, which is defined on $\mathcal{E} \times \mathcal{K}(\mathcal{H})$.

Let utility have the form

$$\mathcal{U}(e_0, A) = \max_{h \in A} \left[U(e_0, h) + \tfrac{\kappa(e_0)}{1-\kappa(e_0)} \, V(e_0, h) \right] - \tfrac{\kappa(e_0)}{1-\kappa(e_0)} \, \max_{h' \in A} V(e_0, h'), \quad (5.1)$$

where

$$U(e_0, h) = -v(e_0) + \delta \; p \cdot u(h), \text{ and} \quad (5.2)$$

$$V(e_0, h) = -v(e_0) + \delta \max_{q \in Q} \; q \cdot u(h). \quad (5.3)$$

[9] The extensions are in terms of functional forms. We have not provided axiomatic foundations, though we believe it would be straightforward to do so.

Here $0 \leq \kappa(e_0) < 1$, p is a probability measure on S, Q is a convex and compact set of probability measures on S containing p, $u : \mathcal{C} \longrightarrow \mathbb{R}^1$ is mixture linear and continuous, $v : \mathcal{E} \rightarrow \mathbb{R}^1$ gives the utility cost of effort, and $0 < \delta < 1$ is a discount factor. When restricted to singletons,

$$\mathcal{U}(e_0, \{h\}) = -v(e_0) + \delta p \cdot u(h).$$

For nonsingletons, ex post choice out of A solves

$$\max_{h \in A} \max_{q \in (1-\kappa(e_0))\{p\}+\kappa(e_0)Q} q \cdot u(h),$$

which depends on e_0 via $\kappa(\cdot)$.

Suppose that

$$\kappa(e_0) = \widehat{\kappa}(v(e_0)),$$

where $\widehat{\kappa}(\cdot)$ is increasing. Then an increase in $v(e_0)$, corresponding to greater effort, renders the agent more self-justifying, but leaves the level of dissonance unchanged.[10] More generally, we could also specify Q as a function of $v(e_0)$, for example,

$$Q = (1 - \epsilon(v(e_0)))\{p\} + \epsilon(v(e_0))\Delta(S).$$

If $\widehat{\kappa}(\cdot)\epsilon(\cdot)$ is increasing, then greater effort implies both greater dissonance and greater self-justification.[11]

5.2. Response to Information

The justification of a past decision may also influence the reaction to information - dissonance theory predicts that information is interpreted in a way that is favorable to past choices. By adding a signal realized at time 1 and building on Epstein [10], we can extend our model to capture also the response to information.

An outline of the model is as follows: let S_1 denote the (finite) space of signals, one of which is realized at time 1. Ex ante, the agent chooses a *contingent menu* - a mapping F from signals into menus of Anscombe-Aumann acts. At time 1, she observes the realized signal, updates her beliefs about S, and then chooses

[10] We are using the formal comparative notions defined in the preceding section applied to the preferences on $\mathcal{K}(\mathcal{H})$ induced by $\succeq$ and the two levels of consumption.

[11] If $Q_i = (1-\epsilon_i)\{p\} + \epsilon_i \Delta(S)$, $i = 1, 2$, with $\epsilon_1 \geq \epsilon_2$, then $Q_2 = (1-\epsilon)\{p\} + \epsilon Q_1$ with $\epsilon = \epsilon_2/\epsilon_1$. Thus Theorem 4.3 implies that preference 1 is more self-justifying (and has greater dissonance) than preference 2 if $\epsilon_1\kappa_1 \geq \epsilon_2\kappa_2$.

an act from the realized menu $F(s_1)$. Denote by p prior beliefs on $S_1 \times S$, by p_1 its first marginal, and, for each signal s_1, let Q_{s_1} be a (closed and convex) set of probability measures on S containing $p(\cdot \mid s_1)$, the Bayesian update of p. Then the utility of any contingent menu F is given by

$$\mathcal{W}(F) = \int_{S_1} \mathcal{U}(F(s_1); s_1) \; dp_1(s_1),$$

where, for any menu A,

$$\mathcal{U}(A; s_1) = \max_{h \in A} \left[(1-\kappa) U(h; s_1) + \kappa V(h; s_1)\right] - \kappa \max_{h' \in A} V(h'; s_1),$$

$$U(h; s_1) = \; p(\cdot \mid s_1) \cdot u(h), \text{ and}$$
$$V(h) = \max_{q \in Q_{s_1}} \; q \cdot u(h).$$

The interpretation is clear given the parallel with our model (2.1)-(2.3). The key is that at the ex post stage, the agent does not rely simply on the Bayesian update $p(\cdot \mid s_1)$ of her prior beliefs, but rather behaves as though she adjusts the latter in a direction that renders the realized menu $F(s_1)$ attractive, as indicated by the maximization over Q_{s_1}. As a result the signal is interpreted so as to justify the past choice of an action (that is, F).[12]

A. Appendix: Proof of the Representation Theorem

For necessity, verification of the axioms is straightforward.

The proof of sufficiency proceeds roughly as follows: apply the Anscombe-Aumann Theorem to derive an expected utility function $U : \mathcal{H} \to \mathbb{R}^1$ for preference restricted to singleton menus. This delivers a linear utility index $u : \mathcal{C} \to \mathbb{R}^1$ and a prior p on S, such that $U(f) = p \cdot u(f)$. Next, for any $f \in \mathcal{H}$, let

$$\mathcal{H}_f = \{\alpha c + (1-\alpha) f : t \in [0,1], c \in \mathcal{C}\},$$

and let $\mathcal{A}_f$ be the class of menus in $\mathcal{H}_f$. Then $\mathcal{H}_f$ is a compact mixture space, and $\succeq$ restricted to $\mathcal{A}_f$ satisfies Independence (because $\succeq$ satisfies Collinear Independence) and Set-Betweenness. Thus, by Kopylov's [19, Theorem 2.1] extension of GP's theorem to mixture spaces, one obtains a continuous and linear function $V_f : \mathcal{H}_f \to \mathbb{R}$ such that

$$\mathcal{U}(A) = \max_{h \in A} (U(h) + V_f(h)) - \max_{h \in A} V_f(h)$$

[12] A closely related bias, called confirmatory bias, states that people tend to interpret evidence in ways that confirm prior beliefs, as opposed to past actions (see [29], for example).

represents $\succeq$ on $\mathcal{A}_f$. The critical step is to extend the local functions V_f to a global temptation function V. The remaining step is to show that V has the form (2.3) for some Q, which is done by analogy with the proof of Gilboa and Schmeidler [14] (suitably modified for the maxmax model rather than maxmin).

Turn to the detailed proof. Throughout abbreviate the domain $\mathcal{K}(\mathcal{H})$ by $\mathcal{A}$, and assume that $\succeq$ is *non-degenerate*, that is, $A \succ B$ for some $A, B \in \mathcal{A}$. (Otherwise, the desired representation holds trivially with $u \equiv 0$.)

Lemma A.1. *There exist a continuous function $\mathcal{U} : \mathcal{A} \to \mathbb{R}$, a probability measure p on S, and a non-constant expected utility function $u : \mathcal{C} \to \mathbb{R}$ such that $\mathcal{U}$ represents $\succeq$ and*

$$\mathcal{U}(\{f\}) = p \cdot u(f) \quad \text{for all } f \in \mathcal{H}. \tag{A.1}$$

Such p is unique, and u is unique up to a positive linear transformation.

Proof. By the Anscombe–Aumann Theorem, the axioms of Order, Continuity, Monotonicity, and Commitment Independence imply that the preference $\succeq$ restricted to singleton menus can be represented by $\mathcal{U}(\{f\}) = p \cdot u(f)$, where p is a probability measure on S, and $u : \mathcal{C} \to \mathbb{R}$ is a continuous vNM expected utility function. As $\mathcal{C}$ is compact, there exist lotteries $c_+, c_- \in \mathcal{C}$ such that $u(c_+) \geq u(c) \geq u(c_-)$ for all $c \in \mathcal{C}$. Then $\{c_+\} \succeq \{f\} \succeq \{c_-\}$ for all $f \in \mathcal{H}$. By SSB, $\{c_+\} \succeq A \succeq \{c_-\}$ for all finite menus A; by Continuity, $\{c_+\} \succeq A \succeq \{c_-\}$ for all menus $A \in \mathcal{A}$. As $\succeq$ is non-degenerate, $\{c_+\} \succ \{c_-\}$ and hence, u is non-constant.

By Continuity, for any $A \in \mathcal{A}$, there exists a unique $\alpha \in [0, 1]$ such that $A \sim \{\alpha c_+ + (1 - \alpha)\, c_-\}$. Let

$$\mathcal{U}(A) = u(\alpha c_+ + (1 - \alpha)\, c_-).$$

Then $\mathcal{U}$ represents $\succeq$ on $\mathcal{A}$ and inherits continuity from $\succeq$. □

Hereafter, fix $c_+, c_- \in \mathcal{C}$ as in the proof of the above lemma, and fix the unique u (and the unique corresponding $\mathcal{U}$) such that $u(c_+) = 1$ and $u(c_-) = -1$. Let $c_0 = \frac{c_+ + c_-}{2}$; then $u(c_0) = 0$.

For every act $f \in \mathcal{H}$, let

- $U(f) = \mathcal{U}(\{f\}) = p \cdot u(f)$
- $e(f) = \frac{1+U(f)}{2} c_+ + \frac{1-U(f)}{2} c_-$; then $e(f) \in \mathcal{C}$ and $\{f\} \sim \{e(f)\}$
- $f + \alpha = \alpha c_+ + (1 - \alpha) f$ and $f - \alpha = \alpha c_- + (1 - \alpha) f$

Take an arbitrary act $f \in \mathcal{H}$ and invoke [19, Theorem 2.1]: $\mathcal{H}_f$ is a compact mixture space satisfying properties M1–M4 in [19], and $\succeq$ restricted to $\mathcal{A}_f$ satisfies Order, Continuity, Binary Independence, and Set-Betweenness, the axioms in the cited theorem. Thus $\succeq$ can be represented on $\mathcal{A}_f$ by

$$\mathcal{U}_f(A) = \max_{g \in A}(U_f(g) + V_f(g)) - \max_{g \in A} V_f(g),$$

where the functions $U_f : \mathcal{H}_f \to \mathbb{R}$ and $V_f : \mathcal{H}_f \to \mathbb{R}$ are continuous, linear, non-constant, and normalized by $U_f(c_+) = 1$ and $U_f(c_0) = V_f(c_0) = 0$. The normalization of U_f implies that $U_f \equiv u \equiv U$ on $\mathcal{C} \subset \mathcal{H}_f$, and hence, $U_f \equiv U$ on $\mathcal{H}_f$. It follows that $\mathcal{U}_f \equiv \mathcal{U}$ on $\mathcal{A}_f$. (To see this, for any A in $\mathcal{A}_f$, let $e(A) \in \mathcal{C}$ satisfy $A \sim \{e(A)\}$. Then $\mathcal{U}(A) = u(e(A)) = \mathcal{U}_f(e(A)) = \mathcal{U}_f(A)$.) Thus

$$\mathcal{U}(A) = \max_{g \in A}(U(g) + V_f(g)) - \max_{g \in A} V_f(g) = \max_{g \in A} W_f(g) - \max_{g \in A} V_f(g), \tag{A.2}$$

where $W_f(\cdot) = U(\cdot) + V_f(\cdot)$ on $\mathcal{H}_f$.

Show that V_f is monotonic. Take any $h, h' \in \mathcal{H}_f$ such that h dominates h'. For all $\alpha \in (0,1)$, Monotonicity and Lemma A.1 imply $\{h+\alpha\} \sim \{h+\alpha, h'-\alpha\} \succ \{h'-\alpha\}$ and hence, $V_f(h+\alpha) \geq V_f(h'-\alpha)$. Let $\alpha \to 0$; then $V_f(h) \geq V_f(h')$ by continuity.

The monotonicity of V_f implies that for all $c, c' \in \mathcal{C}$, if $\{c\} \succeq \{c'\}$, then $V_f(c) \geq V_f(c')$. By the vNM theorem, V_f is a positive linear transformation of U on $\mathcal{C}$. As $U(c_0) = V_f(c_0) = 0$, there exists $\kappa_f \in (0,1)$ such that $V_f(\cdot) = \frac{\kappa_f}{1-\kappa_f} U(\cdot)$ on $\mathcal{C}$.

Say that $f \in \mathcal{H}$ is *never tempting* if $\{c, f\} \succeq \{c\}$ for all $c \in \mathcal{C}$; otherwise call f *potentially tempting*.

Lemma A.2. *If $f \in \mathcal{H}$ is never tempting, then $V_f(f) = V_f(e(f))$, and $\mathcal{U}(A) = \max_{g \in A} U(g)$ for all $A \in \mathcal{A}_f$. If $f \in \mathcal{H}$ is potentially tempting, then $V_f(f) > V_f(e(f))$, and the representation (A.2) is unique.*

Proof. The act f must satisfy exactly one of the following three cases.

Case 1. $V_f(f) = V_f(e(f))$. Let $g = \alpha f + (1-\alpha)c \in \mathcal{H}_f$. Then

$$V_f(g) = \alpha V_f(f) + (1-\alpha)V_f(c) = \tfrac{\kappa_f}{1-\kappa_f}(\alpha U(e(f)) + (1-\alpha)U(c)) = \tfrac{\kappa_f}{1-\kappa_f} U(g),$$

and $\mathcal{U}(A) = \max_{g \in A}(U(g) + V_f(g)) - \max_{g \in A} V_f(g) = \max_{g \in A} U(g)$ for all $A \in \mathcal{A}_f$. Thus $\succeq$ satisfies Strategic Rationality on $\mathcal{A}_f$, and f is never tempting.

Case 2. $V_f(f) > V_f(e(f))$. The monotonicity of V_f implies that $V_f(c_+) \geq V_f(f) > V_f(e(f))$ and $U(c_+) > U(e(f))$. Then

$$1 \geq \tfrac{V_f(f) - V_f(e(f))}{V_f(c_+) - V_f(e(f))} > \tfrac{V_f(f) - V_f(e(f))}{V_f(c_+) - V_f(e(f)) + U(c_+) - U(e(f))} = \tfrac{W_f(f) - W_f(e(f))}{W_f(c_+) - W_f(e(f))} > 0. \tag{A.3}$$

Take any $\alpha \in (0,1)$ such that $\frac{V_f(f) - V_f(e(f))}{V_f(c_+) - V_f(e(f))} > \alpha > \frac{W_f(f) - W_f(e(f))}{W_f(c_+) - W_f(e(f))}$. Then $W_f(e(f)+\alpha) > W_f(f)$ and $V_f(f) > V_f(e(f)+\alpha)$ because W_f and V_f are linear. By (A.2), $\{e(f)+\alpha\} \succ \{e(f)+\alpha, f\} \succ \{f\}$. This ranking implies that f is potentially tempting, and by [19, Theorem 2.1], that the representation (A.2) on $\mathcal{H}_f$ is unique.

Case 3. $V_f(e(f)) > V_f(f)$. Then by Continuity, there exists α such that $V_f(e(f)-\alpha) > V_f(f+\alpha)$. By (A.2), $\{f+\alpha\} \succ \{f+\alpha, e(f)-\alpha\} \succeq \{e(f)-\alpha\}$, which contradicts Constants-Do-Not-Tempt. So this case is impossible.

It follows that f is never tempting if and only if Case 1 holds, and f is potentially tempting if and only if Case 2 holds. $\square$

Lemma A.3. *If $f, g \in \mathcal{H}$ are potentially tempting, then $\kappa_f = \kappa_g$.*

Proof. Given the potentially tempting acts $f, g \in \mathcal{H}$, let

$$f_0 = \begin{cases} \frac{1}{U(f)+1} f + \frac{U(f)}{U(f)+1} c_- & \text{if } U(f) \geq 0 \\ \frac{1}{1-U(f)} f + \frac{-U(f)}{1-U(f)} c_+ & \text{if } U(f) < 0, \end{cases} \qquad g_0 = \begin{cases} \frac{1}{U(g)+1} g + \frac{U(g)}{U(g)+1} c_- & \text{if } U(g) \geq 0 \\ \frac{1}{1-U(g)} g + \frac{-U(g)}{1-U(g)} c_+ & \text{if } U(g) < 0. \end{cases}$$

Then $f_0 \in \mathcal{H}_f$, $g_0 \in \mathcal{H}_g$, and $e(f_0) = e(g_0) = c_0$. By Lemma A.2, $V_f(f_0) > V_f(c_0) = 0$ and $W_f(f_0) > 0$. Fix $\gamma \in (0,1)$ such that $V_f(f_0 - \gamma) > 0$ and $W_f(f_0 - \gamma) > 0$. By (A.2), $\mathcal{U}(\{c_0, f_0 - \gamma\} = \mathcal{U}(\{f_0 - \gamma\}$ and $\mathcal{U}(\{c_0, f_0 - 1\}) = \mathcal{U}(\{c_0, c_-\}) = 0$.

Define a function ϕ on $(0,1]$ by

$$\phi(\alpha) = \frac{\mathcal{U}(\{c_0, f_0 - \alpha\})}{\mathcal{U}(\{f_0 - \alpha\})} = \frac{\mathcal{U}(\{c_0, f_0 - \alpha\})}{-\alpha}.$$

Then ϕ is continuous and satisfies $\phi(\gamma) = 1$ and $\phi(1) = 0$. By continuity, there exists $\alpha_f \in (\gamma, 1)$ such that $\phi(\alpha_f) = \frac{1}{2}$. Then $U(f_0 - \alpha_f) = -\alpha_f$ and $V_f(f_0 - \alpha_f) = \frac{\alpha_f}{2}$. Analogously, find $\alpha_g \in (0,1)$ such that $U(g_0 - \alpha_g) = -\alpha_g$ and $V_g(g_0 - \alpha_g) = \frac{\alpha_g}{2}$.

Let $f' = \alpha_g(f_0 - \alpha_f) + (1 - \alpha_g)c_0$ and $g' = \alpha_f(g_0 - \alpha_g) + (1 - \alpha_f)c_0$. Then

$$V_f(f') = V_g(g') = \tfrac{\alpha_f \alpha_g}{2} > 0 > W_f(f') = W_g(g') = -\tfrac{\alpha_f \alpha_g}{2} > U(f') = U(g') = -\alpha_f \alpha_g.$$

SSB and the representation (A.2) imply the rankings

$$\{c_0\} \succ \{c_0, f'\} \sim \{c_0, f', g'\} \sim \{c_0, g'\} \succ \{f'\} \sim \{g'\}.$$

By SSB, $\{c_0, f'\} \cup B \sim \{c_0, f', g'\} \cup B \sim \{c_0, g'\} \cup B$ for all menus B. Take $\varepsilon > 0$ such that $V_f(c_0 + \varepsilon) < V_f(f')$ and $V_g(c_0 + \varepsilon) < V_g(g')$. Then $\{c_0 + \varepsilon, c_0, f'\} \sim \{c_0 + \varepsilon, c_0, g'\}$. By (A.2), $\mathcal{U}(\{c_0 + \varepsilon, c_0, f'\}) = \frac{\varepsilon}{1-\kappa_f} - \frac{\alpha_f \alpha_g}{2}$, and $\mathcal{U}(\{c_0 + \varepsilon, c_0, g'\}) = \frac{\varepsilon}{1-\kappa_g} - \frac{\alpha_f \alpha_g}{2}$. Thus $\frac{\varepsilon}{1-\kappa_f} - \frac{\alpha_f \alpha_g}{2} = \frac{\varepsilon}{1-\kappa_g} - \frac{\alpha_f \alpha_g}{2}$, that is, $\kappa_f = \kappa_g$. □

Let $\kappa \in (0,1)$ be such that $\kappa_h = \kappa$ for all potentially tempting acts $h \in \mathcal{H}$ (κ exists by Lemma A.3). For every $f \in \mathcal{H}$, let $W(f) = U(f) + V(f)$, where

- $V(f) = V_f(f)$ if f is potentially tempting,
- $V(f) = \frac{\kappa}{1-\kappa} U(f)$ if f is never tempting.

For every menu $A \in \mathcal{A}$, let

$$\mathcal{U}_{WV}(A) = \max_{g \in A} W(g) - \max_{g \in A} V(g).$$

Later we show that both W and V are continuous and hence, the maxima in the above definition are obtained even if A is not finite.

Lemma A.4. *If $f \in \mathcal{H}$ is potentially tempting, then:*

(i) $V(f) > \frac{\kappa}{1-\kappa} U(f)$,

(ii) $V(\cdot) = V_f(\cdot)$ *and* $W(\cdot) = W_f(\cdot)$ *on* $\mathcal{H}_f$,

(iii) *for all finite menus* $A \in \mathcal{A}_f$, $\mathcal{U}(A) = \mathcal{U}_{WV}(A)$.

Proof. Let f be potentially tempting. (i) By Lemma A.2,

$$V(f) = V_f(f) > V_f(e(f)) = \tfrac{\kappa}{1-\kappa} U(e(f)) = \tfrac{\kappa}{1-\kappa} U(f).$$

(ii) Let $g = \alpha f + (1-\alpha)c \in \mathcal{H}_f$. If $\alpha = 0$, then $g = c \in \mathcal{C}$ and $V(g) = \frac{\kappa}{1-\kappa} U(c) = V_f(g)$. If $\alpha > 0$, then g is potentially tempting because there exists $c' \in \mathcal{C}$ such that $\{c'\} \succ \{c', f\}$, which implies $\{\alpha c' + (1-\alpha)c\} \succ \{\alpha c' + (1-\alpha)c, g\}$ by Collinear Independence. By Lemma A.2, the function V_g in representation (A.2) is unique, and hence, $V_g(\cdot) = V_f(\cdot)$ on $\mathcal{H}_g \subseteq \mathcal{H}_f$. In particular, $V(g) = V_g(g) = V_f(g)$.
(iii) If $A \in \mathcal{A}_f$, then, by Lemma A.2,

$$\mathcal{U}(A) = \max_{g\in A} W_f(g) - \max_{g\in A} V_f(g) = \max_{g\in A} W(g) - \max_{g\in A} V(g) = \mathcal{U}_{WV}(A)$$

because $V(g) = V_f(g)$ and $W(g) = W_f(g)$ for all $g \in A$. $\square$

Lemma A.5. *For all finite menus* $A \in \mathcal{A}$ *and for all acts* $f, g \in \mathcal{H}$, *if* $U(f) = U(g)$ *and* $V(f) = V(g)$, *then* $\mathcal{U}(\{f\} \cup A) = \mathcal{U}(\{g\} \cup A)$.

Proof. Fix A, f and g as in the hypothesis and consider two possible cases.
Case 1. f is never tempting. By Lemma A.2, for all $\alpha \in (0,1)$,

$$\{f+\alpha\} \sim \{f+\alpha, e(f)-\alpha\} \succ \{e(f)-\alpha\},$$
$$\{e(f)+\alpha\} \sim \{e(f)+\alpha, f-\alpha\} \succ \{f-\alpha\},$$

and by SSB, $\{f+\alpha\} \cup A \sim \{f+\alpha, e(f)-\alpha\} \cup A$ and $\{e(f)+\alpha\} \cup A \sim \{e(f)+\alpha, f-\alpha\} \cup A$. Let $\alpha \to 0$; by Continuity $\{f\} \cup A \sim \{f, e(f)\} \cup A \sim \{e(f)\} \cup A$. The equality $V(g) = V(f) = \frac{\kappa}{1-\kappa} U(f) = \frac{\kappa}{1-\kappa} U(g)$ implies, by Lemma A.4(i), that g is never tempting. Therefore, a similar argument proves that $\{g\} \cup A \sim \{g, e(g)\} \cup A \sim \{e(g)\} \cup A$. Finally, $\{f\} \cup A \sim \{e(f)\} \cup A = \{e(g)\} \cup A \sim \{g\} \cup A$, that is, $\mathcal{U}(\{f\} \cup A) = \mathcal{U}(\{g\} \cup A)$.
Case 2. f is potentially tempting. By (A.3) and Lemma A.4.(ii),

$$1 \geq \tfrac{V(f)-V(e(f))}{V(c_+)-V(e(f))} > \tfrac{W(f)-W(e(f))}{W(c_+)-W(e(f))} > \tfrac{U(f)-U(e(f))}{U(c_+)-U(e(f))} = 0.$$

Let $c = e(f) + \frac{W(f)-W(e(f))}{W(c_+)-W(e(f))}$. Then $U(c) > U(f)$, $V(c) < V(f)$, and $W(c) = W(f)$ because the functions V and W are linear on $\mathcal{H}_f$. Thus for any sufficiently small $\gamma > 0$, $U(c) > U(f+\gamma)$, $V(f+\gamma) > V(f-\gamma) > V(c)$, and $W(f+\gamma) > W(c) > W(f-\gamma)$. By Lemma A.4(iii),

$$\{c\} \succ \{f+\gamma, c\} \sim \{f+\gamma\} \sim \{f+\gamma, f-\gamma\} \succ \{f-\gamma\}. \tag{A.4}$$

By Lemma A.4(i), g is potentially tempting. Therefore, V and W are linear on $\mathcal{H}_g$ as well; hence, $V(g-\gamma) = V(f-\gamma)$ and $W(g-\gamma) = W(f-\gamma)$. By Lemma A.4(iii) and SSB,

$$\{c, f-\gamma\} \sim \{c, f-\gamma, g-\gamma\} \sim \{c, g-\gamma\} \succ \{g-\gamma\}. \tag{A.5}$$

It follows from SSB and the rankings (A.4) and (A.5) that

$$\{f+\gamma\} \sim \{f+\gamma, c, f-\gamma\} \sim \{f+\gamma, c, f-\gamma, g-\gamma\} \sim \{f+\gamma, f-\gamma, g-\gamma\} \sim \{f+\gamma, g-\gamma\}.$$

Thus $\{f+\gamma\} \sim \{f+\gamma, g-\gamma\} \succ \{g-\gamma\}$. Analogously, $\{g+\gamma\} \sim \{g+\gamma, f-\gamma\} \succ \{f-\gamma\}$. By SSB, $\{f+\gamma\} \cup A \sim \{f+\gamma, g-\gamma\} \cup A$ and $\{g+\gamma\} \cup A \sim \{g+\gamma, f-\gamma\} \cup A$. Let $\gamma \to 0$; then $\{f\} \cup A \sim \{f, g\} \cup A \sim \{g\} \cup A$ by Continuity. □

Lemma A.6. *For all finite menus $A \in \mathcal{A}$, $\mathcal{U}(A) = \mathcal{U}_{WV}(A)$.*

Proof. Fix a finite menu $A \in \mathcal{A}$ and consider several possible cases.

Case 1. $A = \{f, g\}$, where both f and g are never tempting. Wlog $U(f) \geq U(g)$. As $U(g) = U(e(g))$ and $V(g) = \frac{\kappa}{1-\kappa}U(g) = V(e(g))$, then by Lemmas A.5 and A.2,

$$\mathcal{U}(\{f, g\}) = \mathcal{U}(\{f, e(g)\}) = \max\{U(f), U(e(g))\} = U(f).$$

On the other hand, the equality $\mathcal{U}_{WV}(\{f, g\}) = W(f) - V(f)U(f)$ follows from the definitions of the functions V, W, and $\mathcal{U}_{WV}$. Thus, $\mathcal{U}(A) = \mathcal{U}_{WV}(A)$.

Case 2. $A = \{f, g\}$, where f is potentially tempting, and g is never tempting. Then $U(g) = U(e(g))$ and $V(g) = V(e(g))$. By Lemmas A.5 and A.4 (iv),

$$\mathcal{U}(\{f, g\}) = \mathcal{U}(\{f, e(g)\}) = \mathcal{U}_{WV}(\{f, e(g)\}) = \mathcal{U}_{WV}(\{f, g\}).$$

Case 3. $A = \{f, g\}$, where both f and g are potentially tempting. Wlog $U(f) \geq U(g)$. Consider three possible subcases.

Subcase 3.1. $U(f) = U(g)$. Wlog $V(f) \geq V(g)$. By SSB,

$$\mathcal{U}(\{f, g\}) = \mathcal{U}(\{f\}) = W(f) - V(f) = \mathcal{U}_{WV}(\{f, g\}).$$

Subcase 3.2. $U(f) > U(g)$ and $V(g) \geq V(f)$. We claim that there exist $\alpha \in (0, 1)$ and $c \in \mathcal{C}$ such that $U(\alpha g + (1-\alpha)c) = U(f)$ and $V(\alpha g + (1-\alpha)c) = V(f)$. To construct such α and c, let $Y(f) = (1-\kappa)V(f) - \kappa U(f)$ and $Y(g) = (1-\kappa)V(g) - \kappa U(g)$. Then $Y(f)$ and $Y(g)$ are both positive by Lemma A.4(i) and satisfy the identity

$$\kappa[U(f)Y(g) - U(g)Y(f)] = (1-\kappa)[V(f)Y(g) - V(g)Y(f)].$$

The inequalities $U(f) > U(g)$ and $V(g) \geq V(f)$ imply that $Y(g) > Y(f)$, and hence,

$$-1 \leq U(g) < \tfrac{U(f)Y(g) - U(g)Y(f)}{Y(g) - Y(f)} = \tfrac{1-\kappa}{\kappa} \cdot \tfrac{V(f)Y(g) - V(g)Y(f)}{Y(g) - Y(f)} \leq \tfrac{1-\kappa}{\kappa} \cdot V(f) \leq 1.$$

Take $\alpha = \frac{Y(f)}{Y(g)} \in (0, 1)$ and $c \in \mathcal{C}$ such that $U(c) = \frac{U(f)Y(g) - U(g)Y(f)}{Y(g) - Y(f)} \in [-1, 1]$. Then $U(\alpha g + (1-\alpha)c) = U(f)$ by linearity of U, $V(c) = \frac{\kappa}{1-\kappa}U(c) = \frac{V(f)Y(g) - V(g)Y(f)}{Y(g) - Y(f)}$, and hence, $V(\alpha g + (1-\alpha)c) = V(f)$ by linearity of V on $\mathcal{H}_g$.

Conclude by Lemmas A.5 and A.4(iv) that

$$\mathcal{U}(\{f, g\}) = \mathcal{U}(\{\alpha g + (1-\alpha)c, g\}) = \mathcal{U}_{WV}(\{\alpha g + (1-\alpha)c, g\}) = \mathcal{U}_{WV}(\{f, g\}).$$

Subcase 3.3. $U(f) > U(g)$ and $V(f) > V(g)$. As $V(g) > V(e(g))$ and V is linear on $\mathcal{H}_f$, there exists $\alpha \in (0,1)$ such that $V(\alpha f + (1-\alpha)e(g)) = V(g)$. Let $f' = \alpha f + (1-\alpha)e(g)$. As $0 < \alpha < 1$, f' is potentially tempting and satisfies $U(f) > U(f') > U(g)$, $V(f) > V(f') = V(g)$, and $W(f) > W(f') > W(g)$. By Lemma A.4(iii) and by Subcase 3.2,

$$\mathcal{U}(\{f,f'\}) = \mathcal{U}_{WV}(\{f,f'\}) = W(f) - V(f) = U(f) > U(f') \quad \text{and}$$
$$\mathcal{U}(\{f',g\}) = \mathcal{U}_{WV}(\{f',g\}) W(f') - V(f') = U(f') > U(g).$$

By SSB, $\{f,g\} \sim \{f,f',g\} \sim \{f,f'\} \sim \{f\}$. Thus, $\mathcal{U}(\{f,g\}) = U(f) = W(f) - V(f) = \mathcal{U}_{WV}(\{f,g\})$.

Case 4. A is an arbitrary finite menu. Take $g_A \in \arg\max_{f\in A} W(f)$ and $h_A \in \arg\max_{f\in A} V(f)$. Then for all $f \in A$,

$$\mathcal{U}_{WV}(\{g_A,f\}) \geq \mathcal{U}_{WV}(\{g_A,h_A\}) \geq \mathcal{U}_{WV}(\{f,h_A\}).$$

Cases 1–3 imply that $\mathcal{U}(\{g_A,f\}) \geq \mathcal{U}(\{g_A,h_A\}) \geq \mathcal{U}(\{f,h_A\})$, that is, $\{g_A,f\} \succeq \{g_A,h_A\} \succeq \{f,h_A\}$. From SSB, it follows by induction with respect to the size of the set A that

$$A = \bigcup_{f\in A} \{g_A,f\} \succeq \{g_A,h_A\} \succeq \bigcup_{f\in A} \{f,h_A\} = A,$$

that is, $A \sim \{g_A,h_A\}$. Thus, $\mathcal{U}(A) = \mathcal{U}(\{g_A,h_A\}) = \mathcal{U}_{WV}(\{g_A,h_A\}) = \mathcal{U}_{WV}(A)$. $\square$

Lemma A.7. *There exists a convex and closed set Q of probability measures on S such that for all $f \in \mathcal{H}$,*

$$V(f) = \tfrac{\kappa}{1-\kappa} \max_{q\in Q} q \cdot u(f). \tag{A.6}$$

Moreover, Q is unique and $p \in Q$.

Proof. First show that V is monotonic, continuous, and quasi-convex.
Monotonicity: Take any $f, f' \in \mathcal{H}$ such that f dominates f'. For all $\alpha \in (0,1)$, Monotonicity and Lemma A.1 imply that $\{f+\alpha\} \sim \{f+\alpha, f'-\alpha\} \succ \{f'-\alpha\}$. It follows from Lemma A.6 that $V(f+\alpha) \geq V(f'-\alpha)$, that is,

$$\alpha V(c_+) + (1-\alpha)V(f) \geq \alpha V(c_-) + (1-\alpha)V(f').$$

Take $\alpha \to 0$ to deduce that $V(f) \geq V(f')$.
Continuity. Let a sequence of acts f_n converge to f as $n \to \infty$. There exist sequences α_n and β_n both converging to zero such that $f + \alpha_n$ dominates f_n, and f_n dominates $f - \beta_n$. As V is monotonic,

$$\alpha_n V(c_+) + (1-\alpha_n) V(f) \geq V(f_n) \geq \beta_n V(c_-) + (1-\beta_n) V(f).$$

It follows that $V(f) = \lim_{n\to\infty} V(f_n)$.
Quasi-Convexity. Suppose that $V(\alpha f + (1-\alpha)g) > V(f) = V(g)$ for some $f, g \in \mathcal{H}$ and $\alpha \in (0,1)$. Take $c \in \mathcal{C}$ such that

$$V(\alpha f + (1-\alpha)g) > V(c) > V(f) = V(g).$$

Then $V(c) > V(e(f))$ and $V(c) > V(e(g))$. By monotonicity of V, $U(c) > U(f)$, $U(c) > U(g)$ and hence, $U(c) > U(\alpha f + (1-\alpha)g)$. By Lemma A.6, $\{c\} \sim \{c, f\} \succ \{f\}$ and $\{c\} \sim \{c, g\} \succ \{g\}$. However,

$$\{c\} \succ \{c, \alpha f + (1-\alpha)g\} \succeq \{\alpha f + (1-\alpha)g\},$$

contradicting Convex Temptation.

The preceding shows that the ranking on $\mathcal{H}$ represented by V satisfies all the axioms of the maxmax model—these are the axioms of Gilboa and Schmeidler's multiple-priors model [14], with the exception that "Uncertainty Aversion", which is convexity of weakly better-than sets, is replaced by convexity of weakly worse-than sets. It follows from [14] that V has the form (A.6), and that Q is unique. The inclusion $p \in Q$ follows from the fact that for all $f \in \mathcal{H}$, $V(f) \geq V(e(f)) = p \cdot u(f)$. $\square$

The continuity of V and W implies that $\mathcal{U}_{WV}$ is continuous on $\mathcal{A}$. Lemma A.6 asserts that $\mathcal{U} \equiv \mathcal{U}_{WV}$ on the set of all finite menus, which is dense in $\mathcal{A}$. Thus $\mathcal{U} \equiv \mathcal{U}_{WV}$ on all of $\mathcal{A}$.

To show the required uniqueness of (u, p, κ, Q) in representation (2.1)-(2.3), suppose that this tuple can be replaced by (u', p', κ', Q'). Then u' is a positive linear transformation of u, and hence, (u, p, κ, Q) can be replaced by (u, p', κ', Q') as well. The uniqueness statements in Lemmas A.1, A.2 and A.7 imply that if $\succeq$ is not strategically rational, then $p = p'$, $\kappa = \kappa'$, and $Q = Q'$.

B. Appendix: Proofs for Comparative Dissonance

Proof of Theorem 4.2: Let $\succeq^*$ and $\succeq$ conform to our model with corresponding tuples $(u^*, p^*, Q^*, \kappa^*)$ and (u, p, Q, κ). Suppose that neither preference is strategically rational. Then $\kappa, \kappa^* > 0$ and sufficiency of (4.2) and (4.3) is immediate:

$$\{f\} \succ \{f, g\} \quad \Rightarrow \quad [p \cdot u(f) > p \cdot u(g) \wedge Q \cdot u(g) > Q \cdot u(f)] \quad \Rightarrow$$
$$[p^* \cdot u^*(f) > p^* \cdot u^*(g) \wedge Q^* \cdot u^*(g) > Q^* \cdot u^*(f)] \quad \Rightarrow \quad \{f\} \succ^* \{f, g\}.$$

For necessity, let $\succeq^*$ have greater dissonance than $\succeq$. For all vectors $a \in \mathbb{R}^S$, let

$$Q \cdot a = \max_{q \in Q} q \cdot a \quad \text{and} \quad Q^* \cdot a = \max_{q \in Q^*} q \cdot a. \tag{B.1}$$

Lemma B.1. *(i) u and u^* are cardinally equivalent.*
(ii) For all $a, b \in \mathbb{R}^S$,

$$p \cdot a > p \cdot b \text{ and } Q \cdot b > Q \cdot a \quad \Rightarrow \quad p^* \cdot a > p^* \cdot b \text{ and } Q^* \cdot b > Q^* \cdot a. \tag{B.2}$$

(iii) $p = p^*$.

Proof. First, show that for all $c, c' \in \mathcal{C}$,

$$u(c) = u(c') \quad \Rightarrow \quad u^*(c) = u^*(c'). \tag{B.3}$$

Suppose to the contrary that $u(c) = u(c')$ and $u^*(c) > u^*(c')$ for some $c, c' \in \mathcal{C}$. Take $f, g \in \mathcal{H}$ such that $\{f\} \succ \{f, g\}$. The equality $u(c) = u(c')$ implies

$$\{\alpha f + (1-\alpha)c\} \succ \{\alpha f + (1-\alpha)c, \alpha g + (1-\alpha)c'\}.$$

Because $\succeq^*$ has greater dissonance, $\{f\} \succ^* \{f, g\}$. Therefore, the inequality $u^*(c) > u^*(c')$ implies that for sufficiently small $\alpha > 0$,

$$\{\alpha f + (1-\alpha)c\} \sim^* \{\alpha f + (1-\alpha)c, \alpha g + (1-\alpha)c'\}.$$

But this contradicts the hypothesis that $\succeq^*$ has greater dissonance than $\succeq$.

Take $c_+, c_- \in \mathcal{C}$ such that $u(c_+) > u(c_-)$ and $u(c_+) \geq u(c) \geq u(c_-)$ for all $c \in \mathcal{C}$. Then for all $c \in \mathcal{C}$,

$$c \sim \frac{u(c) - u(c_-)}{u(c_+) - u(c_-)} c_+ + \frac{u(c_+) - u(c)}{u(c_+) - u(c_-)} c_-, \quad \text{and by (B.3),}$$
$$u^*(c) = \frac{u^*(c_+) - u^*(c_-)}{u(c_+) - u(c_-)} u(c) + \frac{u^*(c_-)u(c_+) - u^*(c_+)u(c_-)}{u(c_+) - u(c_-)}.$$

Note that $u^*(c_+) \neq u^*(c_-)$ because $\succeq^*$ is not strategically rational and hence non-degenerate. Thus, either u^* is a positive linear transformation of u, or u^* is a negative linear transformation of u. Next, we show that the former case implies statements (ii) and (iii), and that the latter case is impossible.

Case 1. u^* is a positive linear transformation of u. Wlog assume that $u = u^*$ and $u(\mathcal{C}) = u^*(\mathcal{C}) = [-1, 1]$. Fix any $a, b \in \mathbb{R}^S$ such that $p \cdot a > p \cdot b$ and $Q \cdot b > Q \cdot a$. Take $\alpha > 0$ such that $|\ \alpha a(s)\ |,\ \ |\ \alpha b(s)\ |\ \ \leq 1$ for all $s \in S$. Then $\alpha a = u(f)$ and $\alpha b = u(g)$ for some $f, g \in \mathcal{H}$. (Here $u(f)$ and $u(g)$ are vectors in $\mathbb{R}^S$.) Then

$$p \cdot a > p \cdot b \text{ and } Q \cdot b > Q \cdot a \quad \Rightarrow \quad p \cdot u(f) > p \cdot u(g) \text{ and } Q \cdot u(g) > Q \cdot u(f) \quad \Rightarrow$$
$$\{f\} \succ \{f, g\} \quad \Rightarrow \quad \{f\} \succ^* \{f, g\} \quad \Rightarrow$$
$$p^* \cdot u(f) > p^* \cdot u(g) \text{ and } Q^* \cdot u(g) > Q^* \cdot u(f) \quad \Rightarrow \quad p^* \cdot a > p^* \cdot b \text{ and } Q^* \cdot b > Q^* \cdot a,$$

which proves (ii).

To show (iii), suppose that $p \neq p^*$. Let

$$R = \{q \in \mathbb{R}^S : q = p + \alpha(p - p^*) \text{ for } \alpha \geq 0\} = \{q \in \mathbb{R}^S : p \in [q, p^*]\}.$$

Consider two subcases.

(1) $Q \not\subset R$: Let $p' \in Q \backslash R$. Take a hyperplane $b \in \mathbb{R}^S$ that separates the singleton p and the segment $[p', p^*]$:

$$p \cdot b < 0, \quad p' \cdot b > 0, \quad p^* \cdot b > 0.$$

These inequalities violate (B.2) for $a = 0$.

(2) $Q \subset R$: Then Q is a segment with end points p and $p' = p + \alpha(p - p^*)$ for some $\alpha > 0$. Note that p is an interior point of the segment $[p^*, p']$. Take a hyperplane $a \in \mathbb{R}^S$ that separates p^* and p' and passes through p:

$$p^* \cdot a > 0, \quad p \cdot a = 0, \quad p' \cdot a < 0.$$

Take a hyperplane $b \in \mathbb{R}^S$ that separates p' and the segment $[p, p^*]$:

$$p' \cdot b > 0, \quad p \cdot b < 0, \quad p^* \cdot b < 0.$$

Wlog $p^* \cdot a > Q^* \cdot b$ (multiply a by a positive scalar if needed). Thus $p \cdot a > p \cdot b$,

$$Q \cdot b \geq p' \cdot b > 0 = \max\{p \cdot a, p' \cdot a\} = Q \cdot a,$$

but $Q^* \cdot a \geq p^* \cdot a > Q^* \cdot b$. This contradicts (B.2).

Case 2. u^* is a negative linear transformation of u. We show this is impossible.

Wlog assume that $u^* = -u$ and $u(\mathcal{C}) = u^*(\mathcal{C}) = [-1, 1]$. Then, paralleling (B.2) in the previous case,

$$p \cdot a > p \cdot b \text{ and } Q \cdot b > Q \cdot a \quad \Rightarrow \quad p^* \cdot (-a) > p^* \cdot (-b) \text{ and } Q^* \cdot (-b) > Q^* \cdot (-a). \tag{B.4}$$

for all $a, b \in \mathbb{R}^S$. It follows that for some $a, b \in \mathbb{R}^S$, $p \cdot a > p \cdot b$ but $p^* \cdot a < p^* \cdot b$. Thus $p \neq p^*$. Consider two subcases.

(1) $Q \not\subset [p, p^*]$: Let $p' \in Q \backslash [p, p^*]$. Take a hyperplane $b \in \mathbb{R}^S$ that separates p' and $[p, p^*]$:

$$p' \cdot b > 0, \quad p \cdot b < 0, \quad p^* \cdot b < 0.$$

This contradicts (B.4) for $a = 0$.

(2) $Q \subset [p, p^*]$: Then Q is a segment with end points p and $p' = \alpha p^* + (1 - \alpha)p$ for some $\alpha > 0$. Take a hyperplane $a \in \mathbb{R}^S$ that separates p and $[p', p^*]$:

$$p \cdot a = 0, \quad p' \cdot a < 0, \quad p^* \cdot a < 0.$$

Take another hyperplane $b \in \mathbb{R}^S$ that separates p and $[p', p^*]$:

$$p \cdot b < 0, \quad p' \cdot b > 0, \quad p^* \cdot b > 0.$$

Wlog $p^* \cdot (-a) > Q^* \cdot (-b)$ (multiply a by a positive scalar if needed). Then $p \cdot a = 0 > p \cdot b$,

$$Q \cdot b \geq p' \cdot b > 0 = \max\{p' \cdot a, p \cdot a\} = Q \cdot a,$$

but $Q^* \cdot (-a) \geq p^* \cdot (-a) > Q^* \cdot (-b)$. This contradicts (B.4). □

The following method of proof is analogous to the one used by Kopylov [20]. Let $\mathbb{D}$ be the set of all points $a \in \mathbb{R}^S$ at which the convex functions $Q \cdot a$ and $Q^* \cdot a$ are both differentiable. By [30, Theorem 25.5], the complement of the set $\mathbb{D}$ has measure zero. Thus $\mathbb{D}$ is dense. For every $a \in \mathbb{D}$, let

$$q(a) = \nabla(Q \cdot a) \quad \text{and} \quad q^*(a) = \nabla(Q^* \cdot a)$$

be the derivatives of $Q \cdot a$ and $Q^* \cdot a$ respectively. Let $\vec{1} = (1, \ldots, 1) \in \mathbb{R}^S$.

Lemma B.2. *The functions $q(\cdot), q^*(\cdot) : \mathbb{D} \to \mathbb{R}^S$ have the following properties:*

(i) *For all $a \in \mathbb{D}$ and $q \in Q$, $q = q(a)$ iff $Q \cdot a = q \cdot a$.*

(ii) *For all $a \in \mathbb{D}$ and $q \in Q^*$, $q = q^*(a)$ iff $Q^* \cdot a = q \cdot a$.*

(iii) *If $a \in \mathbb{D}$, $\alpha > 0$ and $\gamma \in \mathbb{R}$, then*

$$\alpha a + \gamma \vec{1} \in \mathbb{D}, \quad q(\alpha a + \gamma \vec{1}) = q(a), \quad q^*(\alpha a + \gamma \vec{1}) = q^*(a).$$

(iv) *For any $a \in \mathbb{D}$, there exists $\epsilon_a \in [0, 1]$ such that $q(a) = \epsilon_a q^*(a) + (1 - \epsilon_a)p$.*

(v) *There exists $\epsilon \in [0, 1]$ such that $q(a) = \epsilon q^*(a) + (1 - \epsilon)p$ for all $a \in \mathbb{D}$.*

Proof.

(i) Fix $a \in \mathbb{D}$ and $q \in Q$ such that $Q \cdot a = q \cdot a$. For all $b \in \mathbb{R}^S$ and $\delta \in \mathbb{R}$,

$$Q \cdot a + \delta(q \cdot b) = q \cdot (a + \delta b) \leq Q \cdot (a + \delta b) = Q \cdot a + \delta(q(a) \cdot b) + o(\delta).$$

Then $q \cdot b = q(a) \cdot b$ for all $b \in \mathbb{R}^S$, that is, $q = q(a)$. Similarly for (ii).

(iii) Fix $a \in \mathbb{D}$, $\alpha > 0$ and $\gamma \in \mathbb{R}$. Then $\alpha a + \gamma \vec{1} \in \mathbb{D}$ because the superposition $Q \cdot b = \alpha Q \cdot \left(\frac{b - \gamma \vec{1}}{\alpha}\right) + \gamma$ is differentiable at $\alpha a + \gamma \vec{1}$. By (i), $q(\alpha a + \gamma \vec{1}) = q(a)$ because $Q \cdot (\alpha a + \gamma \vec{1}) = \alpha(Q \cdot a) + \gamma = q(a) \cdot (\alpha a + \gamma \vec{1})$. Similarly for Q^* and $q^*(\cdot)$.

(iv) Suppose that for some a no such ϵ_a exists. Let b separate $q(a)$ from the segment $[q^*(a), p]$, so that $q^*(a) \cdot b < 0$, $p \cdot b < 0$, but $q(a) \cdot b > 0$. Then for sufficiently small $\delta > 0$, $Q^* \cdot (a + \delta b) = Q^* \cdot a + \delta(q^*(a) \cdot b) + o(\delta) < Q^* \cdot a$, but also

$$p \cdot a > p \cdot (a + \delta b) \quad \text{and} \quad Q \cdot (a + \delta b) \geq q(a) \cdot (a + \delta b) > q(a) \cdot a = Q \cdot a.$$

By (B.2), $Q^* \cdot (a + \delta b) > Q^* \cdot a$, a contradiction.

(v) Let $a, b \in \mathbb{D}$ be such that $q^*(a) \neq p$ and $q^*(b) \neq p$, and prove $\epsilon_a = \epsilon_b$. (Note that if $q^*(a) \neq p$, then ϵ_a is unique, and if $q^*(a) = p$, then $\epsilon_a \in [0, 1]$ is arbitrary.) As $q^*(a) \neq p$ and $p = p^* \in Q^*$, then by (iii), $Q^* \cdot a > p \cdot a$. Similarly, $Q^* \cdot b > p \cdot b$. Let

$$a' = \frac{a - (p \cdot a)\vec{1}}{Q^* \cdot a - p \cdot a} \quad \text{and} \quad b' = \frac{b - (p \cdot b)\vec{1}}{Q^* \cdot b - p \cdot b}.$$

By (iii) and (iv), $a', b' \in \mathbb{D}$, $q^*(a') = q^*(a)$, $q^*(b') = q^*(b)$, and

$$q(a') = q(a) = \epsilon_a q^*(a) + (1 - \epsilon_a)p \quad \text{and} \quad q(b') = q(b) = \epsilon_b q^*(b) + (1 - \epsilon_b)p.$$

By construction, $p \cdot a' = p \cdot b' = 0$, $Q^* \cdot a' = Q^* \cdot b' = 1$, $Q \cdot a' = \epsilon_a$, and $Q \cdot b' = \epsilon_b$. Suppose that $\epsilon_a \neq \epsilon_b$; wlog let $\epsilon_a < \epsilon_b$. Then for sufficiently small $\gamma > 0$,

$$p \cdot (a' + \gamma \vec{1}) = \gamma > p \cdot b', \quad Q \cdot (a' + \gamma \vec{1}) = \epsilon_a + \gamma < \epsilon_b = Q \cdot b',$$

but $Q^* \cdot (a' + \gamma \vec{1}) = 1 + \gamma > Q^* \cdot b'$. This contradicts (B.2). Thus $\epsilon_a = \epsilon_b$. □

Conclude that $Q \cdot a = \epsilon(Q^* \cdot a) + (1-\epsilon)(p \cdot a)$ for all $a \in \mathbb{D}$ and hence, by continuity, for all $a \in \mathbb{R}^S$. It follows that $Q = \epsilon Q^* + (1-\epsilon)p$; $\epsilon > 0$ because $\succeq$ is not strategically rational. This completes the proof of Theorem 4.2.

Proof of Theorem 4.3: Let $\succeq^*$ and $\succeq$ conform to our model with corresponding tuples $(u^*, p^*, Q^*, \kappa^*)$ and (u, p, Q, κ). Suppose that neither preference is strategically rational.

Let $P = (1-\kappa)\{p\} + \kappa Q$ and $P^* = (1-\kappa^*)\{p^*\} + \kappa^* Q^*$. The conditions (4.2), (4.3), and $\kappa^* \geq \epsilon\kappa$ imply

$$P = \left(1 - \frac{\epsilon\kappa}{\kappa^*}\right)\{p\} + \frac{\epsilon\kappa}{\kappa^*}P^*.$$

Sufficiency of these conditions now follows from:

$$\begin{aligned}\{f\} \succ \{f,g\} \sim \{g\} \quad &\Rightarrow \quad [p \cdot u(f) > p \cdot u(g) \wedge P \cdot u(g) > P \cdot u(f)] \quad \Rightarrow \\ &[p^* \cdot u^*(f) > p^* \cdot u^*(g) \wedge P^* \cdot u^*(g) > P^* \cdot u^*(f)] \quad \Rightarrow \quad \{f\} \succ^* \{f,g\} \sim^* \{g\}.\end{aligned}$$

For necessity, let $\succeq^*$ be more self-justifying than $\succeq$. Then $\succeq^*$ has more dissonance than $\succeq$, and Theorem 4.3 implies (4.2) and (4.3). Moreover, for all $a, b \in \mathbb{R}^S$,

$$p \cdot a > p \cdot b \text{ and } P \cdot b > P \cdot a \quad \Rightarrow \quad p^* \cdot a > p^* \cdot b \text{ and } P^* \cdot b > P^* \cdot a. \tag{B.5}$$

To prove this claim, fix any $a, b \in \mathbb{R}^S$. Take $\alpha > 0$ and $f, g \in \mathcal{H}$ such that $\alpha a = u(f)$ and $\alpha b = u(g)$. Then

$$\begin{aligned}p \cdot a > p \cdot b \text{ and } P \cdot b > P \cdot a \quad &\Rightarrow \quad p \cdot u(f) > p \cdot u(g) \text{ and } P \cdot u(g) > P \cdot u(f) \quad \Rightarrow \\ \{f\} \succ \{f,g\} \sim \{g\} \quad &\Rightarrow \quad \{f\} \succ^* \{f,g\} \sim^* \{g\} \quad \Rightarrow \\ p^* \cdot u(f) > p^* \cdot u(g) \text{ and } P^* \cdot u(g) > P^* \cdot u(f) \quad &\Rightarrow \quad p^* \cdot a > p^* \cdot b \text{ and } P^* \cdot b > P^* \cdot a.\end{aligned}$$

Use the condition (B.5) to replace Q and Q^* by P and P^* in Lemma B.2 and obtain $0 < \theta \leq 1$ such that $P = (1-\theta)\{p\} + \theta P^*$. In particular, $P \subseteq P^*$ and therefore also

$$(1-\kappa\epsilon)\{p\} + \kappa\epsilon Q^* \subseteq (1-\kappa^*)\{p\} + \kappa^* Q^*.$$

As Q^* is a nonsingleton, $\kappa\epsilon \leq \kappa^*$. $\square$

References

[1] G. Akerlof and W.T. Dickens, The economic consequences of cognitive dissonance, *Amer. Econ. Rev.* 72 (1982), 307-319.

[2] C.D. Aliprantis and K.C. Border, *Infinite Dimensional Analysis*, Springer, 1994.

[3] E. Aronson, *The Social Animal*, 7th edition, W.H. Freeman, New York, 1996.

[4] R. Baumeister, The optimal margin of illusion, *J. Soc. and Clinical Psych.* 8 (1989), 176-189.

[5] S.H. Chew, A generalization of the quasilinear mean with application to the measurement of income inequality and decision theory resolving the Allais Paradox, *Econometrica* 51 (1983), 1065-1092.

[6] E. Dekel, B. Lipman and A. Rustichini, A unique subjective state space for unforeseen contingencies, *Econometrica* 69 (2001), 891-934.

[7] E. Dekel, B. Lipman and A. Rustichini, Temptation driven preferences, 2005.

[8] W.T. Dickens, Crime and punishment again: the economic approach with a psychological twist, *J. Pub. Econ.* 30 (1986), 97-107.

[9] J.H. Dreze, Decision theory with moral hazard and state-dependent preferences, Ch. 2 in *Essays on Economic Decisions under Uncertainty*, Cambridge, U.K. 1987.

[10] L.G. Epstein, An axiomatic model of non-Bayesian updating, *Rev. Econ. Stud.*, forthcoming.

[11] L.G. Epstein and M. Marinacci, Coarse contingencies, 2005.

[12] H. Ergin and T. Sarver, A unique costly contemplation representation, 2005.

[13] L. Festinger, *A Theory of Cognitive Dissonance*, Stanford U. Press, 1957.

[14] I. Gilboa and D. Schmeidler, Maxmin expected utility with non-unique priors, *J. Math. Econ.* 18 (1989), 141-153.

[15] W.N. Goetzmann and N. Peles, Cognitive dissonance and mutual fund investors, *J. Financial Research* 20 (1997), 145-158.

[16] F. Gul and W. Pesendorfer, Temptation and self-control, *Econometrica* 69 (2001), 1403-35.

[17] R. Haagsma and N. Koning, Endogenous mobility-reducing norms, *J. Econ. Behavior and Organization* 49 (2002), 523-547.

[18] R.E. Knox and J.A. Inkster, Postdecision dissonance and post time, *J. Personality and Social Pysch.* 8 (1968), 319-323.

[19] I. Kopylov, Temptations in general settings, 2005.

[20] I. Kopylov, A parametric model of ambiguity aversion, 2005.

[21] D. Kreps, A representation theorem for 'preference for flexibility', *Econometrica* 47 (1979), 565-577.

[22] D. Kreps, *Notes on the Theory of Choice*, Westview, 1988.

[23] D. Kreps, Static choice in the presence of unforeseen contingencies, in *Essays in Honour of F. Hahn*, P. Dasgupta et al eds., MIT Press, 1992.

[24] M. Machina, Temporal risk and the nature of induced preferences, *J. Econ. Theory* 33 (1984), 199-231.

[25] J. Noor, Temptation, welfare and revealed preference, Boston U., 2005.

[26] J. Noor, Menu-dependent self-control, Boston U., 2006.

[27] R.J. Oxoby, Cognitive dissonance, status and growth of the underclass, *Economic Journal* 114 (2004), 727-749.

[28] M. Rabin, Cognitive dissonance and social change, *J. Econ. Behav. and Organization* 23 (1994), 177-194.

[29] M. Rabin and J.L. Schragg, First impressions matter: a model of confirmatory bias, *Quart. J. Econ.* 114 (1999), 37-82.

[30] R.T. Rockafeller, *Convex Analysis*, Princeton U. Press, 1972.

[31] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica* 57, 571-587, 1989.

[32] J. Smith, Cognitive dissonance and the overtaking anomaly: psychology in the principal-agent relationship, Princeton, 2004.

[33] R.H. Strotz, Myopia and inconsistency in dynamic utility maximization, *Rev. Ec. Studies* 23 (1955), 165-180.

[34] S.E. Taylor, *Positive Illusions: Creative Self-Deception and the Healthy Mind*, Basic Books, New York: 1989.

[35] S.E. Taylor and J.B. Brown, Illusion and well-being: a social psychological perspective on mental health, *Psych. Bulletin* 103 (1988), 193-210.