Half empty, half full and why we can *agree to disagree* forever*

Alexander Zimper[†]

Revised version: June 5, 2008

Abstract

Aumann (1976) derives his famous *we cannot agree to disagree* result under the assumption that people are expected utility (=EU) decision makers. Motivated by empirical evidence against EU theory, we study the possibility of *agreeing to disagree* within the framework of Choquet expected utility (=CEU) theory which generalizes EU theory by allowing for ambiguous beliefs. As our first main contribution, we show that people may well *agree to disagree* if their Bayesian updating of ambiguous beliefs is psychologically biased in our sense. Remarkably, this finding holds regardless of whether people with identical priors apply the same psychologically biased Bayesian update rule or not. As our second main contribution, we develop a formal model of Bayesian learning under ambiguity. As a key feature of our approach the posterior subjective beliefs do, in general, not converge to "true" probabilities which is in line with psychological evidence against converging learning behavior. This finding thus formally establishes that CEU decision makers may even agree to disagree in the long-run despite the fact that they always received the same information.

Keywords: Common Knowledge, No-Trade Results, Bayesian Learning, Ambiguity, Choquet Expected Utility Theory, Dynamic Inconsistency

JEL Classification Numbers: D81, G20.

---

*I would like to thank Jayant Ganguli, Alexander Ludwig, Elias Khalil and two anonymous referees for helpful comments and suggestions. Financial support from ERSA (Economic Research South Africa) is gratefully acknowledged.

[†]Department of Economics and Econometrics, University of Johannesburg, PO Box 524, Auckland Park, 2006, South Africa. E-mail: azimper@uj.ac.za

1 Introduction

Aumann (1976) proves that "If two people have the same priors, and their posteriors for an event A are common knowledge, then these posteriors are equal" (p. 1236). This celebrated *we cannot agree to disagree* result has been derived under the implicit assumption that people are subjective expected utility (EU) maximizers whose beliefs are given as additive probability measures. While there exist several studies on the possibility of agreeing-to-disagree in decision-theoretic frameworks that generalize EU theory (cf. section 2 "Related literature" in this paper), to the best of our knowledge none of these studies addresses the possibility that people may interpret information in different ways.

The contributions of this paper are two-fold. As our first contribution, we demonstrate that people may agree to disagree if their beliefs express ambiguity attitudes as, e.g., elicited in paradoxes of the Ellsberg type. This result holds regardless of whether people receive the same information or not since our decision theoretic framework allows for the possibility that identical information may be interpreted differently. More precisely, we differentiate between optimistically, resp. pessimistically, Bayesian updating of ambiguous beliefs. With these definitions we formally describe the difference between "half empty" versus "half full" attitudes in the context of interpreting new information. In order to investigate the question whether agents may forever agree to disagree if they are always fed the same information, we develop, as our second contribution, a formal model of Bayesian learning with ambiguous beliefs.

Key to our analysis is the assumption that people are Choquet expected utility (CEU) rather than EU decision-makers. CEU theory (Schmeidler 1989, Gilboa 1987) is a generalization of EU theory that admits for the integration of a vNM function with respect to non-additive probability measures (capacities). Properties of such capacities are used for the formal description of ambiguity attitudes which may explain Ellsberg (1961) paradoxes. Ellsberg paradoxes demonstrate systematic violations of Savage's (1954) "sure thing principle". The sure thing principle, however, ensures that there is a unique way of deriving ex-post preferences from ex-ante preferences, implying a unique Bayesian update rule for the additive probabilities of subjective EU theory. The picture is different for the non-additive probability measures of CEU theory for which several perceivable Bayesian update rules exist (cf. Gilboa and Schmeidler 1993, Sarin and Wakker (1998), Eichberger, Grant and Kelsey 2006, Siniscalchi 2001, 2006). Following Gilboa and Schmeidler's (1993) psychological interpretation we consider the extreme cases of the optimistic, resp. pessimistic, update rule, which we apply to non-additive probability measures defined as neo-additive capacities in the sense of Chateauneuf, Eichberger and Grant (2006). Our resulting definition of optimistically, resp. pessimistically, biased agents combines the standard model of rational Bayesian learning with an optimistic, resp. pessimistic, attitude towards the interpretation of new information.

We present two different results of the type that people may agree to disagree if their update rules are psychologically biased. Our first result (proposition 1) shows that if two people have the same prior, apply different updating rules, and their posteriors for an event $A \notin \{\varnothing, \Omega\}$ are common knowledge, then these posteriors will be different even in case they have identical information partitions. Within the appropriate framework this result is easily derived. However, beyond the mere formal result our finding addresses an important behavioral issue. Aumann (1976) writes "In private conversation, Tversky has suggested that people may also be biased because of psychological factors, that may make them disregard information that is unpleasant or does not conform to previously formed notions" (p. 1238). There is no way of describing such psychological biases of real-life people within Aumann's framework. Within our approach, however, the resulting "myside bias" has a straightforward interpretation as people's different attitudes towards the interpretation of information due to psychological predispositions such as the "half-empty glass" versus the "half-full glass" attitude.

Whereas our first result applies to people who use different rules of Bayesian updating, our second result (proposition 2) refers to the case of identical updating rules. We find that if two people have the same prior, apply the same learning rule, and their posteriors for an event $A \notin \{\varnothing, \Omega\}$ are common knowledge, then these posteriors can be different in case they have different information partitions. Thus, neither in the case where people have the same information partitions nor in the case where people apply the same update rule does Aumann's conclusion obtain when Bayesian learning is psychologically biased in our sense. To the contrary, according to our results a difference in posteriors that are common knowledge is the rule rather than the exception when people are psychologically biased.

Standard models of Bayesian learning with additive beliefs show that people's beliefs must converge in the long-run to the same belief if they observe identical information drawn from an i.i.d. stochastic process. If this convergence result also holds true in our environment of CEU decision-makers, the relevance of proposition 1 would be restricted to a short-run argument only. In order to investigate the long-run relevance of proposition 1 we therefore develop a model of Bayesian learning based on our decision-theoretic framework. In contrast to the standard model of rational Bayesian learning (e.g., Tonks 1983, Viscusi and O'Connor 1984, Viscusi 1985), which obtains as a special case of our model, the posterior beliefs of our learning model do in general not converge to "true" probabilities. As a consequence, agreeing to disagree becomes possible in our framework even in the long run. This finding also contributes to the discussion about the plausibility of the common priors assumption which typically presumes (cf. Aumann 1987, 1998, Gul 1998) that common priors are justified for agents with symmetric information. Our finding demonstrates that this is not necessarily the case for ambiguous beliefs.

Our model of Bayesian learning with respect to ambiguous beliefs is also in line with several studies in the psychological literature which show that real-life agents' learning behavior does not necessarily imply convergence since learning may be prone to effects such as "myside bias" or "irrational belief persistence" (cf., e.g., the references in chapter 9, Baron 2007). For example, in an early contribution to this literature, Lord, Ross, and Lepper (1979) conduct an experiment in which agents' posteriors diverge despite the fact that all agents have received the same information.[1] Moreover, definitions of several psychological phenomenons such as delusions, depressions etc. are based on the observation that different subjects may interpret identical information in different ways (cf. Beck 1976). We regard the findings in this psychological literature as further evidence in support of this paper's main theme; namely, that the interpretation of new information may be prone to some psychological bias. If this is true, Aumann's *we cannot agree to disagree* result would apply to idealized rather than to real-life decision-makers even in case the common priors assumption is satisfied.

The subsequent analysis is structured as follows. Section 2 discusses the relationship of our approach to the existing literature. In section 3 we describe our decision-theoretic framework. Section 4 recalls Aumann's (1976) epistemic framework and presents our first *agreeing to disagree* result. A simple example in section 5 about the possibility of ex-post asset trade illustrates this first result. In section 6 we introduce our model of Bayesian learning under ambiguity which demonstrates that our first result is also relevant in the long-run. Our second *agreeing to disagree* result is stated and proved in section 7. Section 8 concludes.

2 Related literature

2.1 No-trade results

Combined with Harsanyi's (1967) common priors doctrine Aumann's *we cannot agree to disagree* result has been very influential in information economics. Especially the so-called no-trade theorems - basically stating that there should be no ex-post trade

[1] The subjects in this experiment were confronted with two purported statistical studies, one study supporting the other study rejecting the claim that capital punishment has a crime deterrence effect. For analogous results in the context of Bayesian updating of subjective probabilities see Pitz, Downing, and Reinhold (1967), Pitz (1969), and Chapman (1973).

in financial assets if the agents are rational - are based on Aumann's approach (cf., e.g., Milgrom 1981, Milgrom and Stokey 1982, Samet 1990, Morris 1994, Bonanno and Nehring 1999). The connection between Aumann's *we cannot agree to disagree* result and the impossibility of ex-post trade in financial assets is straightforward. Under the assumption that agents have different preferences for such assets if and only if they have different beliefs about the assets' future returns, there are strict incentives for ex-post trade if and only if the agents have different posterior beliefs. Since the market-price of such assets is common knowledge between the trading agents, any trade would result in the traders' common knowledge that their posteriors must be different.[2]

Since no trade-results are seemingly at odds with reality, there are several contributions in the literature investigating the robustness of no-trade results with respect to a weakening of Aumann's assumptions. One line of research discusses concepts of bounded rationality that weaken the rationality assumptions of Aumann's epistemic framework. For example, information structures have been considered that are non-partitional (Bacharach 1985, Samet 1990, Geneakoplos 1992, Rubinstein and Wolinsky 1990) or concepts of "almost" common-knowledge have been introduced (Neeman 1996). In contrast to this literature our approach fully adopts Aumann's epistemic framework. The agents of our model are boundedly rational not with respect to their logical capability but with respect to their psychological bias in interpreting new information.

Closer to our own approach is a second line of research on no-trade results that considers decision theoretic alternatives to EU theory. In an early contribution Dow, Madrigal and Werlang (1990) already provide an example in which ex-post trading becomes possible because agents update their non-additive beliefs according to the Dempster-Shafer rule which is at the heart of our definition of pessimistically biased Bayesian learning.[3] Dow et al. thereby assume asymmetric information and common non-additive priors so that their example can be regarded as an illustration of our proposition 2 for the special case of pessimistically biased agents.

Halevy (1998, 2004) claims that the finding of Dow et al. can be extended to the case of symmetric information so that there might occur ex-post trading between agents with common priors and identical information partitions if their beliefs are non-additive. More precisely, Halevy writes:

> "A similar result appears in Dow et al (1990). Their result, as noted by Epstein and Le Breton (1993) and as our present example illustrates, relies

[2] Note that the typical assumption of "strictly risk averse" traders (e.g., Milgrom 1981, Milgrom and Stokey 1982) is thus not necessary for obtaining "no-trade" results.

[3] Within the multiple prior framework Kajii and Ui (2007) show that full-Bayesian updating may result (under asymmetric information) in different sets of posteriors whose intersection is non-empty.

merely on dynamic inconsistency of the individual agents. Their claim that trade is a result of asymmetric information is not accurate: we show below that it could be reached with completely symmetric information and even with a common prior." (footnote 17, p. 20 in Halevy 1998)

In the light of our propositions 1 and 2, we take a somewhat different view from Halevy. Namely, Dow et al.'s conclusion is indeed accurate under their assumption of an identical update rule for all agents: our analysis demonstrates that agents with an identical update rule cannot agree to disagree if they have identical information partitions. Our own asset-trade example in section 5 of this paper therefore establishes the existence of ex-post trade between agents with symmetric information and common priors if and only if the agents have different update rules. Since Halevy's example does not consider different update rules, his finding appears to be at odds with our own results. As it turns out, however, the difference between Dow et al. and our conclusion, on the one hand, and Halevy's conclusion, on the other hand, is only due to different notions of beliefs arising from the fact that Halevy considers an environment with objective probability distributions, i.e., lotteries, whereas Dow et al. and our approach work within the subjective Savage-framework.[4] More specifically, Halevy's example is based on Yaari's (1987) dual theory in which objective additive probabilities are transformed into non-additive beliefs by some transformation function so that he can separate between (objective) probabilities and ambiguity attitudes arising from the transformation. Halevy then speaks of common priors whenever his agents have common objective probabilities. Since these agents apply different transformation functions, however, their resulting non-additive priors are no longer identical with respect to our notion of subjective beliefs which incorporates ambiguity attitudes. According to our and Dow et al.'s notion of subjective beliefs the assumption of common priors would therefore be violated in Halevy's example.

Rubinstein and Wolinsky (1990, Remark p. 190) argue that Milgrom and Stokey's no-trade result applies to all decision theories under uncertainty which satisfy dynamic consistency. Similarly, Ma (2001) (for general preferences over Savage-acts), Kajii and Ui (2007) (for the multiple prior framework) and Wakai (2001) (for the recursive multiple prior framework) establish formal links between dynamic consistency and the impossibility of speculative trade. Halevy (2004) reports the interesting fact that there might even be ex-post trading between dynamically consistent agents if these agents violate consequentialism. While EU decision-makers satisfy, by the sure-thing principle, dynamic consistency as well as consequentialism, the CEU decision-makers of our model

[4] I am grateful to an anonymous referee for pointing me to this interpretation of Halevy's approach.

only satisfy consequentialism and the possibility of agreeing to disagree exclusively results from their dynamically inconsistent preferences (cf. Epstein and Le Breton 1993, Sarin and Wakker 1998).

2.2 Learning under ambiguity

Epstein and Schneider (2007) also consider a model of learning under ambiguity which shares with our learning model the feature that ambiguity does not necessarily vanish in the long run. Their learning model is based on the *recursive multiple priors* approach (Epstein and Wang 1994; Epstein and Schneider 2003) which, basically, restricts conditional *max min expected utility* (MMEU) preferences of Gilboa and Schmeidler (1989) in such a way that dynamic consistency is satisfied. While MMEU theory is closely related to CEU theory restricted to *convex* capacities (e.g., neo-additive capacities for which the degree of optimism is zero), the similarity between Epstein and Schneider's approach on the one hand and our learning model on the other hand ends here. Epstein and Schneider establish ambiguity in the long run under the assumption that the decision-maker permanently receives ambiguous signals, which they formalize via a multitude of different likelihood functions at each information stage in addition to the existence of multiple priors.[5] This introduction of multiple likelihoods is rather ad hoc and it would be interesting to see an axiomatic or/and psychological foundation of this approach which goes beyond the mere technical property that multiple likelihoods can sustain long-run ambiguity in the recursive multiple priors framework. In the meantime, our - comparably simple - model of a Bayesian learner who is prone to psychological attitudes in the interpretation of new information offers a rather straightforward explanation for biased long-run beliefs even in the case that the decision-maker receives signals that are not ambiguous. Finally notice that the restriction of Epstein and Schneider's approach to dynamically consistent preferences excludes preferences that violate Savage's sure-thing principle as elicited in Ellsberg paradoxes. More specifically, by observation 2 in this paper we show that any updating rule for preferences over Savage-acts must be dynamically inconsistent if the ex-ante preferences strictly violate the sure-thing principle. Since our learning model does not exclude dynamically inconsistent decision behavior, it can accommodate a broader notion of ambiguity attitudes than the Epstein-Schneider approach, including ambiguity attitudes that are not compatible with the sure-thing principle.

Marinachi (1999) studies a decision-maker who repeatedly observes an experiment such that the outcomes at each trial are identically and independently distributed with

[5] In the case of learning from ambiguous urns without multiple likelihoods, ambiguity obviously vanishes in the learning process; (for a formal result see Marinachi 2002).

respect to the decision-maker's non-additive belief; (Marinachi thereby defines independent events as events such that the product of their marginal non-additive probabilities coincides with the non-additive probability of their intersection.) For such i.i.d. random variables with respect to non-additive beliefs Marinachi derives a law of large numbers as counterpart to the additive case thereby admitting for the possibility that ambiguity does not vanish in the long-run. While Marinachi's approach may thus be regarded as a Frequentist's approach towards non-additive probabilities, our approach is a subjectivist Bayesian one according to which an agent has a subjective prior belief over the whole event space while she uses sample information from an (objective) i.i.d. process in order to update her subjective belief. In contrast to Marinachi's approach, the distributions of outcomes in different Bernoulli trials are, in general, not i.i.d. with respect to the non-additive posteriors of our Bayesian approach. Unlike in our approach the learning behavior of different agents in Marinachi's model must therefore converge to the same limit if they have identical priors.

3 Preliminaries: The decision-theoretic framework

3.1 Choquet decision theory

As in Aumann (1976) we consider a measurable space $(\Omega, \mathcal{F})$ with $\mathcal{F}$ denoting a σ-algebra on the state space Ω. As a generalization of Aumann's assumption of EU decision-makers, however, we consider a CEU rather than an EU decision-maker.[6] In contrast to EU theory, CEU theory admits for non-additive probability measures, i.e., capacities, whereby a capacity $\nu : \mathcal{F} \rightarrow [0,1]$ must satisfy

(i) $\nu(\emptyset) = 0$, $\nu(\Omega) = 1$

(ii) $A \subset B \Rightarrow \nu(A) \leq \nu(B)$ for all $A, B \in \mathcal{F}$.

Additional properties of capacities are used in the literature for formal definitions of, e.g., *ambiguity* and *uncertainty attitudes* (Schmeidler 1989, Epstein 1999, Ghirardato and Marinacchi 2002), *pessimism* and *optimism* (Eichberger and Kelsey 1999, Wakker 2001), as well as *sensitivity to changes in likelihood* (Wakker, 2004).

[6] CEU theory was first axiomatized by Schmeidler (1986, 1989) within the Anscombe and Aumann (1963) framework, which assumes preferences over objective probability distributions. Subsequently, Gilboa (1987) as well as Sarin and Wakker (1992) have presented CEU axiomizations within the Savage (1954) framework, assuming a purely subjective notion of likelihood. CEU theory is equivalent to *cumulative prospect theory* (Tversky and Kahneman 1992, Wakker and Tversky 1993) restricted to the domain of gains (compare Tversky and Wakker 1995). Moreover, as a representation of preferences over lotteries, CEU theory coincides with *rank dependent utility theory* as introduced by Quiggin (1981, 1982).

In our model of non-rational Bayesian learning we restrict attention to a class of capacities that are defined as *neo-additive capacities* in the sense of Chateauneuf, Eichberger, and Grant (2006). Neo-additive capacities stand for deviations from additive probabilities such that a parameter δ (*degree of ambiguity*) measures the lack of confidence the decision-maker has in some subjective additive probability measure μ. In addition, a second parameter λ measures the degree of optimism versus pessimism by which the decision-maker resolves his ambiguity.

Definition: *A* neo-additive capacity, ν, *is defined, for some* $\delta, \lambda \in [0,1]$, *by*

$$\nu(A) = \delta\left(\lambda\omega^{o}(A) + (1-\lambda)\,\omega^{p}(A)\right) + (1-\delta)\,\mu(A) \tag{1}$$

for all $A \in \mathcal{F}$ *such that* π *is some additive probability measure and we have for the non-additive capacities* ω^{o}

$$\begin{aligned} \omega^{o}(A) &= 1 \text{ if } A \neq \emptyset \\ \omega^{o}(A) &= 0 \text{ if } A = \emptyset \end{aligned} \tag{2}$$

and ω^{p} *respectively*

$$\begin{aligned} \omega^{p}(A) &= 0 \text{ if } A \neq \Omega \\ \omega^{p}(A) &= 1 \text{ if } A = \Omega. \end{aligned} \tag{3}$$

Recall that the Choquet integral of a bounded function $f : \Omega \to \mathbb{R}$ with respect to capacity ν is defined as the following Riemann integral extended to domain Ω (Schmeidler 1986):

$$E[f,\nu] = \int_{-\infty}^{0} \left(\nu\left(\{s \in \Omega \mid f(s) \geq z\}\right) - 1\right) dz + \int_{0}^{+\infty} \nu\{s \in \Omega \mid f(s) \geq z\}\, dz. \tag{4}$$

The following observation extends a result (Lemma 3.1) of Chateauneuf, Eichberger, and Grant (2007) for finite random variables to the more general case of random variables with a closed and bounded range.

Observation 1. *Let* f *be a real-valued function with closed and bounded range. Then the Choquet expected value (4) of* f *with respect to a neo-additive capacity (1) is given by*

$$E[f,\nu] = \delta\left(\lambda \max f + (1-\lambda)\min f\right) + (1-\delta)\,E[f,\mu]\,. \tag{5}$$

Proof: Relegated to the appendix.

Let f denote a Savage-act, i.e., a mapping from the state space Ω into some set of consequences X. The *Choquet expected utility* (Schmeidler 1989; Gilboa 1987) of a bounded and closed Savage-act f with respect to a neo-additive capacity is

$$E\left[u\left(f\right),\nu\right]=\delta\left(\lambda\max_{\omega\in\Omega}u\left(f\left(\omega\right)\right)+\left(1-\lambda\right)\min_{\omega\in\Omega}u\left(f\left(\omega\right)\right)\right)+\left(1-\delta\right)E\left[u\left(f\right),\mu\right],\quad(6)$$

with $u:X\rightarrow\mathbb{R}$ denoting a von Neumann-Morgenstern utility function.[7] Obviously, in case there is no ambiguity, i.e., $\delta=0$, (6) reduces to the standard subjective expected utility representation of preferences over Savage-acts.

3.2 Bayesian updating of capacities

CEU theory has been developed in order to accommodate paradoxes of the Ellsberg type which show that real-life decision-makers violate Savage's *sure-thing principle*. In this subsection we demonstrate that the abandoning of the sure-thing principle bears two important implications for conditional CEU preferences over Savage-acts. Firstly, in contrast to Bayesian updating of additive probability measures, there exist several perceivable Bayesian update rules for non-additive probability measures (cf. Gilboa and Schmeidler 1993, Sarin and Wakker 1998, Pires 2002, Eichberger, Grant and Kelsey 2006, Siniscalchi 2001, 2006). Secondly, any preferences that (strictly) violate the sure-thing principle cannot be updated in a dynamically consistent way. That is, there does not exist any updating rule for capacities such that ex-ante CEU preferences that (strictly) violate the sure-thing principle are updated in a dynamically consistent manner to ex-post CEU preferences.

To see this define the Savage-act $f_B h:\Omega\rightarrow X$ such that

$$f_B h\left(\omega\right)=\begin{cases}f\left(\omega\right) & \text{for }\omega\in B\\ h\left(\omega\right) & \text{for }\omega\in\neg B\end{cases}$$

where B is some non-empty event. Recall that Savage's sure-thing principle states that, for all acts f,g,h,h' and all events $B\in\mathcal{F}$,

[7]Ludwig and Zimper (2006a) show that the CEU of an act with respect to a neo-additive capacity can be equivalently described by the *α-maxmin expected utility with respect to multiple priors* (α-MEU) of an act which encompasses the original multiple priors approach of Gilboa and Schmeidler (1989) as a special case (see, e.g., Ghirardato et al., 1998; Ghirardato et al., 2004; Siniscalchi, 2005). In particular, we have equivalence between the CEU with respect to neo-additive capacities and the α-MEU with respect to so-called *ε-contaminated priors* used in Bayesian statistics (Berger and Berliner, 1986) that may be interpreted as neo-additive capacities.

$$f_B h \succeq g_B h \text{ implies } \; f_B h' \succeq g_B h'. \tag{7}$$

Let us interpret event B as new information received by the agent. The sure-thing principle then implies a straightforward way for deriving ex-post preferences $\succeq_B$, conditional on the new information B, from the agent's original preferences $\succeq$ over Savage-acts. Namely, we have

$$f \succeq_B g \text{ if and only if } f_B h \succeq g_B h \text{ for any } h, \tag{8}$$

implying for a subjective EU decision-maker

$$f \succeq_B g \Leftrightarrow E\left[u\left(f\right), \mu\left(\cdot \mid B\right)\right] \geq E\left[u\left(g\right), \mu\left(\cdot \mid B\right)\right]$$

where $\mu\left(\cdot \mid B\right)$ is a conditional additive probability measure defined, for all $A, B \in \mathcal{F}$ such that $\mu\left(B\right) > 0$, by

$$\mu\left(A \mid B\right) = \frac{\mu\left(A \cap B\right)}{\mu\left(B\right)}.$$

It is well known that the updating of EU preferences satisfies *dynamic consistency*, which - informally - states that there are no strict ex-post incentives for deviating from an ex-ante optimal plan of actions. Formally, we define dynamic consistency in terms of update rules, i.e., rules that derive conditional preferences, $\{\succeq_B\}$ for all events B, from an ex-ante preference ordering $\succeq$.

Definition: Dynamic Consistency. *We speak of a dynamically consistent update rule if for all partitions $\mathcal{P} \subseteq \mathcal{F}$ and all Savage-acts f, g, $f \succeq_B g$ for all $B \in \mathcal{P}$ implies $f \succeq g$.*

Observation 2. *There does not exist any dynamically consistent update rule for preferences $\succeq$ that strictly violate the sure-thing principle.*

Proof: For preferences that strictly violate the sure-thing principle we have, for some f and g,

$$f_B h \succ g_B h \text{ and } g_B h' \succ f_B h' \text{ for some } h \neq h' \text{ and some } B.$$

Observe that any update rule for preferences must result in conditional preferences $f \succeq_B g$ or $g \succeq_B f$. Consider at first the case $f \succeq_B g$. Since $h' \succeq_{\neg B} h'$, dynamic consistency implies $f_B h' \succeq g_B h'$, a contradiction to $g_B h' \succ f_B h'$ by the definition of a preference ordering. Now consider the case $g \succeq_B f$. Since $h \succeq_{\neg B} h$, dynamic consistency implies $g_B h \succeq f_B h$, a contradiction to $f_B h \succ g_B h$.□

In case the sure-thing principle does not hold, the specification of act h in (8) is no longer arbitrary so that there exist for CEU preferences several possibilities of deriving ex post preferences from ex ante preferences. Following Gilboa and Schmeidler (1993) we focus on so-called *h-Bayesian update rules* for preferences $\succeq$ over Savage acts. That is, we consider some collection of conditional preference orderings, $\{\succeq_B^h\}$ for all events B, such that for all acts f, g

$$g \succeq_B^h h \Leftrightarrow f_B h \succeq g_B h. \tag{9}$$

In the light of observation 2, none of these h-update rules is dynamically consistent if the ex-ante preferences strictly violate Savage's sure-thing principle. Gilboa and Schmeidler show that CEU preferences $\succeq$ on Savage acts are updated to conditional CEU preferences $\{\succeq_B^h\}$ for all events B if and only if h is an act such that for some event $E \in \mathcal{F}$

$$h = (x^*, E; x_*, \neg E), \tag{10}$$

where x^* denotes the best and x_* denotes the worst consequence possible. The different possible specifications of E in (10) can result in a multitude of different h-Bayesian update rules. For example, for the so-called *optimistic* update rule h is the constant act where $E = \emptyset$. That is, under the optimistic update rule the null-event becomes associated with the worst consequence possible. Gilboa and Schmeidler (1993) offer the following psychological motivation for this update rule:

> "[...] when comparing two actions given a certain event A, the decision-maker implicitly assumes that had A not occurred, the worst possible outcome [...] would have resulted. In other words, the behavior given A [...] exhibits 'happiness' that A has occurred; the decisions are made as if we are always in 'the best of all possible worlds'."

As corresponding optimistic Bayesian update rule for conditional beliefs of CEU decision-makers obtains

$$\nu^{opt}(A \mid B) = \frac{\nu(A \cap B)}{\nu(B)}. \tag{11}$$

For the *pessimistic* (=Dempster-Shafer) update rule h is the constant act where $E = \Omega$, associating with the null-event the best consequence possible. Gilboa and Schmeidler:

"[...] we consider a 'pessimistic' decision-maker, whose choices reveal the hidden assumption that all the impossible worlds are the best conceivable ones."

The corresponding pessimistic Bayesian update rule for CEU decision-makers is

$$\nu^{pess}(A \mid B) = \frac{\nu(A \cup \neg B) - \nu(\neg B)}{1 - \nu(\neg B)}. \tag{12}$$

Observation 3: *Let* $A, B \notin \{\varnothing, \Omega\}$.

(i) *An application of the optimistic update rule (11) to a prior belief (1) results in the conditional belief*

$$\nu^{opt}(A \mid B) = \delta_B^{opt} + \left(1 - \delta_B^{opt}\right) \cdot \mu(A \mid B)$$

with

$$\delta_B^{opt} = \frac{\delta \cdot \lambda}{\delta \cdot \lambda + (1 - \delta) \cdot \mu(B)}.$$

(ii) *An application of the pessimistic update rule (12) to a prior belief (1) results in the conditional belief*

$$\nu^{pess}(A \mid B) = (1 - \delta_B^{pess}) \cdot \mu(A \mid B)$$

with

$$\delta_B^{pess} = \frac{\delta \cdot (1 - \lambda)}{\delta \cdot (1 - \lambda) + (1 - \delta) \cdot \mu(B)}.$$

Proof: Relegated to the appendix.

Let $A, B \notin \{\varnothing, \Omega\}$ and observe that

$$\nu^{pess}(A \mid B) < \nu^{opt}(A \mid B), \tag{13}$$

if $\delta > 0$, and

$$\nu^{pess}(A \mid B) < \mu(A \mid B) < \nu^{opt}(A \mid B)$$

if $\delta > 0$, $\lambda \in (0, 1)$, and $\mu(A \mid B) \in (0, 1)$. For the ex post evaluation of any Savage act f we therefore have

$$E[u(f), \nu^{pess}(A \mid B)] \leq E[u(f), \mu(A \mid B)] \leq E\left[u(f), \nu^{opt}(A \mid B)\right],$$

whereby these inequalities are strict in non-trivial cases.

Remark: Observe that the optimistic and the pessimistic update rules are extreme benchmark cases (e.g., compared against the so-called full Bayesian update rule) whose application has a strong impact on the agents' ambiguity attitudes. Following Schmeidler (1989), ambiguity aversion (resp. proneness) of a CEU decision-maker (at least in the Anscombe-Aumann framework) is usually associated with *convex* (resp. *concave*) capacities, i.e., $\nu(A \cup B) + \nu(A \cap B) \geq$ (resp. $\leq$) $\nu(A) + \nu(B)$ for all events A, B. While neo-additive capacities are neither convex nor concave for $\delta, \lambda \in (0,1)$, the resulting posteriors are convex (resp. concave) after an application of the pessimistic (resp. optimistic) update rule. To see this, notice that $\nu^{pess}(\cdot \mid B)$ is a convex combination of the convex capacities $\mu(\cdot \mid B)$ and (2) whereas $\nu^{opt}(\cdot \mid B)$ is a convex combination of the concave capacities $\mu(\cdot \mid B)$ and (3).

4 A first result: Identical information partitions

Throughout this paper we adopt Aumann's (1976) original epistemic framework. We consider two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of a non-empty state-space Ω which are interpreted as the information space of agent 1, respectively 2. Denote by $P_i(\omega)$, with $i \in \{1,2\}$, the member of $\mathcal{P}_i$ that contains $\omega \in \Omega$. We say that *i knows event $A \in \mathcal{F}$ in state ω* iff $P_i(\omega) \subseteq A$. Moreover, let $\mathcal{P}_1 \wedge \mathcal{P}_2$ denote the finest partition of $\mathcal{F}$ that is coarser than $\mathcal{P}_1$ and $\mathcal{P}_2$ (i.e., the *meet* of $\mathcal{P}_1$ and $\mathcal{P}_2$). Following Aumann's definition, we say that *event $A \in \mathcal{F}$ is common knowledge between agent 1 and 2 in state ω* iff $P(\omega) \subseteq A$ whereby $P(\omega)$ is the member of $\mathcal{P}_1 \wedge \mathcal{P}_2$ containing $\omega \in \Omega$.

Our first *agreeing to disagree* result considers the situation in which agents have identical information partitions but apply different update rules.

Proposition 1: *Consider the following assumptions:*

(A1) *The agents have identical neo-additive priors, i.e., $\nu_1 = \nu_2 \equiv \nu$, such that $\delta > 0$.*

(A2) *The agents have identical information partitions $\mathcal{P}_1 = \mathcal{P}_2 \neq \{\Omega\}$.*

(A3) *The agents' posteriors are common knowledge in some state of the world $\omega^* \in \Omega$ in the sense that it is common knowledge in ω^* that the agents have neo-additive priors with parameter-values λ, δ, μ and agent1 applies the optimistic whereas agent 2 applies the pessimistic update rule.*

Then the agents' posterior beliefs about any event $A \notin \{\varnothing, \Omega\}$ are different.

Proof: Suppose that the posteriors are common-knowledge in $\omega^* \in \Omega$. By assumption, agent 1 is optimistically and agent 2 is pessimistically biased, implying

$$\begin{aligned}\nu_1 (A \mid P_1 (\omega^*)) &= \nu^{opt} (A \mid P_1 (\omega^*)) \\ \nu_2 (A \mid P_2 (\omega^*)) &= \nu^{pess} (A \mid P_2 (\omega^*)) .\end{aligned}$$

Moreover, $\mathcal{P}_1 = \mathcal{P}_2$ implies $\mathcal{P}_1 = \mathcal{P}_2 = \mathcal{P}_1 \wedge \mathcal{P}_2$ so that $P(\omega^*) = P_1(\omega^*) = P_2(\omega^*)$. By inequality (13), the agents' posteriors $\nu_1 (A \mid P(\omega^*))$ and $\nu_2 (A \mid P(\omega^*))$ are therefore different for every event $A \notin \{\varnothing, \Omega\}$.$\square$

Proposition 1 shows that, except for degenerate cases, optimistically and pessimistically biased agents have in the ex-post situation always strict incentives to bet with each other. While the formal proof of proposition 1 is simple, it reveals a fundamental difference between Aumann's concept of information and our approach. According to Aumann, any differences in the beliefs of different agents are caused by different information received by the agents. According to our approach, differences in the beliefs of agents may also result because of different psychological attitudes with respect to the *interpretation* of new information. That is, while one agent might have a "half-full" attitude, another agent may have a "half-empty" attitude when interpreting the same fact.

5 An illustrative example: Asset-trading

We illustrate proposition 1 by a simple example in which ex-post asset-trading happens in every state of the world due to different ex-post evaluations of the asset. Moreover, these different ex-post evaluations are common knowledge to the agents despite the fact that their ex ante evaluations and their information partitions are identical.

Assume that agent 2 owns in period 1 a financial asset which gives vNM utility of 1 in case an investment project is successful and an utility of 0 in case it is not. Before it will be revealed in period 3 whether the project is successful or not, there will be news about the project's progress, either *good* or *bad*, in period 2. Let the relevant state space be given as

$$\Omega = \{SG, SB, FG, FB\}$$

whereby the event $G = \{SG, FG\}$ stands for *good* and the event $B = \{SB, FB\}$ stands for *bad* news in period 2. Accordingly, $S = \{SG, SB\}$ is the event of *success* and $F =$

$\{FG, FB\}$ is the event of *failure*. The information partitions $\mathcal{P}_1(t), \mathcal{P}_2(t)$, $t \in \{1, 2\}$, in period $t = 1$ are

$$\mathcal{P}_1(1) = \mathcal{P}_2(1) = \{\Omega\}.$$

Under the assumption of identical neo-additive priors $\nu_1 = \nu_2 = \nu$, both agents therefore (ex-ante) evaluate the Savage-act f of *holding the asset* by the same CEU (6), namely,

$$E_1[u(f), \nu] = E_2[u(f), \nu] = \delta \cdot \lambda + (1 - \delta) \cdot \mu(S).$$

As a consequence, there is no strict incentive for the agents to trade the asset in the ex-ante situation.

Consider now the following information partitions at period 2

$$\mathcal{P}_1(2) = \mathcal{P}_2(2) = \{\{SG, FG\}, \{SB, FB\}\}$$

and assume that agent 1 applies optimistically and agent 2 applies pessimistically biased Bayesian learning upon learning the news $x \in \{G, B\}$. Agent 1, resp. 2, then evaluates *holding the asset* in the ex-post situation as

$$E_1\left[u(f), \nu^{opt}(\cdot \mid x)\right] = \nu^{opt}(S \mid x),$$

resp.

$$E_2\left[u(f), \nu^{pess}(\cdot \mid x)\right] = \nu^{pess}(S \mid x).$$

By (13), we have for $\delta, \lambda \in (0, 1)$

$$E_2\left[u(f), \nu^{pess}(\cdot \mid x)\right] < \mu(S \mid x) < E_1\left[u(f), \nu^{opt}(\cdot \mid x)\right].$$

Thus, regardless of whether the news turn out *good* or *bad* agent 1 ex-post evaluates the asset strictly higher than agent 2. As a consequence there will be ex-post trade in the asset in every state of the world. For example, at price $\mu(S \mid x)$ agent 2 would strictly prefer to sell the asset while agent 1 would strictly prefer to buy it.

Remark. Observe that if both agents were EU decision-makers, i.e., $\delta = 0$, there would be no strict incentive for ex-post trading if there is no strict incentive for ex-ante trading. As this example shows, this is not necessarily true for CEU decision-makers because of the possibility of dynamically inconsistent CEU preferences.[8] According to our concept of psychologically biased Bayesian learning, the incentive for ex-post trading results in the example from the agents' different psychological attitudes with respect to the interpretation of new information.

[8] Ludwig and Zimper (2006b) demonstrate that *sophisticated* (Strotz 1956, Pollak 1968) CEU decision makers may have even stronger incentives for *intrapersonal commitment* than *sophisticated hyperbolic discounting* decision makers in the sense of Laibson (1997) and Frederick, Loewenstein, and O'Donoghue (2002).

6 Psychologically biased Bayesian learning

According to Bayesian learning models with additive probability measures subjective beliefs converge in the learning process towards objective probabilities if the agents observe a stable stochastic process. The question thus arises whether the agents of proposition 1 can agree to disagree in the long-run if they receive the same sample information drawn from an i.i.d. process. Or put differently: Is the difference in posterior beliefs according to proposition 1 only relevant in the short-run but vanishes in the long-run? In order to address this question about the relevance of proposition 1 we apply in this section our decision-theoretic framework to a formal model of Bayesian learning with neo-additive capacities. Our main result will demonstrate that CEU decision-makers may agree to disagree in the long-run even if they forever receive identical information. Moreover, this finding casts doubts on the standard argument that people with identical information should have common priors (e.g., Aumann 1987, 1998, Gul 1998).

6.1 The benchmark case: Rational Bayesian learning

Consider the situation of an agent who is uncertain about the probability of an outcome, H, but can observe a statistical experiment with n independent trials where H is a possible outcome in each trial that occurs in every trail with identical probability. Let

$$S = \times_{i=0}^{\infty} \{H, T\}$$

denote the experiment's sample space. We consider the state space $\Omega = [0,1] \times S$ and the event space $\mathcal{F} = \mathcal{B} \times \mathcal{S}$ whereby $\mathcal{B}$ denotes the Borel σ-algebra of the unit-interval $[0,1]$ and $\mathcal{S}$ is the power-set of S. Let $\boldsymbol{\pi}$ denote the event in $\mathcal{B} \times \mathcal{S}$ that $\pi \in [0,1]$ is the probability of outcome H, i.e.,

$$\boldsymbol{\pi} = \{\omega \in \Omega \mid \omega = (\pi, ...)\}.$$

Similarly, let $\mathbf{I}_n^k$ denote the event in $\mathcal{B} \times \mathcal{S}$ that outcome H has occurred k-times in the first n trials. In our framework π is the unknown parameter of a Binomial-distribution that stands for the objective probability by which outcome H occurs. Let us suppose that the agent has a subjective additive probability measure, μ, defined on $\mathcal{B} \times \mathcal{S}$ whereby we interpret $\mu\left([\underline{\boldsymbol{\pi}}, \overline{\boldsymbol{\pi}}], \mathbf{I}_n^k\right)$ as the agent's subjective (joint) probability that the "true" probability of outcome H lies in the interval $[\underline{\boldsymbol{\pi}}, \overline{\boldsymbol{\pi}}]$ and that she observes information $\mathbf{I}_n^k$. We further assume that the agent's prior (marginal) distribution $\mu(\cdot)$ over the parameter π is a Beta distribution so that her estimator for the "true" value of π is the expected value of this Beta-distribution, i.e., $E[\pi, \mu(\cdot)] = \frac{\alpha}{\alpha+\beta}$ for given distribution parameters

$\alpha, \beta > 0$. That is, the prior distribution over π is characterized by the probability density[9]

$$\mu(\boldsymbol{\pi}) = \begin{cases} K_{\alpha,\beta}\pi^{\alpha-1}(1-\pi)^{\beta-1} & \text{for } 0 \leq \pi \leq 1 \\ 0 & \text{else} \end{cases} \tag{14}$$

where $K_{\alpha,\beta}$ is a normalizing constant.[10] Since the probability of receiving information $\mathbf{I}_n^k$ for a given π (=likelihood function) is, by the Binomial-assumption,

$$\mu\left(\mathbf{I}_n^k \mid \boldsymbol{\pi}\right) = \binom{n}{k}\pi^k(1-\pi)^{n-k},$$

we obtain by Bayes' rule

$$\begin{aligned} \mu\left(\boldsymbol{\pi} \mid \mathbf{I}_n^k\right) &= \frac{\mu\left(\boldsymbol{\pi}, \mathbf{I}_n^k\right)}{\mu\left(\mathbf{I}_n^k\right)} \\ &= \frac{\mu\left(\mathbf{I}_n^k \mid \boldsymbol{\pi}\right)\mu(\boldsymbol{\pi})}{\mu\left(\mathbf{I}_n^k\right)} \\ &= K_{\alpha+k,\beta+n-k}\pi^{\alpha+k-1}(1-\pi)^{\beta+n-k-1} \end{aligned}$$

whenever $\mu\left(\mathbf{I}_n^k\right) = \int_{[0,1]} \mu\left(\mathbf{I}_n^k \mid \boldsymbol{\pi}\right)\mu(\boldsymbol{\pi})\,d\pi > 0$.

Observe that the agent's subjective posterior distribution over π is a Beta-distribution with parameters $\alpha + k, \beta + n - k$. Accordingly, the agent's posterior belief is given by the expected value of the posterior distribution, $E\left[\pi, \mu\left(\cdot \mid \mathbf{I}_n^k\right)\right] = \frac{\alpha+k}{\alpha+\beta+n}$, which, using that the prior belief is $E[\pi, \mu(\cdot)] = \frac{\alpha}{\alpha+\beta}$, we can rewrite as

$$E\left[\pi, \mu\left(\cdot \mid \mathbf{I}_n^k\right)\right] = \left(\frac{\alpha+\beta}{\alpha+\beta+n}\right)E[\pi, \mu(\cdot)] + \left(\frac{n}{\alpha+\beta+n}\right)\frac{k}{n} \tag{15}$$

where $\frac{k}{n}$ is the observed sample mean according to information $\mathbf{I}_n^k$. That is, the agent's posterior estimator about the probability of H is a weighted average of her prior estimator and the sample mean whereby the weight attached to the sample mean increases in the number of trials.[11] Let π^* denote the "true" probability of outcome H. Since, for every $c > 0$, $\lim_{n\to\infty} prob\left(\left|\frac{k}{n} - \pi^*\right| \leq c\right) = 1$ we obtain the following result for this standard model of rational Bayesian learning.

---

[9] For the ease of exposition we somewhat abuse notation in that we write μ interchangeably for an additive probability measure, which is a set function, and for a density function, which is defined on the real line.

[10] In particular, $K_{\alpha,\beta} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ where $\Gamma(y) = \int_0^\infty x^{y-1}e^{-x}dx$ for $y > 0$.

[11] Tonks (1983) introduces a similar model of rational Bayesian learning in which the agent has a normally distributed prior over the mean of some normal distribution and receives normally distributed information.

Observation 4: *Under the assumption of* $\mu\left(\mathbf{I}_n^k\right) > 0$ *for all* n *the posterior estimator* $E\left[\pi, \mu\left(\cdot \mid \mathbf{I}_n^k\right)\right]$ *converges in probability to the true probability* π^* *of outcome* H *if the number of trials,* n*, approaches infinity.*

6.2 Learning with ambiguous beliefs

In this section we formally link the updating of ambiguous beliefs to Bayesian learning behavior. As a generalization of the Bayesian learning model discussed above, we consider now a neo-additive prior about the unknown parameter π

$$\nu(\boldsymbol{\pi}) = \begin{cases} \delta\lambda + (1-\delta) \cdot K_{\alpha,\beta}\pi^{\alpha-1}(1-\pi)^{\beta-1} & \text{for } 0 \leq \pi \leq 1 \\ 0 & \text{else} \end{cases} \tag{16}$$

whereby the additive part is given by (14). Accordingly, the agent's prior estimator for the true value of π is now given as the Choquet expected value (6)

$$\begin{aligned} E[\pi, \nu(\cdot)] &= \delta(\lambda \max \pi + (1-\lambda) \min \pi) + (1-\delta) E[\pi, \mu(\cdot)] \\ &= \delta\lambda + (1-\delta) E[\pi, \mu(\cdot)]. \end{aligned}$$

The following lemma uses our general results on Bayesian updating of neo-additive capacities (observation 3) in order to derive conditional neo-additive capacities for the special case (16) and to characterize the corresponding conditional Choquet expected values which stand for the agent's posterior beliefs about the probability of outcome H.

Lemma. *Suppose the agent receives information* $\mathbf{I}_n^k \in \mathcal{B} \times \mathcal{S}$*. Contingent on the applied update rule we obtain the following conditional neo-additive beliefs and posterior estimators about parameter* π*.*

(i) *Optimistic Bayesian updating.*

$$\nu^{opt}\left(\boldsymbol{\pi} \mid \mathbf{I}_n^k\right) = \delta_{\mathbf{I}_n^k}^{opt} + \left(1 - \delta_{\mathbf{I}_n^k}^{opt}\right) \cdot K_{\alpha+k,\beta+n-k}\pi^{\alpha+k-1}(1-\pi)^{\beta+n-k-1}$$

with

$$\delta_{\mathbf{I}_n^k}^{opt} = \frac{\delta \cdot \lambda}{\delta \cdot \lambda + (1-\delta) \cdot \mu\left(\mathbf{I}_n^k\right)}$$

so that

$$E\left[\pi, \nu^{opt}\left(\cdot \mid \mathbf{I}_n^k\right)\right] = \delta_{\mathbf{I}_n^k}^{opt} + \left(1 - \delta_{\mathbf{I}_n^k}^{opt}\right) \cdot E\left[\pi, \mu\left(\cdot \mid \mathbf{I}_n^k\right)\right]$$

(ii) *Pessimistic Bayesian updating.*

$$\nu^{pess}\left(\boldsymbol{\pi} \mid \mathbf{I}_n^k\right) = \left(1 - \delta_{\mathbf{I}_n^k}^{pess}\right) \cdot K_{\alpha+k,\beta+n-k}\pi^{\alpha+k-1}(1-\pi)^{\beta+n-k-1}$$

with

$$\delta^{pess}_{\mathbf{I}^k_n} = \frac{\delta \cdot (1-\lambda)}{\delta \cdot (1-\lambda) + (1-\delta) \cdot \mu\left(\mathbf{I}^k_n\right)}$$

so that

$$E\left[\pi, \nu^{pess}\left(\cdot \mid \mathbf{I}^k_n\right)\right] = \left(1 - \delta^{pess}_{\mathbf{I}^k_n}\right) \cdot E\left[\pi, \mu\left(\cdot \mid \mathbf{I}^k_n\right)\right]$$

If the number of trials approaches infinity, i.e., $n \rightarrow \infty$, the sample information $\mathbf{I}^k_n$ converges in probability to the information $\mathbf{I}^*$ according to which outcome H has occurred with frequency π^*. In the limit of a Bayesian learning process the agent's posterior estimators about π will therefore converge to the belief $E\left[\pi, \nu\left(\cdot \mid \mathbf{I}^*\right)\right]$ which depends on the applied Bayesian update rule. The following observation, which combines the above lemma with observation 4, characterizes these limit beliefs.

Observation 5. *Let $n \rightarrow \infty$. Contingent on the applied update rule the agents limit beliefs converge in probability to the following posteriors.*

(i) *Optimistic Bayesian learning.*

$$E\left[\pi, \nu^{opt}\left(\cdot \mid \mathbf{I}^*\right)\right] = \delta^{opt}_{\mathbf{I}^*} + \left(1 - \delta^{opt}_{\mathbf{I}^*}\right) \cdot \pi^*$$

such that

$$\delta^{opt}_{\mathbf{I}^*} = \frac{\delta \cdot \lambda}{\delta \cdot \lambda + (1-\delta) \cdot \mu\left(\mathbf{I}^*\right)}$$

(ii) *Pessimistic Bayesian learning.*

$$E\left[\pi, \nu^{pess}\left(\cdot \mid \mathbf{I}^*\right)\right] = \left(1 - \delta^{pess}_{\mathbf{I}^*}\right) \cdot \pi^*$$

such that

$$\delta^{pess}_{\mathbf{I}^*} = \frac{\delta \cdot (1-\lambda)}{\delta \cdot (1-\lambda) + (1-\delta) \cdot \mu\left(\mathbf{I}^*\right)}$$

Observation 5 demonstrates that psychologically biased Bayesian learning in our sense violates the two standard paradigms of rational Bayesian learning. Firstly, the posterior "subjective" beliefs do not converge to the "objective" probabilities in an infinite learning process. Secondly, since

$$E\left[\pi, \nu^{pess}\left(\cdot \mid \mathbf{I}^*\right)\right] < E\left[\pi, \nu^{opt}\left(\cdot \mid \mathbf{I}^*\right)\right]$$

for $\delta \neq 0$, the posterior estimators of two different agents do not converge to the same limit belief if these agents receive the same information but interpret it differently. Since these posterior estimators of our learning model, can be interpreted - in the appropriate event-space - as the agent's neo-additive belief that some event H occurs, the *agreeing to disagree* result of proposition 1 is not limited to the short-run but also holds in the long-run for CEU decision-makers.

7 A second result: Identical learning rules

Our second *agreeing to disagree* result applies to people who use the same psychologically biased learning rule but have different information partitions.

Proposition 2: *Consider the following assumptions:*

(A1') *The agents have identical neo-additive priors, i.e., $\nu_1 = \nu_2 \equiv \nu$, such that $\delta > 1$.*

(A2') *The agents' posteriors are common knowledge in some state of the world $\omega^* \in \Omega$ in the sense that it is common knowledge in ω^* that the agents have neo-additive priors with parameter-values λ, δ, μ and both agents apply the optimistic (resp. pessimistic) update rule.*

(A3') *The agents' priors satisfy $\nu(P_1(\omega^*)) \neq \nu(P_2(\omega^*))$ whereby $P_1(\omega^*), P_2(\omega^*) \neq \Omega$.*

Then the agents' posterior beliefs about any event $A \notin \{\varnothing, \Omega\}$ are different.

Observe that assumption (A3'), i.e., $\nu(P_1(\omega^*)) \neq \nu(P_2(\omega^*))$, cannot hold if the agents have identical priors **and** identical information partitions. That is, the result of proposition 2 only applies in situations of asymmetric information, i.e., $\mathcal{P}_1 \neq \mathcal{P}_2$, such that the two events $P_1(\omega^*)$ and $P_2(\omega^*)$ are not equally likely according to the agents' common prior.

Before we turn to the proof of proposition 2 consider the following example which illustrates the intuition behind our formal proof.

Example. Consider the following information structure

$$\begin{aligned}\mathcal{P}_1 &= \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, ...\} = \{P_1^1, P_1^2, ...\}, \\ \mathcal{P}_2 &= \{\{\omega_1, \omega_2, \omega_3, \omega_4\}, ...\} = \{P_2^1, ...\}\end{aligned}$$

so that

$$\mathcal{P}_1 \wedge \mathcal{P}_2 = \{\{\omega_1, \omega_2, \omega_3, \omega_4\}, ...\}.$$

Suppose agent 1 and 2 have a common neo-additive prior ν with $\delta \in (0,1)$ such that

$$\mu(\{\omega_1\}) = ... = \mu(\{\omega_4\}) > 0.$$

Further suppose that both agents are optimistically biased. Let

$$A = \{\omega_2, \omega_3\}$$

and observe that

$$\nu_1\left(A \mid P_1^1\right) = \nu_1\left(A \mid P_1^2\right) = \delta_1^{opt} + \left(1 - \delta_1^{opt}\right) \cdot \frac{1}{2} \tag{17}$$

with

$$\delta_1^{opt} = \frac{\delta \cdot \lambda}{\delta \cdot \lambda + (1-\delta) \cdot \mu(\{\omega_1, \omega_2\})}$$

and

$$\nu_2\left(A \mid P_2^1\right) = \delta_2^{opt} + \left(1 - \delta_2^{opt}\right) \cdot \frac{1}{2} \tag{18}$$

with

$$\delta_2^{opt} = \frac{\delta \cdot \lambda}{\delta \cdot \lambda + (1-\delta) \cdot \mu(\{\omega_1, \omega_2, \omega_3, \omega_4\})}.$$

Observe that the posterior of each agent is the same in every state belonging to $P(\omega^*) \in \mathcal{P}_1 \wedge \mathcal{P}_2$ with $\omega^* \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$ so that we can stipulate, in accordance with assumption (A2'), that the agents' posteriors as well as their parameter values λ, δ, μ are common knowledge in every state $\omega^* \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Since $\delta > 0$, the posteriors (17) and (18) coincide if and only if

$$\begin{aligned} \delta_1^{opt} &= \delta_2^{opt} \Leftrightarrow \\ \mu(\{\omega_1, \omega_2\}) &= \mu(\{\omega_1, \omega_2, \omega_3, \omega_4\}) \Leftrightarrow \\ \nu(P_1(\omega^*)) &= \nu(P_2(\omega^*)), \end{aligned}$$

which is not the case in this example. Thus, despite identical priors and identical Bayesian learning rules, both agents have different posterior beliefs which are common knowledge.

Proof of proposition 2. Our proof builds on Aumann's (1976) original proof for the case of an additive probability measure, i.e., $\delta = 0$.

Step 1. Aumann (1976): For an additive common prior μ the agents' posteriors must be identical when they are common knowledge at some state of the world.

Suppose on the contrary that there is some $\omega^* \in \Omega$ in which it is common knowledge that

$$\mu_1 (A \mid P_1 (\omega^*)) = q_1 \text{ and } \mu_2 (A \mid P_2 (\omega^*)) = q_2$$

such that $q_1 \neq q_2$ for some event $A \in \mathcal{F}$. Then

$$\mu_1 \left(A \mid P_1^j\right) = q_1, \tag{19}$$

for all $P_1^j \subseteq P(\omega^*)$ whereby $P(\omega^*)$ is the member of $\mathcal{P}_1 \wedge \mathcal{P}_2$ containing ω^*. Denote by $P_1^1, ..., P_1^n$ the members of $\mathcal{P}_1$ such that

$$P_1^1 \cup ... \cup P_1^n = P(\omega^*) \ .$$

By additivity,

$$\mu\left(P_1^1\right) + ... + \mu\left(P_1^n\right) = \mu\left(P(\omega^*)\right) \tag{20}$$

since $P_1^1, ..., P_1^n$ is a partition of $P(\omega^*)$. Also by additivity,

$$\mu_1\left(A \mid P_1^j\right) = \frac{\mu\left(A \cap P_1^j\right)}{\mu\left(P_1^j\right)},\ j = 1, .., n$$

so that, by (19),

$$\mu\left(P_1^1\right) + ... + \mu\left(P_1^n\right) = \frac{\mu\left(A \cap P_1^1\right)}{q_1} + ... + \frac{\mu\left(A \cap P_1^n\right)}{q_1}.$$

Since, by additivity,

$$\mu\left(A \cap P_1^1\right) + ... + \mu\left(A \cap P_1^n\right) = \mu\left(A \cap P(\omega^*)\right)$$

we have

$$\mu\left(P_1^1\right) + ...\mu\left(P_1^n\right) = \frac{\mu\left(A \cap P(\omega^*)\right)}{q_1}.$$

Thus, by (20),

$$\frac{\mu\left(A \cap P(\omega^*)\right)}{q_1} = \mu\left(P(\omega^*)\right).$$

An analogous argument for agent 2 results in

$$\frac{\mu\left(A \cap P(\omega^*)\right)}{q_2} = \mu\left(P(\omega^*)\right)$$

implying the desired result $q_1 = q_2$.□

Step 2. Consider now the case of identical non-additive priors (1), i.e., $\delta > 0$. Let $A \notin \{\varnothing, \Omega\}$ and suppose both agents are optimistically biased; (there is an analogous argument for pessimistically biased agents). Then, for $\omega \in \Omega$,

$$\nu_1^{opt}\left(A \mid P_1(\omega)\right) = \delta_{1,B}^{opt} + \left(1 - \delta_{1,B}^{opt}\right) \cdot \mu\left(A \mid P_1(\omega)\right) \tag{21}$$

with

$$\delta_{1,B}^{opt} = \frac{\delta \cdot \lambda}{\delta \cdot \lambda + (1-\delta) \cdot \mu \left(P_1 \left(\omega\right)\right)}$$

and

$$\nu_2^{opt} \left(A \mid P_2 \left(\omega\right)\right) = \delta_{2,B}^{opt} + \left(1 - \delta_{2,B}^{opt}\right) \cdot \mu \left(A \mid P_2 \left(\omega\right)\right) \tag{22}$$

with

$$\delta_{2,B}^{opt} = \frac{\delta \cdot \lambda}{\delta \cdot \lambda + (1-\delta) \cdot \mu \left(P_2 \left(\omega\right)\right)}.$$

Assume now that the posteriors (21) and (22) as well as their parameter values λ, δ, μ are common knowledge in some state $\omega^* \in \Omega$ in accordance with assumption (A2'). By the argument of step 1, the posteriors must coincide for the special case of $\delta = 0$. We therefore have for the additive part of the posteriors

$$\mu \left(A \mid P_1 \left(\omega^*\right)\right) = \mu \left(A \mid P_2 \left(\omega^*\right)\right),$$

so that the agents' posteriors (21) and (22) are different if and only if

$$\begin{aligned} \delta_{1,B}^{opt} &\neq \delta_{2,B}^{opt} \Leftrightarrow \\ \mu \left(P_1 \left(\omega^*\right)\right) &\neq \mu \left(P_2 \left(\omega^*\right)\right) \Leftrightarrow \\ \nu \left(P_1 \left(\omega^*\right)\right) &\neq \nu \left(P_2 \left(\omega^*\right)\right). \end{aligned}$$

This proves the proposition. □□

8 Concluding remarks and outlook

We consider CEU decision-makers who may interpret new information in an optimistic or pessimistic way. As our first main contribution we apply these two benchmark cases of psychologically biased Bayesian updating to the epistemic situation studied in Aumann (1976). Two formal main results emerge:

1. Even if people receive the same information, they may agree to disagree if their psychologically attitudes about the interpretation of new information are different.

2. Even if people have the same psychological attitudes, they may agree to disagree if they receive different information.

Both results are in contrast to Aumann's famous conclusion that agents cannot agree to disagree regardless of whether they receive the same information or not. Our concept of psychologically biased Bayesian learning can therefore offer a possible explanation for the existence of ex-post trade in financial assets.

As our second main contribution, we address the question in how far the first agreeing-to-disagree result may be relevant in the long-run. To this end we introduce a formal model of psychologically biased Bayesian learning based on optimistic respectively pessimistic interpretation of new information which encompasses the standard model of rational Bayesian learning with additive beliefs as a special case. According to this standard model people's posteriors coincide in the limit if they observe the same sample information drawn from an i.i.d. process so that they cannot agree to disagree forever. We demonstrate that the case is different for CEU decision-makers with non-additive beliefs since their posteriors will not converge to the same limit beliefs if they apply different update rules. While our learning model thus offers a long-run justification for our first agreeing-to-disagree result, it also shows that identical information does not necessarily induce identical priors.

In order to provide straightforward psychological interpretations of our decision-theoretic framework, we have restricted attention to the special case of CEU preferences with neo-additive capacities. Especially with respect to updating and learning under ambiguity, the assumption of neo-additive capacities turned out as formally and interpretational extremely convenient. For the sake of mathematical generality, however, it would be interesting to obtain similar results for more general classes of capacities. This holds especially true since Aumann's argument has been so far - with the notable exception of Dow, Madrigal and Werlang (1990) - foremostly investigated within the multiple-priors but not within the CEU framework. Another avenue for future research would be to look into alternative updating rules for capacities and their empirical relevance. The optimistic and the pessimistic update rule are, admittedly, very extreme update rules whose main virtue is to capture the difference between "half full" and "half empty" attitudes in a clear-cut way. Less extreme update rules, such as the full-Bayesian update rule, may be more appropriate for the description of real-life decision-makers' updating and learning behavior.

Appendix

Proof of observation 1: By an argument in Schmeidler (1986), it suffices to restrict attention to a non-negative valued function f so that

$$E[f,\nu] = \int_0^{+\infty} \nu\{\omega \in \Omega \mid f(\omega) \geq z\}\, dz, \tag{23}$$

which is equivalent to

$$E[f,\nu] = \int_{\min f}^{\max f} \nu\{\omega \in \Omega \mid f(\omega) \geq z\}\, dz$$

since the range of f is closed and bounded. We consider a partition P_n, $n = 1, 2, ...,$ of Ω with members

$$A_n^k = \{\omega \in \Omega \mid a_{k,n} < X(\omega) \leq b_{k,n}\} \text{ for } k = 1, ..., 2^n$$

such that

$$\begin{aligned} a_{k,n} &= [\max f - \min f] \cdot \frac{(k-1)}{2^n} + \min f \\ b_{k,n} &= [\max f - \min f] \cdot \frac{k}{2^n} + \min f. \end{aligned}$$

Define the step functions $a_n : \Omega \to \mathbb{R}$ and $b_n : \Omega \to \mathbb{R}$ such that, for $\omega \in A_n^k$, $k = 1, ..., 2^n$,

$$\begin{aligned} a_n(\omega) &= a_{k,n} \\ b_n(\omega) &= b_{k,n}. \end{aligned}$$

Obviously,

$$E[a_n,\nu] \leq E[f,\nu] \leq E[b_n,\nu]$$

for all n and

$$\lim_{n\to\infty} E[b_n,\nu] - E[a_n,\nu] = 0.$$

That is, $E[a_n,\nu]$ and $E[b_n,\nu]$ converge to $E[f,\nu]$ for $n \to \infty$. Furthermore, observe that

$$\begin{aligned} \min a_n &= \min f \text{ for all } n, \text{ and} \\ \max b_n &= \max f \text{ for all } n. \end{aligned}$$

Since $\lim_{n\to\infty} \min b_n = \lim_{n\to\infty} \min a_n$ and $E[b_n,\mu]$ is continuous in n, we have

$$\begin{aligned} \lim_{n\to\infty} E[b_n,\nu] &= \delta\left(\lambda \lim_{n\to\infty} \max b_n + (1-\lambda) \lim_{n\to\infty} \min b_n\right) + (1-\delta) \lim_{n\to\infty} E[b_n,\mu] \\ &= \delta(\lambda \max f + (1-\lambda)\min f) + (1-\delta) E[f,\mu]. \end{aligned}$$

In order to prove proposition 3, it therefore remains to be shown that, for all n,

$$E\left[b_n,\nu\right]=\delta\left(\lambda\max b_n+(1-\lambda)\min b_n\right)+(1-\delta)\,E\left[b_n,\mu\right].$$

Since b_n is a step function, (23) becomes

$$\begin{aligned}E\left[b_n,\nu\right] &= \sum_{A_n^k\in P_n}\nu\left(A_n^{2^n}\cup...\cup A_n^k\right)\cdot\left(b_{k,n}-b_{k-1,n}\right)\\ &= \sum_{A_n^k\in P_n}b_{k,n}\cdot\left[\nu\left(A_n^{2^n}\cup...\cup A_n^k\right)-\nu\left(A_n^{2^n}\cup...\cup A_n^{k-1}\right)\right],\end{aligned}$$

implying for a neo-additive capacity

$$\begin{aligned}E\left[b_n,\nu\right] &= \max b_n\left[\delta\lambda+(1-\delta)\,\mu\left(A_n^{2^n}\right)\right]+\sum_{k=2}^{2^n-1}b_{k,n}\,(1-\delta)\,\mu\left(A_n^k\right)\\ &\quad+\min b_n\left[1-\delta\lambda-(1-\delta)\sum_{k=2}^{2^n}\mu\left(A_n^k\right)\right]\\ &= \delta\lambda\max b_n+(1-\delta)\sum_{k=1}^{2^n}b_{k,n}\mu\left(A_n^k\right)+\min b_n\left[\delta-\delta\lambda\right]\\ &= \delta\left(\lambda\max b_n+(1-\lambda)\min b_n\right)+(1-\delta)\,E\left[b_n,\mu\right].\end{aligned}$$

□

Proof of observation 3:

Applying the optimistic Bayesian update rule to a neo-additive capacity gives, for $A\notin\{\varnothing,\Omega\}$,

$$\begin{aligned}\nu\left(A\mid B\right) &= \frac{\delta\cdot\lambda+(1-\delta)\cdot\mu\left(A\cap B\right)}{\delta\cdot\lambda+(1-\delta)\cdot\mu\left(B\right)}\\ &= \frac{\delta\cdot\lambda}{\delta\cdot\lambda+(1-\delta)\cdot\mu\left(B\right)}+\frac{(1-\delta)\cdot\mu\left(B\right)}{\delta\cdot\lambda+(1-\delta)\cdot\mu\left(B\right)}\cdot\mu\left(A\mid B\right)\\ &= \delta_B^{opt}+\left(1-\delta_B^{opt}\right)\cdot\mu\left(A\mid B\right)\end{aligned}$$

such that

$$\delta_B^{opt}=\frac{\delta\cdot\lambda}{\delta\cdot\lambda+(1-\delta)\cdot\mu\left(B\right)}.$$

Applying the pessimistic Bayesian update rule to a neo-additive capacity gives, for $A \notin \{\varnothing, \Omega\}$,

$$\begin{aligned}
\nu^{pess}(A \mid B) &= \frac{\nu(A \cup \neg B) - \nu(\neg B)}{1 - \nu(\neg B)} \\
&= \frac{\delta \cdot \lambda + (1-\delta) \cdot \mu(A \cup \neg B) - \delta \cdot \lambda - (1-\delta) \cdot \mu(\neg B)}{1 - \delta \cdot \lambda - (1-\delta) \cdot \mu(\neg B)} \\
&= \frac{(1-\delta) \cdot \mu(A)}{1 - \delta \cdot \lambda - (1-\delta) \cdot (\mu(\neg B))} - \frac{(1-\delta)\,\mu(A \cap \neg B)}{1 - \delta \cdot \lambda - (1-\delta) \cdot (\mu(\neg B))} \\
&= \frac{(1-\delta) \cdot \mu(A)}{1 - \delta \cdot \lambda - (1-\delta) \cdot (\mu(\neg B))} - \frac{(1-\delta)\,\mu(\neg B)}{1 - \delta \cdot \lambda - (1-\delta) \cdot (\mu(\neg B))}\mu(A \mid \neg B) \\
&= \frac{(1-\delta) \cdot \mu(A)}{1 - \delta \cdot \lambda - (1-\delta) \cdot (\mu(\neg B))} \\
&\quad - \frac{(1-\delta)\,\mu(\neg B)}{1 - \delta \cdot \lambda - (1-\delta) \cdot (\mu(\neg B))}\left[\frac{\mu(A) - \mu(A \mid B) \cdot \mu(B)}{\mu(\neg B)}\right] \\
&= \frac{(1-\delta) \cdot \mu(B)}{\delta \cdot (1-\lambda) + (1-\delta) \cdot \mu(B)} \cdot \mu(A \mid B) \\
&= (1 - \delta_B^{pess}) \cdot \mu(A \mid B)
\end{aligned}$$

such that

$$\delta_B^{pess} = \frac{\delta \cdot (1-\lambda)}{(\delta \cdot (1-\lambda) + (1-\delta) \cdot \mu(B))}.$$

□

References

Anscombe, F.J, and R.J. Aumann (1963), "A Definition of Subjective Probability", *Annals of American Statistics* **34**, 199-205.

Aumann, R. (1976), "Agreeing to disagree," *Annals of Statistics* **4**, 1236-1239.

Aumann, R. (1987), "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica* **55**, 1-18.

Aumann, R. (1998), "Common Priors: A Reply to Gul," *Econometrica* **66**, 929-938.

Bacharach, M. (1985), "Some Extension of a Claim of Aumann in an Axiomatic Model of Knowledge", *Journal of Economic Theory* **37**, 167–190.

Baron, J. (2007), *Thinking and Deciding*, Cambridge University Press: New York, Melbourne, Madrid.

Beck, A.T. (1976), *Cognitive Therapy and the Emotional Disorders*, International Universities Press: New York.

Berger, J., and L.M. Berliner (1986), "Robust Bayes and Empirical Bayes Analysis with ε-Contaminated Priors", *Annals of Statistics* **14**, 461-486.

Bonanno, G., and K. Nehring (1999), "How to Make Sense of the Common Prior Assumption under Incomplete Information", *International Journal of Game Theory* **28**, 409-434.

Chapman, C.R. (1973), "Prior Probability Bias in Information Seeking and Opinion Revision," *The American Journal of Psychology* **86**, 269-282.

Chateauneuf, A., Eichberger, J., and S. Grant (2006), "Choice under Uncertainty with the Best and Worst in Mind: Neo-additive Capacities", *Journal of Economic Theory* forthcoming.

Dow, J., Madrigal, V., and S.R. Werlang (1990), "Preferences, Common Knowledge, and Speculative Trade", mimeo

Eichberger, J., and D. Kelsey (1999), "E-Capacities and the Ellsberg Paradox", *Theory and Decision* **46**, 107-140.

Eichberger, J., Grant, S., and D. Kelsey (2006), "Updating Choquet Expected Utility Preferences", mimeo

Ellsberg, D. (1961), "Risk, Ambiguity and the Savage Axioms", *Quarterly Journal of Economics* **75**, 643-669.

Epstein, L.G. (1999), "A Definition of Uncertainty Aversion", *The Review of Economic Studies* **66**, 579-608.

Epstein, L.G. and M. Le Breton (1993), "Dynamically Consistent Beliefs Must Be Bayesian", *Journal of Economic Theory* **61**, 1-22.

Epstein, L.G. and M. Schneider (2003), "Recursive Multiple-Priors ", *Journal of Economic Theory* **113**, 1-31.

Epstein, L.G. and M. Schneider (2007), "Learning Under Ambiguity", *Review of Economic Studies* **74**, 1275-1303.

Epstein, L.G. and T. Wang (1994), "Intertemporal Asset Pricing Under Knightian Uncertainty," *Econometrica* **62**, 283-322.

Frederick, S., Loewenstein, G., and T. O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review", *Journal of Economic Literature* **Vol. XL**, 351-401.

Geanakoplos, J. (1992), "Common Knowledge", *Journal of Economic Perspectives* **6**, 53–82.

Ghirardato, P., and M. Marinacci (2002), "Ambiguity Made Precise: A Comparative Foundation", *Journal of Economic Theory* **102**, 251–289.

Ghirardato, P., Klibanoff, P., and M. Marinacci (1998), "Additivity with Multiple Priors", *Journal of Mathematical Economics* **30**, 405-420.

Ghirardato, P., Maccheroni, F., and M. Marinacci (2004), "Differentiating Ambiguity and Ambiguity Attitude", *Journal of Economic Theory* **118**, 133-173.

Gilboa, I. (1987), "Expected Utility with Purely Subjective Non-Additive Probabilities", *Journal of Mathematical Economics* **16**, 65-88.

Gilboa, I., and D. Schmeidler (1989), "Maxmin Expected Utility with Non-Unique Priors," *Journal of Mathematical Economics* **18**, 141-153.

Gilboa, I., and D. Schmeidler (1993), "Updating Ambiguous Beliefs", *Journal of Economic Theory* **59**, 33-49.

Gul, F. (1998), "A Comment on Aumann's Bayesian View," *Econometrica* **66**, 923-927.

Halevy, Y. (1998), "Trade between Rational Agents as a Result of Asymmetric Information," mimeo.

Halevy, Y. (2004), "The Possibility of Speculative Trade between Dynamically Consistent Agents," *Games and Economic Behavior* **46**, 189-198.

Harsanyi, J.C. (1967), "Games with Incomplete Information Played by 'Bayesian' Players. Part I: The Basic Model," *Management Science* **14**, 159-182.

Kahneman, D., and A. Tversky (1979), "Prospect theory: An Analysis of Decision under Risk," *Econometrica* **47**, 263-291.

Kajii, A. and T. Ui (2005), "Incomplete Information Games with Multiple Priors," *Japanese Economic Review* **56**, 332-351.

Kajii, A. and T. Ui (2007), "Interim Efficient Allocations under Uncertainty," KIER discussion paper 642.

Laibson, D. (1997), "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics* **112**, 443-477.

Lord, C.G., Ross, L., and M.R. Lepper (1979), "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," *Journal of Personality and Social Psychology*, **37**, 2098-2109.

Ludwig, A., and A. Zimper (2006a), "Rational Expectations and Ambiguity: A Comment on Abel (2002)," *Economics Bulletin*, **4**(2), 1-15.

Ludwig, A., and A. Zimper (2006b), "Investment Behavior under Ambiguity: The Case of Pessimistic Decision Makers," *Mathematical Social Sciences* **52**, 111-130.

Ma, C. (2001), "No Trade Theorem under Knightian Uncertainty with General Preferences," *Theory and Decision* **51**, 173–181.

Marinacci, M. (1999), "Limit Laws for Non-additive Probabilities and Their Frequentist Interpretation," *Journal of Economic Literature* **84**, 145–195.

Marinacci, M. (2002), "Learning from Ambiguous Urns," *Statistical Papers* **43**, 145–151.

Milgrom, P. (1981), "An Axiomatic Characterization of Common Knowledge," *Econometrica* **49**, 219-222.

Milgrom, P., and N. Stockey (1982), "Information, Trade and Common Knowledge," *Journal of Economic Theory* **26**, 17-27.

Morris, S. (1994), "Trade with Heterogeneous Prior Beliefs and Asymmetric Information," *Econometrica* **62**, 1327-1347.

Neeman, Z. (1996), "Approximating Agreeing to Disagree Results with Common p-Beliefs," *Games and Economic Behavior* **16**, 77-96.

Pitz, G.F. (1969), "An Inertia Effect (Resistance to Change) in the Revision of Opinion," *Canadian Journal of Psychology* **23**, 24-33.

Pitz, G.F., Downing, L., and H. Reinhold (1967), "Sequential Effects in the Revision of Subjective Probabilities," *Canadian Journal of Psychology* **21**, 381-393.

Pollak, R.A. (1968), "Consistent Planning," *The Review of Economic Studies* **35**, 201-208.

Quiggin, J.P. (1981), "Risk Perception and Risk Aversion among Australian Farmers," *Australian Journal of Agricultural Economics* **25**, 160-169.

Quiggin, J.P. (1982), "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization* **3**, 323-343.

Rubinstein, A., and A. Wollinski (1990), "On the Logic of "Agreeing to Disagree" Type Results," *Journal of Economic Theory* **51**, 184-193.

Samet, D. (1990), "Ignoring Ignorance and Agreeing to Disagree," *Journal of Economic Theory* **52**, 190-207.

Sarin, R., and P.P. Wakker (1992), "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica* **60**, 1255-1272.

Sarin, R., and P.P. Wakker (1998), "Revealed Likelihood and Knightian Uncertainty," *Journal of Risk and Uncertainty* **16**, 223-250.

Savage, L.J. (1954), *The Foundations of Statistics*, John Wiley and & Sons, Inc.: New York, London, Sydney.

Schmeidler, D. (1986), "Integral Representation without Additivity," *Proceedings of the American Mathematical Society* **97**, 255-261.

Schmeidler, D. (1989), "Subjective Probability and Expected Utility without Additivity," *Econometrica* **57**, 571-587.

Siniscalchi, M. (2001), "Bayesian Updating for General Maxmin Expected Utility Preferences," mimeo.

Siniscalchi, M. (2005), "A Behavioral Characterization of Plausible Priors," *Journal of Economic Theory*, forthcoming.

Siniscalchi, M. (2006), "Dynamic Choice under Ambiguity," mimeo.

Strotz, R.H. (1956), "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economic Studies* **23**, 165-180.

Tonks, I. (1983), "Bayesian Learning and the Optimal Investment Decision of the Firm," *The Economic Journal* **93**, 87-98.

Tversky, A., and D. Kahneman (1992), "Advances in Prospect Theory: Cumulative Representations of Uncertainty," *Journal of Risk and Uncertainty* **5**, 297-323.

Tversky, A., and P.P. Wakker (1995), "Risk Attitudes and Decision Weights," *Econometrica* **63**, 1255-1280.

Viscusi, W. K. (1985). A Bayesian Perspective on Biases in Risk Perception. Economics Letters 17, 59{62.

Viscusi, W. K., and C.J. O'Connor (1984), "Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?," *The American Economic Review* **74**, 942-956.

Wakai, K. (2002), "Linking Behavioral Economics, Axiomatic Decision Theory and General Equilibrium Theory," Ph.D. dissertation. Yale University.

Wakker, P.P. (2001), "Testing and Characterizing Properties of Nonadditive Measures through Violations of the Sure-Thing Principle," *Econometrica* **69**, 1039-1059.

Wakker, P.P (2004), "On the Composition of Risk Preference and Belief," *Psychological Review* **111**, 236-241.

Wakker, P.P, and A. Tversky (1993), "An Axiomatization of Cumulative Prospect Theory," *Journal of Risk and Uncertainty* **7**, 147-176.

Yaari, M.E. (1987), "The Dual Theory of Choice under Risk," *Econometrica* **55**, 95-115.