Reference Dependence and Attitudes towards Uncertainty *

Juan Sebastián Lleras[†], Evan Piermont[‡], Richard Svoboda[§]

March 11, 2015

Abstract

This paper characterizes a model of reference-dependence, where a state-contingent contract (act) is evaluated by its expected value and its expected gain-loss utility. The expected utility of an act serves as the reference point, hence gains (resp., losses) occur in states where the act provides an outcome that is better (worse) than expected. Beliefs, preferences over outcomes, and a degree of reference-dependence characterize the utility representation, and all are uniquely identified from behavior. Moreover, we establish a link between reference-dependence and attitudes towards uncertainty. In particular, we show that within our framework loss aversion and reference dependence are equivalent to max-min expected utility. Finally, we apply our model to sealed bid auctions where equilibrium bidding strategies are higher bids than risk-neutrality, and hence are more closely aligned with experimental findings.

Keywords: Reference Dependent Preferences, Endogenous Reference Points, Gain-Loss Attitudes, Subjective Expected Utility, Belief Distortion.

*We would like to thank David Ahn, Haluk Ergin, Shachar Kariv, Jian Li, Omar Nayeem, Luca Rigotti, Chris Shannon, and seminar participants at UC Berkeley, Pitt, NC State, FSU, and CMU for helpful discussion and comments. We are particularly grateful to Roee Teper for constant discussions and comments throughout all stages of this project.

[†] jslleras@gmail.com

[‡] ehp5@pitt.edu, University of Pittsburgh, Department of Economics

[§] rts33@pitt.edu University of Pittsburgh, Department of Economics

1 Introduction

In many circumstances, a decision maker (DM) may evaluate an uncertain prospect not only in absolute terms but also in relative relation to some reference point. Kahneman and Tversky [1979] first introduced the notion of reference-dependence, in the seminal *Prospect Theory*, to explain experimental violations of expected utility. Within Prospect Theory, deviations from the reference point are weighted by a gain-loss value function, which has the feature, referred to as *loss aversion*, that losses have more negative value than equal sized gains have positive value.

A different resolution for empirical deviations from expected utility proposes models of multiple priors, in particular *MaxMin Expected Utility (MMEU)*. MMEU, axiomatized by Gilboa and Schmeidler [1989] as an explanation of the Ellsberg Paradox[1], considers a DM who holds a family of beliefs regarding the likelihood of events. She evaluates uncertain prospects by the minimum expected utility consistent with any of her beliefs. As such, a MMEU DM displays *uncertainty aversion* (or, ambiguity aversion), the feature that she prefers to minimize her exposure to uncertainty.

At a purely intuitive level, there seems to be a connection between loss aversion and uncertainty aversion; both behaviors characterize some form of pessimism in comparison to a subjective expected utility (SEU) maximizer. A loss averse DM places more weight on the utility of "bad" events but leaves the probabilities undistorted, whereas a uncertainty averse DM places more weight on the probability of "bad" events but leaves the utilities undistorted. We show in this paper that this connection is more than superficial; there exists a formal connection between reference dependence and attitude towards uncertainty. In particular, we axiomatize a simple class of reference dependent preferences, called *asymmetric gain-loss (AGL) preferences*, which can be equivalently represented by a MMEU functional. Within our framework, loss aversion and uncertainty aversion produce identical choice data.[2]

1.1 AGL Prefernces

In addition to formalizing the connection between reference dependence and ambiguity aversion, AGL preferences provide a simple model of endogenous reference dependence. Our object of choice is a state contingent contract, or act, which is an assignment of

[1]Because it is well established in the literature, we refer the reader to Gilboa and Schmeidler [1989] for a formal discussion of the Ellsberg Paradox and its resolution by MMEU.

[2]Likewise, "loss seeking" attitudes and uncertainty seeking attitudes produce the same data.

outcomes to each state of the world. AGL preferences evaluate an act according to two components: *first-order expected utility*, and *gain-loss utility*. The first order expected utility of an act is as in the standard SEU model: the decision maker holds a subjective belief, μ, over the state-space, and a preferences over final outcomes given by a utility function, U. Her expected utility of an act f is

$$\mathrm{E}_\mu[U \circ f] = \int_S U(f(s))\mu(ds), \tag{1.1}$$

where S is the state space and $s \in S$ is a generic state. The AGL DM takes this assessment of acts both as the reference point by which gains and losses are measured, and as the baseline level of utility on which gains and losses act as distortions. The main result of this paper is the behavioral characterization of asymmetric gain-loss preferences, which are preferences that can be represented by the functional

$$V(f) = \underbrace{\mathrm{E}_\mu[U \circ f]}_{\text{First Order Expected Utility}} + \underbrace{\int_S \eta_\mu(U \circ f)\mu(ds)}_{\text{Expected Gain-Loss Utility}} . \tag{1.2}$$

The first term is the DM's subjective expected utility without any reference considerations, and the second is the gain-loss utility. In each state, η linearly scales the difference between the expected utility of the act and the utility of the outcome in that particular state. To capture the asymmetry between gains and losses, η can scale gains differently than losses. The scalars associated with gains and losses are λ_g and λ_l, respectively. Formally, for all $s \in S$, the gain-loss utility in that state is defined by

$$\eta_\mu(U \circ f)(s) = \begin{cases} \lambda_g\left(U(f(s)) - \mathrm{E}_\mu[U \circ f]\right) & \text{if} \quad U(f(s)) \geq \mathrm{E}_\mu[U \circ f] \\ \\ \lambda_l\left(\mathrm{E}_\mu[U \circ f] - U(f(s))\right) & \text{if} \quad U(f(s)) < \mathrm{E}_\mu[U \circ f]. \end{cases} \tag{1.3}$$

In our representation results for AGL preferences all the elements are identified from choice behavior: μ is unique, U is cardinally unique, and $\lambda = \lambda_g + \lambda_l$ is identified uniquely.

Note that in this model any gain-loss considerations take place because of uncertainty. The gain-loss function η expresses how the DM will feel for every realization of uncertainty compared to the reference point. Hence, in the absence of uncertainty, the DM behaves as an expected utility maximizer.

1.2 A Simple Example of Asymmetric Gain-Loss Preferences

We employ the following numerical example to explain the intuition behind the representation, and show how asymmetric gain-loss preferences can explain different types of behavior regarding uncertainty.

Consider the environment of a seller selling a single good to a buyer who makes a take-it-or-leave-it offer.[3] The value to the seller is $\hat{v}$ which the buyer believes is distributed on the interval $[0, 1]$ according to the cumulative distribution function F_1, where $F_1(x) = x$.[4] The buyer has an independent private value for the object given by $v \in [0, 1]$. The buyer will submit her offer, b, and the seller will accept or reject the offer. The seller will accept any offer which (weakly) exceeds her value. The buyer's utility associated with the bid b is given by

$$U(b) = \begin{cases} v - b & \text{if} \quad b \geq \hat{v} \\ 0 & \text{otherwise} . \end{cases}$$

If the buyer is a risk neutral expected utility maximizer then her optimal bid solves

$$\max_{b} \quad (v - b)F_1(b).$$

The optimal bid is $b^* = \frac{v}{2}$, which has an expected utility of $\frac{v^2}{4}$ before the bid is placed.

Now suppose that the buyer has gain-loss preferences: in addition to the expected value she wants to avoid losses, so she subtracts any expected losses from the expected first order utility to determine the valuation (this corresponds to the parametrization $\lambda_l = -1$ and $\lambda_g = 0$). A loss for her is any outcome where her ex-post utility is worse than the expected value; therefore, outcomes in which she does not obtain the item are considered losses.

Her expected AGL utility, taking into account her gain-loss preferences, of making the bid $b = \frac{v}{2}$ is given by:

$$\underbrace{\frac{v^2}{4}}_{\text{Expected Utility}} - \underbrace{\left(1 - F_1\left(\frac{v}{2}\right)\right)\left(\frac{v^2}{4} - 0\right)}_{\text{Expected Loss for } b^* = \frac{v}{2}},$$

which is equal to $\frac{v^3}{8}$. The first term is the first order expected utility and the second term is the expected losses. The optimal AGL bid, however, is $b^{**} = \frac{2v}{3}$. The corresponding expected first order utility would be $\frac{2v^2}{9}$, and her valuation, taking into account her gain-

[3]We further explore the implications of our model in standard auction environments in section 4.

[4]I.e., F_1 is the uniform distribution. The reason for this notation will become clear later in the example

loss preferences, would be

$$\frac{2v^2}{9} - \left(1 - \frac{2v}{3}\right)\left(\frac{2v^2}{9} - 0\right) = \frac{4v^3}{27},$$

which is larger than if she had submitted the optimal bid in the absence of any gain-loss considerations, b^*.

When the buyer in this simple take-it-or-leave-it example takes into consideration expected gains and losses in addition to the standard expected utility, she is better off increasing her bid. Intuitively, she sacrifices her payoff in good outcomes (where she obtains the item) in order to decrease the chance of bad outcomes (not obtaining the item). While her payoff is smaller contingent on obtaining the good, the outcome is favorable more often and she increases her ex-ante utility.

A similar behavior could be captured by a buyer who was averse to ambiguity. To see this, assume that the buyer, instead of having AGL preferences, has MMEU preferences. Further, the buyer believes that the distribution of the seller's valuation is contained in the set of cumulative distribution functions,

$$C = \{F_p : F_p(x) = x^p, x \in [0,1], p \in [1,2]\}.$$

As such, the buyer chooses a bid so as to maximize

$$\max_b \left(\min_{F_p \in C} (v - b)F_p(b) \right).$$

Since F_p first order stochastically dominates F_q when $p > q$ it follows that for any bid, the minimum attains at F_2, and hence, the optimal bid solves

$$\max_b (v - b)F_2(b),$$

which has a maximum at $b = \frac{2v}{3}$. Therefore a AGL bidder with $\lambda_l = -1$ and $\lambda_g = 0$ and a MMEU bidder with multiple priors given by C will make the same optimal bid. While at first glance this connection may seem contrived, we show in Section 3 that AGL behavior can always be equivalently described by a MMEU decision maker.

1.3 Reference Point Determination

There are two important features that set AGL preferences apart from other models of endogenous reference dependence, such as Shalev [2000], Kőzsegi and Rabin [2006] (henceforth KR) or Sarver [2011], and make it more tractable for use in applied work. First, in our model the DM has no control over the reference point: the reference point does not depend on any actions taken by the DM. The reference point is identified from beliefs and preferences over outcomes, neither of which are affected by reference dependence. In the aforementioned models, the DM is aware of her reference dependence and how her choices affect the reference point. These papers require that the DM's choices are optimal given the reference point they induce, and so, require an equilibrium condition to account for the mutual relationship between reference points and choices. The model presented here has the benefit that it captures gain-loss attitudes and elicits the reference point, reference effects, and beliefs simultaneously from choices in a way that is easy to compute or estimate, without the need of equilibrium conditions.

The second, and closely related difference, is that in our model, the domain of uncertainty if fixed across decision problems. KR provide the first decision-theoretic model of reference dependence that tackles the determination of an endogenous reference point. In their model, the uncertainty concerns the distribution over possible choice problems. On the other hand, in our model, uncertainty is over states that determine the outcomes of acts, which is the standard domain to analyze choice under uncertainty.

The identification of beliefs about the possible choice problems that a DM might face poses a challenge from a theoretical and empirical point of view. If the DM is uncertain about the problems she might face, she must form expectations about the environment long before seeing the choice set, let along making any choices. There is no standard choice setting that allows for the identification of these beliefs. Any observed choice for a DM in the KR framework will be the choice conditional on a reference point, and the reference point is always the chosen alternative, via the equilibrium requirement. This feedback between choices and the reference point can lead to intransitivity of the revealed preference, as shown in Gul and Pesendorfer [2006]. Therefore, under standard choice settings, beliefs about the possible choice sets (and hence the endogenous reference point) cannot be independently identified from behavior.

1.4 Structure of the Paper

The rest of the paper is structured as follows. Section 2 provides an axiomatic characterization of the preferences, discusses the concept of the alignment of acts, which is instrumental for the endogenous determination of a reference point, and formally defines the utility representation. Section 3 explores the link between gain-loss and ambiguity attitudes. In addition it provides comparative statics results. Section 4 contains an application of AGL preferences to sealed bid auctions and shows how AGL preference can be employed to explain experimental findings. Section 5 contains a literature review. Section 6 concludes. All proofs are in the appendix.

2 Axiomatization

In this section we formally present the choice environment and a set of axioms which prove to be necessary and sufficient for the representation presented later in this section.

Let $S = \{s_1, s_2, ..., s_n\}$ be a finite set of states of the world that represent all possible payoff-relevant contingencies for the DM; any $E \subseteq S$ is called an event. Define $\mathcal{E} = \mathcal{P}(S) \setminus \{\emptyset, S\}$ as the set of all non-trivial events. Let X be a prize space, and $\mathcal{L}$ be the set of distributions on X with finite support. Denote by $\mathcal{F} = \mathcal{L}^S$ the set of all acts, that is, functions $f : S \to \mathcal{L}$. Take the mixture operation on $\mathcal{F}$ as the standard pointwise mixture, where for any $\alpha \in [0,1]$, $\alpha f + (1-\alpha)g \in \mathcal{F}$ gives $\alpha f(s) + (1-\alpha)g(s) \in \mathcal{L}$ for any $s \in S$. Abusing notation, any $c \in \mathcal{L}$ can be identified with the constant act $c(s) = c$ for all $s \in S$. Let $\mathcal{F}_c \cong \mathcal{L}$ be the set of constant acts. Preferences on $\mathcal{F}$ are denoted by the binary relation $\succsim$; $\succ$ and $\sim$ represent respectively the asymmetric and symmetric components of $\succsim$. For each $f \in \mathcal{F}$, if there is some $c_f \in \mathcal{F}_c$, such that $f \sim c_f$, then call c_f the *certainty equivalent* of f. Before we can specify the behavioral restrictions on preference that correspond to the AGL utility representation, we need to consider some particular structures in the choice domain.

Balanced Pairs of Acts

A particularly important type of act to study AGL preference is given by those that provide perfect hedges against uncertainty. Hedging gets rid of uncertainty, and therefore it also removes all possible gain-loss considerations from the act. Call a pair of acts $(f, \bar{f})$

balanced if they provide a perfect hedge and are indifferent to each other.[5] The importance of balanced acts is that eliminating subjective gain-loss considerations allows an analyst to identify beliefs from preferences.

Definition 1. Two acts f and $\bar{f}$ are *balanced* if $f \sim \bar{f}$, and for any states $s, s' \in S$

$$\frac{1}{2}f(s) + \frac{1}{2}\bar{f}(s) \sim \frac{1}{2}f(s') + \frac{1}{2}\bar{f}(s')$$

If there exists $e_f \in \mathcal{F}_c$ such that $e_f = \frac{1}{2}f(s) + \frac{1}{2}\bar{f}(s)$ for all $s \in S$, we call e_f the *hedge* of f. $(f, \bar{f})$ is referred to as a *balanced pair*, and $\bar{f}$ is a *balancing act* of f (and vice-versa).

When the notation $\bar{f}$ is used, it is always in reference to the balancing act of $f \in \mathcal{F}$. The conditions imposed on preferences below guarantee that c_f and e_f are unique and well defined for each f.

Act Alignment: Separating Positive and Negative States

Balanced acts will provide a behavioral way of separating gains and losses. We require that when the outcome in state s is considered a gain for f, the outcome on state s is considered a loss for $\bar{f}$. This is a natural requirement given that f and $\bar{f}$ provide a perfect hedge to the DM. Hence, $\bar{f}$ has the exact opposite gain-loss composition of f. For an act, define positive states as those states that deliver gains, and negative states as those states that deliver losses.

Definition 2. Let $(f, \bar{f})$ be a balanced pair. Say $s \in S$ is a *positive state* for f if $f(s) \succsim \bar{f}(s)$, and a *negative state* for f if $\bar{f}(s) \succsim f(s)$. If a state is both positive and negative (i.e. $f(s) \sim \bar{f}(s)$) say s is a *neutral state* for f.

Any balanced pair of acts induces a set of partitions, each of which splits the state space into two events: the states where $f(s) \succsim \bar{f}(s)$ and states where $\bar{f}(s) \succsim f(s)$. Hence one event contains only positive states for f (negative for $\bar{f}$), and one event that contains only negative states for f (positive for $\bar{f}$). We use the convention that neutral states can be labeled as either positive or negative.[6] When there are no neutral states, each act has a unique way of partitioning the states into positive and negative. These partitions associated with each act are called the *alignment* of the act. We use the convention that

[5] Siniscalchi [2009] calls a pair of acts that provide perfect hedging as *complementary acts*. We strengthen the definition of complementary acts to further require the acts to be indifferent.

[6] We do not allow the neutral states to be labeled both positive and negative, instead when there are neutral states an balanced pair induces more than one partition.

the alignment of the act is represented by the event that includes the positive states E (the complement is the negative states) rather than saying that the alignment is represented by the partition $\{E, E^c\}$.[7]

Definition 3. For any $f \in \mathcal{F}$, say f is aligned with the event $E \in \mathcal{E}$ if for all $s \in E$, $f(s) \succsim \bar{f}(s)$ and for all $s \in E^c$, $\bar{f}(s) \succsim f(s)$.

For every $E \in \mathcal{E}$, there is a set of acts that is aligned with E.

Definition 4. Given any event $E \subset S$, define $\mathcal{F}^E$ be the set of acts where the positive states are contained in E, i.e.

$$\mathcal{F}^E = \{f \in H : \ \forall\ s \in E, f(s) \succsim \bar{f}(s) \text{ and } \forall s \notin E,\ \ \bar{f}(s) \succsim f(s)\}$$

Note that any constant act is its own balancing act, therefore constant acts are aligned with all partitions of the state space. It is useful to consider acts that have only one alignment, which are called *single alignment* acts. These acts are important because they are acts where small perturbations on outcomes do not change the alignment.

Definition 5. Given $f \in \mathcal{F}$, f is *single-alignment* act if for no $s \in S$, $f(s) \sim \bar{f}(s)$.

If the event E represents an alignment of f, every subset of E or E^c is called a *non-overlapping event*. These are the events where all the states are either all positive or all negative for f, so there is no overlap between positive and negative states for f. Non-overlapping events provide a way of specifying situations where there is no tradeoff between positive and negative states, only across one type of state.

Definition 6. Given $f \in \mathcal{F}$, event $F \subset S$ is a *non-overlapping event* for f if every state in F is aligned in the same way. F is non-overlapping for $f \in \mathcal{F}^E$ if

$$F \subseteq E \text{ or } F \subseteq E^c$$

Abusing terminology, we say that F is non-overlapping for E whenever F is non-overlapping for $f \in \mathcal{F}^E$.

With these definitions in mind we can now specify the behavioral restrictions that are necessary and sufficient to be represented by the AGL-functional, as given by equation (1.2).

[7] This makes the notation easier because there is no need to specify which cell in the partition represents the positive states and which cell represents the negative states.

Standard Axioms

The first 4 conditions, A1-A4, are standard axioms in the literature of choice under uncertainty.[8]

A1. (Weak Order). $\succsim$ is complete and transitive.

A2. (Continuity). For all $f \in H$, the sets $\{g \in H : g \succsim f\}$ and $\{g \in H : f \succsim g\}$ are closed.

A3. (Strict Monotonicity). If for all s, $f(s) \succsim g(s)$ then $f \succsim g$. If $f(s) \succ g(s)$ for some s and $f(s) \succsim g(s)$ for all s, then $f \succ g$.

A4. (Unboundedness). For any f, g in $\mathcal{F}$, and any $\alpha \in (0,1)$, there exists w and z in $\mathcal{F}_c$ satisfying $g \succ \alpha w + (1-\alpha)f$ and $\alpha z + (1-\alpha)g \succ f$.

New Axioms: Mixture Conditions

The standard subjective expected utility model from Anscombe and Aumann [1963] is characterized by some version of A1-A4, plus the independence axiom. Independence requires that $f \succsim g$ if and only if $\alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h$ for any $h \in \mathcal{F}$, and any $\alpha \in (0,1)$.

The independence axiom does not hold for AGL preferences because it does not allow for gains and losses to be evaluated differently; convex combinations of acts can change the gain-loss composition of acts, therefore changing the assessments as well. AGL preferences relax independence, but impose three consistency requirements for mixtures of acts.

The first new axiom states that as long as the alignment of acts remains the same when mixing, then independence is preserved. If two acts have the same alignment, taking any mixture of them does not change the composition of gains and losses, so the tradeoff between gains and losses should not change.

A5. (Alignment Independence). If $f, h \in \mathcal{F}^E$ and $g, h \in \mathcal{F}^F$ for some $E, F \in \mathcal{E}$, then $f \succsim g$ if and only if $\alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h$ for all $\alpha \in [0,1]$.

The second condition imposes structure on mixtures for pairs of acts that are not mutually aligned. The condition is motivated by two observations. First, when any

[8] Unboundedness implies non-triviality of preferences (see Kopylov [2007] or Grant and Polak [2011]), and it is a technical condition that guarantees that the range of the utility function over outcomes is $(-\infty, \infty)$. Strict monotonicity implies state-independence and that no state is null, i.e. the DM puts positive probability on every state occurring.

act h is neutral on an event E then mixing f with h should not change the gain-loss consideration of f within E directly. Second, for a single alignment act f, perturbing it slightly (i.e. mixing f with any other act, where the weight on f is close to 1) does not change the alignment of f if preferences are continuous and monotonic.

Consider an event E that is non-overlapping for both f and g. The new axiom states that slightly perturbing f and g on E only, which is represented by mixing with any h that is neutral on E^c, should have the same effect on f as on g (i.e. either the DM prefers the mixture of f and h to f and the mixture of g and h to g, or she prefers both f and g over the respective mixtures). In other words, the DM must not change the direction of the assessment of these particular mixtures, as it does not induce any tradeoff between positive and negative states for either act. This condition is called *Local Mixture Consistency*.

A6. (Local Mixture Consistency). For any single-alignment acts $f \in \mathcal{F}^E$ and $g \in \mathcal{F}^{E'}$, any event F, which is non-overlapping for *both* E, and E', and any $h \in \mathcal{F}$ such that $h(s) \sim \bar{h}(s)$ for all $s \notin F$, there exists $\alpha^* < 1$ such that

$$f \succsim \alpha f + (1-\alpha)h \iff g \succsim \alpha g + (1-\alpha)h$$

for all $\alpha \in (\alpha^*, 1)$.

The last axiom imposes a consistency condition on mixtures of acts when reversing the role of gains and losses. Intuitively, the condition requires that the effect of mixing h and f is the opposite to the effect of mixing h and $\bar{f}$. This condition is called *Antisymmetry*.

A7. (Antisymmetry). For any acts f and g such that $f \sim g$, for every $h \in H$, and for any $\alpha \in (0, 1]$, $\alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h$ implies

(i) $\alpha \bar{f} + (1-\alpha)h \precsim \alpha \bar{g} + (1-\alpha)h$ and

(ii) $\alpha f + (1-\alpha)\bar{h} \precsim \alpha g + (1-\alpha)\bar{h}$

where $\bar{f}$, $\bar{g}$, and $\bar{h}$ are balancing acts of f, g, and h respectively.

2.1 Representation Results

This section provides the main representation results of the paper. Theorem 2.1 introduces the AGL representation as characterized by the above axioms. This section also outlines important preliminary results that highlight the role of particular axioms and are helpful to elucidate the relation between the AGL representation and other decision theoretic models.

Theorem 2.1. *The following conditions are equivalent*

1. *$\succsim$ satisfies A1-A7.*

2. *There exists an affine function $U : \mathcal{L} \to \mathbb{R}$, a probability distribution $\mu \in \Delta(S)$, and a real number $\lambda = \lambda_g + \lambda_l$ such that*

$$\begin{array}{rcl} f \succsim g \iff \int_S \left((U \circ f) - \eta_\mu(U \circ f)\right) \mu(ds) & \geq & \int_S \left((U \circ f) - \eta_\mu(U \circ f)\right) \mu(ds) \\ \iff \quad \mathrm{E}_\mu[(U \circ f)] - \mathrm{E}_\mu[\eta_\mu(U \circ f)] & \geq & \mathrm{E}_\mu[(U \circ g)] - \mathrm{E}_\mu[\eta_\mu(U \circ g)] \end{array}$$

where $\eta_\mu : (U(\mathcal{L}))^{|S|} \times \Delta(S) \to \mathbb{R}^s$ is given by

$$\eta_\mu(U \circ f)(s) = \begin{cases} \lambda_l \left(U(f(s)) - \mathrm{E}_\mu[U \circ f]\right) & \text{if} \quad U(f(s)) \geq \mathrm{E}_\mu[U \circ f] \\ \lambda_g \left(\mathrm{E}_\mu[U \circ f] - U(f(s))\right) & \text{if} \quad U(f(s)) < \mathrm{E}_\mu[U \circ f] \end{cases}$$

Moreover, (i) μ is a unique probability distribution, (ii) $\lambda = \lambda_g + \lambda_l \in \mathbb{R}$ is unique, and $|\lambda| < \min_{s \in S} \left(\frac{1}{\mu(s)}\right)$, and (iii) U is unique up to positive affine transformations.

The parameter λ characterizes gain-loss asymmetry uniquely. It is impossible to identify a gain weight and a loss weight uniquely. For instance, in the case that $|\lambda_l| = |\lambda_g|$, the behavior will be the same as a SEU maximizer. The intuition for the result is natural since the evaluation of gains and losses is not absolute. The bound in λ is a consequence of strict monotonicity, since an act is never be deemed worse than the outcome on the worst possible state.

The parameter λ captures the difference between the weight placed on gains and the weight placed on losses, which for the representation is unique. An important application of the representation result from Theorem 2.1 is that it provides an index for reference dependence (λ) that is decoupled from risk attitudes, and can be easily estimated.

2.1.1 Sketch of the Proof and Preliminary Results

The result is proven in several steps. First, Lemma 2.2 provides a SEU representation on $\mathcal{F}^E$ established by axioms A1-A5 (that is, excluding Local Mixture Consistency and Antisymmetry). Next, Lemma 2.3 shows that this representation can be extended to aggregate preferences across families of mutually aligned acts. Lastly, we utilize the properties of axioms A6 and A7 to generate the final result.

Lemma 2.2. *(SEU Representation for $\succsim$ restricted to $\mathcal{F}^E$) For any $E \in \mathcal{E}$, let $\succsim_E$ be the restriction of $\succsim$ to $\mathcal{F}^E$. The following conditions are equivalent:*

1. *$\succsim$ satisfies A1-A4, and Alignment Independence (A5).*

2. *There exists an affine function $U_E : \mathcal{L} \to \mathbb{R}$, and a probability distribution μ_E, such that for $f, g \in \mathcal{F}^E$,*

$$f \succsim_E g \iff \underbrace{\int_S (U_E \circ f)\mu_E(ds)}_{\mathrm{E}_{\mu_E}[U \circ f]} \geq \underbrace{\int_S (U_E \circ g)\mu_E(ds)}_{\mathrm{E}_{\mu_E}[U \circ g]}$$

Moreover, (i) U_E is unique up to positive affine transformations, and (ii) $\mu_E \in \Delta(S)$ is a unique probability distribution over S, where $\mu_E(s) > 0$ for all s.

Lemma 2.2 shows the implication of Alignment Independence. The axiom states that along mutually aligned acts, independence is preserved. Hence for the subspace of all acts with an alignment represented by E, $\mathcal{F}^E$, the DM behaves as a SEU maximizer with the unique prior μ_E. Mixing acts that have the same alignment keeps the proportion of gains and losses the same and therefore the behavior maintains the linearity associated with expected utility.

Since constants can be aligned with any events, i.e. $\mathcal{F}_c \subset \mathcal{F}^E$ for all $E \in \mathcal{E}$, the representations for $\{\succsim_E\}_{E \in \mathcal{E}}$ can be used to find a representation for $\succsim$. Therefore preferences over $\mathcal{F}$ can be represented with the family of functionals $\{V_E(\cdot)\}_{E \in \mathcal{E}}$, where each $V_E(\cdot)$ represents $\succsim_E$ (from Lemma 2.2). Preferences over constants are the same across all alignments, hence the utility function U_E can be normalized to be the same for all families of mutually aligned acts, i.e. $U_E = U_{E'}$ for all $E, E' \in \mathcal{E}$.

Lemma 2.3. *The following conditions are equivalent:*

1. *$\succsim$ satisfies A1-A5.*

2. *There exists an affine function $U : \mathcal{L} \to \mathbb{R}$, and a set of probability distributions over S, $\{\mu_E\}_{E \in \mathcal{E}}$, such that for $f \in \mathcal{F}^E$ and $g \in \mathcal{F}^F$,*

$$f \succsim g \iff \int_S (U \circ f)\mu_E(ds) \geq \int_S (U \circ g)\mu_F(ds)$$

Moreover U is unique up to positive affine transformations, and the set $\{\mu_E\}_{E \in \mathcal{E}}$ is unique.

Every prior in $\{\mu_E\}_{E\in\mathcal{E}}$ is different and Alignment Independence does not imply any structure on the priors. To derive the main result from the representation of Lemma 2.3, Local Mixture Consistency and Antisymmetry are used to guarantee that every prior in the set $\{\mu_E\}_{E\in\mathcal{E}}$ can be written as functions of one unique prior μ.

Adding Local Mixture Consistency guarantees that for all $E \in \mathcal{E}$, the conditional distributions are the same for all μ_E when conditioned on any F which is non-overlapping with E. Hence for any E, E' in $\mathcal{E}$, $\mu_E(\cdot|F) = \mu_{E'}(\cdot|F)$ whenever F is non-overlapping for E and E'. Moreover, the axioms imply that there exists a unique distribution $\mu \in \Delta(S)$ that generates these conditionals: μ is the unique distribution that represents e_f for all $f \in \mathcal{F}$. In other words, if e_f is the hedge of f then $\int_S (U \circ f)\mu(ds) = U(e_f)$. This implies that for any $E \in \mathcal{E}$, μ_E can be written as a constant distortion γ_E^+ of μ on all $s \in E$ (positive states), and a constant distortion γ_E^- of μ on all $s \in E^c$ (negative states), i.e.

$$\mu_E(s) = \begin{cases} \gamma_E^+\mu(s) & \text{if} \quad s \in E \\ \gamma_E^-\mu(s) & \text{if} \quad s \in E^c \end{cases}$$

Antisymmetry implies a particular relationship between distributions indexed by complementary alignments, which is that for any $E, F \in \mathcal{E}$,

$$\mu_E + \mu_{E^c} = \mu_F + \mu_{F^c}$$

Using this observation we show that for $s \in E$ the distributions on μ_E and $\mu_{E\setminus s}$ depend only on s. In particular, whenever $s \in E \cap F$,

$$\mu_E - \mu_{E\setminus s} = \mu_F - \mu_{F\setminus s}$$

Then, using these conditions about the family $\{\mu_E\}_{E\in\mathcal{E}}$ we show that the difference between the distortion on negative states and positive states is always constant, thus $\gamma_E^- - \gamma_E^+ = \lambda$ for all $E \in \mathcal{E}$. Therefore it is possible to characterize any μ_E, as functions of μ, and this constant λ that captures the difference between the negative and positive distortions. That is,

$$\begin{aligned} &\mu_E(s) = \mu(s)\left(1 - \lambda\left(1 - \mu(E)\right)\right) && \text{if } s \in E && \text{and} \\ &\mu_E(s) = \mu(s)\left(1 + \lambda\left(\mu(E)\right)\right) && \text{if } s \in E^c && \end{aligned}$$

Finally this representation of μ_E is used to rewrite the representation from Lemma 2.3 in terms of μ, which yields the desired result.

3 Alternate Representation and Comparative Statics

3.1 Relation to Ambiguity Attitude

According to Theorem 2.1, the DM who abides by the AGL axioms is probabilistically sophisticated but displays some reference effect. That is, she holds some unique belief, μ, regarding the state space, and evaluates each act according to this belief and her preferences for outcomes. Nonetheless, Lemma 2.3 states that the same preferences can be represented by a family of distributions, $\{\mu_E\}_{E\in\mathcal{E}}$, each of which is a distortion of the original belief, μ. This alludes to a possible relationship between reference effects and attitudes towards uncertainty, which as classically been modeled by a DM who considers a (non-singleton) set of priors.

Maxmin Expected Utility (MMEU), axiomatized by Gilboa and Schmeidler [1989], considers a utility function U and a set of priors C, such that

$$f \succsim g \;\Leftrightarrow\; \min_{\mu\in C}\int_S (U\circ f)\mu(ds) \geq \min_{\mu\in C}\int_S (U\circ g)\mu(ds)$$

MMEU is characterized by two key conditions: *certainty independence* and *uncertainty aversion*. Certainty independence requires that $f \succsim g$ if and only if for all $\alpha \in [0,1]$, $\alpha f + (1-\alpha)c \succsim \alpha g + (1-\alpha)c$ where $c \in \mathcal{F}_c$. Mixing two acts with a common constant act does not reverse the preference between them. Since constant acts are aligned with every $E \in \mathcal{E}$, Alignment Independence implies certainty independence.

Uncertainty aversion requires that for all f, g such that $f \sim g$ for any $\alpha \in (0,1)$, $\alpha f + (1-\alpha)g \succsim f$. If uncertainty aversion is changed for *uncertainty seeking* preferences,[9] then the representation is a Maxmax representation, where the DM evaluates an act according to the prior that maximizes her expectations. It is clear from the representation that if $\lambda \leq 0$ then the DM is uncertainty averse, and if $\lambda \geq 0$ then she is uncertainty seeking.

Uncertainty aversion can be characterized as a preference for hedging as hedging reduces the exposure to uncertainty. While the (loss biased) AGL axioms do not explicitly require preferences to be uncertainty averse, the preference for hedging arises as a consequence. This is because hedging reduces the exposure to, not only uncertainty, but also negative states. Pushing the utility value in each state closer to the average has more effect on the negative states (because of the loss bias) and hence weakly improves the act.

The formal connection is captured by the following result, which states that asym-

[9]For all $f, g \in \mathcal{F}$, $f \sim g$ implies for all $\alpha \in (0,1)$ $f \succsim \alpha f + (1-\alpha)g$

metric gain-loss preferences always admit a Maxmin or Maxmax representation and that the set of priors C has a specific structure that is related to the distortion of the (unique) beliefs of the DM.

Theorem 3.1. *Suppose $\succsim$ admits an AGL representation (U, μ, λ). Then there exists a closed and convex set C of probability distributions on S such that*

(i) If $\lambda \geq 0$

$$f \succsim g \iff \max_{\nu \in C} \int_S (U \circ f)\nu(ds) \geq \max_{\nu \in C} \int_S (U \circ g)\nu(ds)$$

(ii) If $\lambda \leq 0$

$$f \succsim g \iff \min_{\nu \in C} \int_S (U \circ f)\nu(ds) \geq \min_{\nu \in C} \int_S (U \circ g)\nu(ds)$$

Where C is the unique convex set C defined by $C = \mathrm{conv}(\{\nu_E\}_{E \in \mathcal{E}})$, where for every $E \in \mathcal{E}$ and all $s \in S$

$$\nu_E(s) = \begin{cases} \mu(s)\,(1 - \lambda\mu(E^c)) & \text{if} \quad s \in E \\ \mu(s)\,(1 + \lambda\mu(E)) & \text{if} \quad s \in E^c \end{cases} \tag{3.1}$$

Theorem 3.1 has several implications. First, it shows that this form of reference-dependence is always tied to a particular attitude towards uncertainty. The preferences studied in this paper will always be either uncertainty averse or uncertainty seeking. Second, it gives a precise form to the belief distortion that takes place when gain-loss consideration affect a probabilistically sophisticated DM. The distorted belief keeps the relative likelihood of states among gains and among losses unchanged as the original prior; however depending on the sign of λ, it increases or decreases the total weight given to gains (and losses) proportionally to the baseline belief. This distortion is a function of the degree of reference dependence, λ, and the total probability that the DM assigns to gains and losses for each case. Whenever $\lambda < 0$ –losses are weighted more than gains– the DM will behave as if she is a SEU maximizer who always thinks losses are more likely than "she originally thought" (i.e. the probability given by μ).

3.2 Comparative Statics

This section capitalizes on the connection to uncertainty attitude to advance comparative statics results relating behavior to elements of the AGL representation. The structure given to the set of priors makes it possible to establish a comparative notion of reference-dependence even when DMs have different baseline priors and sets of beliefs.[10]

The first result provides a characterization of gain-biased and loss-biased preferences as absolute notions of reference dependence. The second result provides comparative notions of reference dependence with a characterization of "more gain-biased" and "more loss-biased" preferences.

A natural measure for the degree and direction of reference effects is the gap between e_f and c_f, the hedge and the certainty equivalent. In the standard SEU model, e_f and c_f are the same, so the SEU model as the baseline case for reference effects. This suggests the following definition.

Definition 7. Let $\succsim$ be a preference over $\mathcal{F}$. Say $\succsim$ is *gain-biased* if for all $f \in \mathcal{F}$, $c_f \succsim e_f$. Say $\succsim$ is *loss-biased* if for all $f \in \mathcal{F}$, $e_f \succsim c_f$.

For a DM with AGL preferences, the parameter λ captures the bias of the preferences uniquely. In addition, this bias towards gains or losses is related to attitudes towards ambiguity (see Theorem 3.1). For a DM with preferences that admit an AGL representation, gain-bias is equivalent to uncertainty seeking preferences, and loss-bias is equivalent to uncertainty averse preferences. This observation establishes a clear connection between the idea of "loss aversion" that has been prevalent since Prospect Theory, and uncertainty aversion.

Proposition 3.2. *Let $\succsim$ admit an AGL representation. Then the following are equivalent:*

1. *$\succsim$ is gain-biased (respectively, loss-biased)*
2. *$\succsim$ is uncertainty seeking (resp., uncertainty averse)*
3. *$\lambda \geq 0$. (resp., $\lambda \leq 0$)*

Beliefs influence the assessment of f, captured by c_f, and also the expectations of f, captured by the hedge e_f. If we want to be able to compare two DM's degree of reference dependence who can hold different beliefs we want to disentangle reference dependence from beliefs. To do this, we define $f \vee \bar{f}$, the join of a balanced pair $(f, \bar{f})$, as the act that gives the DM the best outcome between f and $\bar{f}$ for each $s \in S$; and show that

[10] We show on this section that to meaningfully compare degrees of reference dependence across DMs, these DMs need to have the same underlying preferences for outcomes.

any balanced pair, its join provides information to distinguish the reference effects from beliefs.[11]

Definition 8. Given any balanced pair $(f, \bar{f})$, define the act $f \vee \bar{f}$, the *join of* $(f, \bar{f})$ as

$$(f \vee \bar{f})(s) = \begin{cases} f(s) & \text{if} \quad f(s) \succsim \bar{f}(s) \\ \bar{f}(s) & \text{if} \quad \bar{f}(s) \succ f(s) \end{cases}$$

From the AGL representation, gain-loss utility depends on how much the act deviates state by state from e_f. $f \vee \bar{f}$ provides indicated the absolute value of the state by state deviations of f from e_f.[12] To determine the assessment of these deviations for each DM (which is given by the beliefs), we consider the hedge of the join, $e_{f \vee \bar{f}}$.

To capture reference-dependence behaviorally across DMs we need to focus on acts that have the same hedge.[13] This is the only way to isolate reference effects across DMs from preferences: if acts have different hedges the reference effects can be confounded by the beliefs.

The intuition behind our comparative notion of "more loss biased" is that the DM prefers an act f with smaller deviations from e_f because the expected losses are smaller. Conversely, a DM with gain-bias prefers acts with larger deviations. Since we want to consider acts that have the same hedge, the comparative notions of "more gain-biased" and "more loss-biased" depends on a possibly different act for each DM: f for DM 1 and g for DM 2. Then the fact that $e_{f \vee \bar{f}}$ is larger than $e_{g \vee \bar{g}}$ provides a sufficient information to conclude that DM 1 is willing to pay more for f than DM 2 is willing to pay for g. For the definition we use the notation where e^i_f denotes the hedge of f for DM i.

Definition 9. Given two preference orders $\succsim_1$ and $\succsim_2$, say that $\succsim_1$ is *more gain-biased* than $\succsim_2$ (and $\succsim_2$ is more *loss-biased* than $\succsim_1$) if for any f, g with $e^1_f = e^2_g$ and $e^1_{f \vee \bar{f}} \succsim_i e^2_{g \vee \bar{g}}$ for $i = 1, 2$, then for any $c \in \mathcal{F}_c$, $c \succsim_1 f$ implies $c \succsim_2 g$ and $c \succ_1 f$ implies $c \succ_2 g$.

The comparative notion of "more gain bias" or "more loss bias" is similar to the notion of uncertainty aversion in Ghirardato and Marinacci [2002]. For gain-loss preference the implication of comparative loss bias is the same as comparative uncertainty aversion, although in this case it holds only for acts that have the same hedge. This provides a behavioral way to compare reference effects across DMs.

[11] If we define the meet (minimum) of balance pair instead, all the arguments where we use the join will still follow (with the appropriate natural reversal of the relationships and intuition). Hence, on here we focus exclusively on the join.

[12] Note, $(f \vee \bar{f})(s) \succsim e_f$ for all s.

[13] And additionally need DMs to agree on the value of constant acts.

Theorem 3.3. *Let $\succsim_i$ admit an AGL representation given by (U_i, μ_i, λ_i) for $i = 1, 2$. If U_1 and U_2 are cardinally equivalent, then the following are equivalent:*

1. *$\succsim_1$ is more gain-biased than $\succsim_2$.*

2. *$\lambda_1 \geq \lambda_2$.*

Moreover, if $\mu_1 = \mu_2$ then so too is,

3. *For all $f \in \mathcal{F}$ and all $c \in \mathcal{F}_c$, $f \succsim_2 c$ implies $f \succsim_1 c$.*

Finally, if $\mu_1 = \mu_2$, and $\succsim_1$ and $\succsim_2$ are both gain biased (or loss biased) then so too is,

4. *$C_2 \subseteq C_1$ for the equivalent Maxmin/Maxmax representation from Theorem 3.1.*

The additional equivalences when $\mu_1 = \mu_2$ stem from the fact that when DMs have the same belief, then for any $f \in \mathcal{F}$, $e^1_f = e^2_f$. The belief completely determines the hedge of any act.[14] Therefore the implication of (3) and (4) of Theorem 3.3 is that when DMs have the same prior, loss bias is equivalent to the comparative notion of ambiguity aversion from Ghirardato and Marinacci [2002]. This is because the priors determine the hedge of an act and also the meet and the join, hence $e^1_f = e^2_f$ for all $f \in \mathcal{F}$. This further implies that whenever $\succsim_i$ is gain-biased or loss-biased for both DMs, the notion of loss bias is consistent with the representation of comparative ambiguity aversion derived from Gilboa and Schmeidler [1989] (that the more ambiguity averse DM should have a larger set of priors).

These comparative statics results establish an unexplored link between the absolute and comparative notions of gain or loss bias, and existing notions of uncertainty aversion which is worth further exploring. The initial motivation for studying uncertainty was due to the Ellsberg [1961] idea that DMs are not able to formulate unique probabilities over uncertain events. Many models with multiple priors have been developed to capture what is considered as "Ellsbergian behavior". Nonetheless, even if the DM is able to form a unique prior, having gain-loss considerations can appear to contaminate her prior in a way that gives rise to behavior embodied by some multiple priors model. Hence, for AGL preference a probabilistically sophisticated DM can appear to have multiple priors due to gain-loss asymmetry.

[14] See the proof of Theorem 2.1

4 AGL-Preferences in Sealed Bid Auctions

In this section we apply the concept of AGL-type reference dependence to sealed bid independent private value auctions. We show that the equilibrium bidding strategy where all bidders have AGL preferences is consistent with empirically verified overbidding (relative to risk neutral Nash equilibrium).

Lange and Ratan [2010] study the link between overbidding and reference dependence in greater depth than we do here. They examine an auction environment in which bidders have a KR type reference dependence. However, to tractably tailor KR to the application they must make a number of simplifying assumptions (see the discussion in section 1.3). Most importantly, they do not characterize the reference point as fixed point. Namely, they assume (as we do here), that for fixed bidding strategies of the other bidders, the reference point given bid b, is simply the expected utility of the auction given, b. So, in equilibrium, each bid is associated with a reference point and the collection of reference points affect the optimal bid, but which bid is optimal *does not* affect the individual reference points.

In spirit, Lange and Ratan [2010] are using AGL preferences (albeit in a more complicated multi-dimensional environment). Thus, it is not surprising that we find similar results. As such, this application is not so much intended to resolve the issue of overbidding in sealed bid auctions, as it is to reinforce the tractability and simplicity of the AGL representation when imported into larger models.

4.1 The Auction Environment

A single seller is selling an indivisible good. There are N potential buyers, represented by $i = \{1, \ldots, N\}$. Each buyer has a private value for the good, θ_i. Each θ_i is independently drawn from $[\underline{\theta}, \overline{\theta}] \subset \mathbb{R}^+$, with cumulative distribution function F. Denote the first derivative of the cumulative distribution function as f. The distribution is commonly known to all buyers. Furthermore, let $G(\theta)$ denote the probability of θ being the highest type out of N bidders, that is $F(\theta)^{N-1}$.

Bidders have AGL preferences over bid profiles, $b : [\underline{\theta}, \overline{\theta}]^N \to \mathbb{R}^N$, and a linear utility index $U_i(\cdot)$ defined over outcomes:

$$U_i(b) = \begin{cases} \theta_i - p_i(b) & \text{if the bidder wins the item and pays } p_i(b) \\ -p_i(b) & \text{if the bidder does not win the item and pays } p_i(b) \end{cases} \tag{4.1}$$

Note that without specifying the assignment rules of the auction, we do not have a full characterization of preference, as the mapping from bid profiles to equation (4.1) is undefined. Therefore we are abusing notation in assuming that U_i is defined over bid-functions, when it is actually defined over the composition of bid functions and, the as of yet unspecified, assignment rule. Given an assignment rule, each bid function constitutes a state contingent contract. In each state (a resolution of the type-space) the bidding profile specifies the bids, and the assignment rule specifies the outcomes to each bidder accordingly.

The reference point, as discussed in Section 1.1, is the expectation of this "first-order" utility with respect to the bidders' beliefs, μ_i. Gain-loss utility is greater than zero if the realized value of $U_i(\cdot)$ is larger than the reference point and less than zero if the value of $U_i(\cdot)$ is smaller than the reference point. Formally, the ex-post AGL utility from a bid profile, b, is as follows:

$$V_i(b) \equiv U_i(b) + \eta_\mu(U_i(b)) \tag{4.2}$$

where η_μ is the gain-loss function defined in equation (1.3). Unless otherwise noted, it is assumed that $\lambda < 0$: losses are weighted more than gains. Note also that if $\lambda_l = \lambda_g = 0$ then the model collapses into the classical case of risk neutral bidders.

We are interested pure strategy, monotone, symmetric Bayes-Nash equilibria. In such equilibria, the bidder with the highest valuation always wins the auction. Therefore, we can without loss of generality assume that a bidder's belief that she is going to win is equivalent to her belief that she is the bidder with the highest value, as given by Bayes' rule and the commonly known distribution over types, F. Thus, we can rewrite equation (4.2) as

$$V_i(b) \equiv U_i(b) + \eta_{F(\theta_{-i})}(U_i(b)) \tag{4.3}$$

Ex-ante, an optimal bid maximizes a bidders expected (with respect to F) AGL utility. The expected utility of bidder i can be written as:

$$E_{F(\theta_{-i})}\left[V_i(b_i)\right] \equiv$$

$$\underbrace{\int_{\theta_{-i}} U_i(b) dF(\theta_{-i})}_{E_{F(\theta_{-i})}[U_i(b)]} + \underbrace{\int_{\theta_{-i}} \eta_{F(\theta_{-i})}(U_i(b)) dF(\theta_{-i})}_{E_{F(\theta_{-i})}\left[U_i(b) - E_{F(\theta_{-i})}(U_i(b))\right]} \tag{4.4}$$

This is precisely the AGL-functional, defined equation (1.2). The first term is the expected utility and the second term is the expected gain-loss utility. Given the strategy of all other bidders, bidder i seeks to maximize (4.4) by choosing b_i (and hence choosing b subject to the restrictions of the other bidders).

Although a bidder's reference point is determined by the equilibrium strategies, the reference point itself is not an equilibrium condition (i.e., given fixed strategies of other bidders). Fixing the opponents strategies and the assignment rule, the choice of a bid is analogous to the choice from a set of acts. The acts pay according to the assignment rule and the probability of each state is determined by the strategies of other bidders, and the distribution over the type space. Each act carries with it a unique reference point (the expected first order utility) irrespective of which act is chosen.

As such we can employ the standard techniques in auction theory to solve for equilibrium bidding behavior. This, of course, would not be possible if the reference point was not uniquely identified, or depended of equilibrium conditions. Given the later, it would require that bidder i not only maximize his expected (reference dependent) utility subject to other bidders strategies, but also with respect to his own choice. In place of the relatively straightforward optimization problem we would be instead tasked with finding a fixed point.

4.2 Equilibrium Bidding Functions

Assume that the seller implements a first price auction. Each buyer, i, makes a bid b_i. Then buyer $j = \underset{i}{\operatorname{argmax}}\{b_i\}$ will receive the item, and pay his own bid, b_i; all other bidder pay nothing. Each bidder maximizes her expected utility, given by (4.4), according to her beliefs over other players strategies and types.

Theorem 4.1. *The unique symmetric monotone equilibrium is for a bidder of type θ to bid according to*

$$b^{AGL}(\theta) = \frac{\int_{\underline{\theta}}^{\theta} g(x)x dx}{G(\theta)\left[1 + \lambda(1 - G(\theta))\right]} + \frac{\int_{\underline{\theta}}^{\theta} g(x)x\left(\lambda\left(1 - 2G(x)\right)\right)dx}{G(\theta)\left[1 + \lambda(1 - G(\theta))\right]}$$

Moreover, if $\lambda \in (-1, 0)$ then there is a $\hat{\theta}$ such that for all $\theta \geq \hat{\theta}$, θ will overbid.

Denote the risk neutral SEU optimal bid of type θ as in a first price auction as $b^{RN}(\theta)$. Recall that symmetric bidding function is given by $b^{RN}(\theta) = \frac{\int_{\underline{\theta}}^{\theta} g(x)x dx}{G(\theta)}$. To see that there exists some threshold, above which all types will bid higher than the RNNE, note that when $\lambda \in (-1, 0)$, the first term is larger than b^{RN} for all θ. Thus, all bidders for which the second term is positive will unambiguously bid larger than the RNNE. Basic continuity arguments suffice to show that this set exists and is nontrivial.

The effect here is due to the asymmetry (between gains and losses) of the preferences. It is the relatively large disutility from not meeting the reference point that motivates overbidding. To see this, note that if $\lambda_g = -\lambda_l > 0$ the optimal bid is the same as the risk neutral case. As a bidder increases her bid, her reference point changes accordingly. Increasing the bid has two opposing effects on the reference point (and therefore on expected utility). First, as the bid increases, the expected payment is increasing. Thus, increasing the bid reduces profit. On the other hand, as the bid increases so does the bidder's probability of winning. Higher type bidders have a higher reference point, and so will have a tendency to overbid by more. The extra cost associated with the higher bid is, so to speak, payment to cover the now increased cost of losing.

We now turn to the mathematically simple environment where bidders types are uniformly distributed on the [0,1] interval. From these examples we can extract more intuition about the optimal bidding strategies. Additionally, most of the experimental evidence of overbidding arises from experiments where types uniformly distributed.

With this distribution in mind, we can greatly simplify the bidding functions.

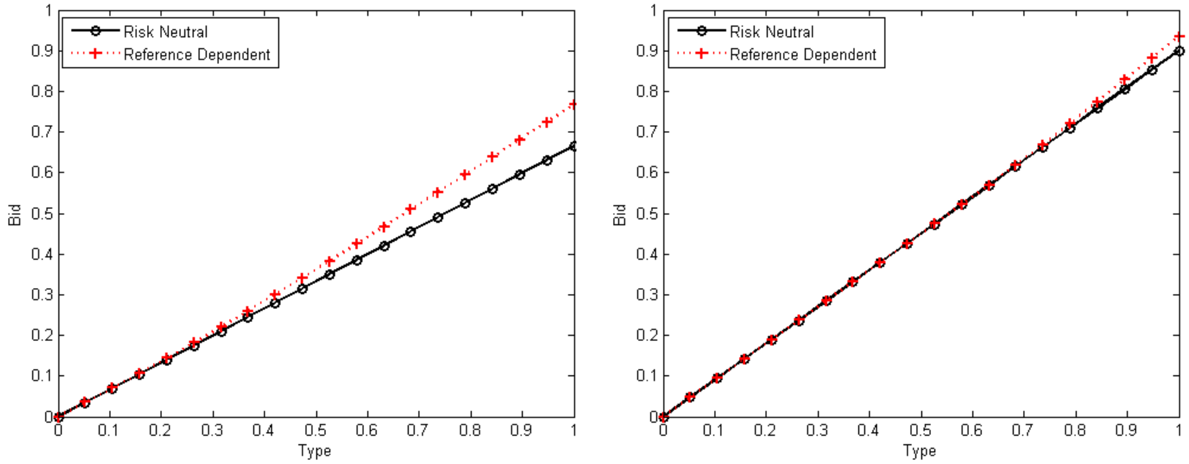
$$b^{RN}(\theta) = \frac{N-1}{N}\theta \qquad \text{and} \qquad b^{AGL}(\theta) = b^{RN}(\theta)\left[\frac{\left(1+\lambda\left[1-\frac{2N\theta^{N-1}}{(2N-1)}\right]\right)}{[1+\lambda(1-\theta^{N-1})]}\right]$$

From this we obtain the following result:

Corollary 4.2 (Theorem 4.1)**.** *When the type space is uniformly distributed and* $\lambda \in (-1, 0)$*, the optimal bid is weakly larger than* $b^{RN}(\theta)$ *for all types, and strictly larger for* $\theta > 0$.

Proof. This follows directly from setting $G(\theta) = \theta^{N-1}$, and verifying that the bracketed term in the simplified expression is always weakly greater than 1. □

By specifying the distribution, we can sign the difference in the optimal bid; AGL type preference can help to explain the persistence of overbidding in auctions. Figure 1 plots the optimal bids given $\lambda_g = \frac{1}{4}$, $\lambda_l = -1$, in auctions with 3 and 10 bidders. It is evident from the graph that the magnitude by which a type overbids is monotonically increasing in type. Further, the overbidding diminishes as the number of bidders increases



(a) $N = 3$. (b) $N = 10$.

Figure 1: Optimal Bids In a First Price Auction, $\lambda_g = \frac{1}{4},\ \lambda_l = -1$

–the reference effect is smaller as the probability of winning (and so, the reference point) is decreasing in the number of bidders.

5 Related Literature

This paper links two topics in the literature that have not been formally linked yet: reference-dependent preferences, and attitudes towards variation and ambiguity. We show that the idea of Kőzsegi and Rabin [2006] where the reference point depends on expectations, provides a clean way to link these two concepts in the domain of choice under uncertainty.

In many decision theory models, the status quo has been interpreted as a reference point. Giraud [2004a], Masatlioglu and Ok [2005], Sugden [2003], Sagi [2006], Rubinstein and Salant [2007], Apesteguia and Ballester [2009], Ortoleva [2010], Riella and Teper [2012] and Masatlioglu and Ok [2012] provide models of reference dependence, where the reference point is exogenously given. Along with Kőzsegi and Rabin [2006] and Kőszegi and Rabin [2007], three other papers that tackle the problem of endogenous reference point determination are Giraud [2004b], Ortoleva et al. [2014] and Sarver [2011].

The approach in Ortoleva et al. [2014] investigates reference point determination problem under a very general framework, where they do not need an equilibrium condition to characterize reference dependence. Nonetheless in their framework is impossible to identify reference points and reference effects uniquely. They identify reference dependence through WARP violations; their representation is a family of functionals. This difference is characterized by the source of reference dependence: in Ortoleva et al. [2014], the reference point is a function of the choice problem (the menu), whereas in the AGL model, the reference point is a function of the object of choice (the act).

For gain-loss preferences the evaluation of acts depends on the state by state variation of the act. Although some papers have studied attitudes toward variation in the context of risk and uncertainty, none relates such attitudes to reference-dependence. In the risk domain, Quiggin and Chambers [1998, 2004] measure attitudes towards risk, which depend on the expectation of the lottery and a risk index of the lottery that depends on the variation of the distribution. A different approach is taken by Gul [1991], who provides a theory of disappointment aversion. In Gul [1991], variation depends on the composition of a lottery among disappointment and elation outcomes.

In the uncertainty domain some models use the idea of a *reference prior*, which is a baseline for adjusting imprecise information, or to measure different specifications of a model like in the multiplier preferences of Hansen and Sargent [2001]. Wang [2003], Gajdos et al. [2004], Gajdos et al. [2008], Siniscalchi [2009], Strzalecki [2011] and Chambers et al. [2012] take this approach. In contrast, in this paper the DM has a unique prior over the states, but there is some contamination due to gain-loss considerations.

Grant and Polak [2011] provide a general model of mean-dispersion preferences, where deviations from the expectation affect utility as well. They show that many well-known families of preferences such as Choquet EU [Schmeidler, 1989], Maxmin EU [Gilboa and Schmeidler, 1989], invariant biseparable preferences [Ghirardato et al., 2004], variational preferences [Maccheroni et al., 2006], and Vector EU [Siniscalchi, 2009], belong to this family of preferences. The asymmetric gain-loss preferences belong to this family of preferences as well, since deviations from the reference point are in fact deviations from the expected utility of each act. However, the general mean-dispersion preferences are so general that it is not possible to provide clear comparative statics results like the ones presented here. Chambers et al. [2012] model mean-dispersion preferences where absolute uncertainty aversion is allowed to vary across acts, unlike in the mean-dispersion preferences of Grant and Polak [2011] where absolute uncertainty aversion is constant.[15]

[15] Other models where the uncertainty aversion is allowed to vary are Klibanoff et al. [2005] and Cerreia-Vioglio et al. [2011].

6 Conclusion

This paper provides a model of reference dependent preferences in the domain of choice under uncertainty, which is characterized by the family of preferences called asymmetric gain-loss preferences. This model provides a behavioral way to identify a unique reference point, and uniquely capture reference-dependent attitudes from choice behavior.

For asymmetric gain-loss preferences an act is evaluated according to its expected utility, and gain-loss utility which is measured using the expectations as a reference point. The AGL model is characterized by a triple (U, μ, λ) such that it captures reference-dependence in a way that deviates minimally from the standard SEU model. The only difference between the standard model and the AGL model is the parameter λ, which provides all the information about reference-dependence. In addition, the preferences studies in this model, allow for the behavioral comparison of gain-loss attitudes of different DMs who might have different beliefs, and also offer an clear behavioral characterization of gain-bias and loss-bias.

The identification of the reference point has the feature that the choices do not affect the reference point. Hence this model provides a simple tool to capture reference dependence from behavior without the need to consider equilibrium conditions which can be problematic to identify from choice data. Because of its simplicity, we are able to easily embed AGL preferences into the standard auction environment. AGL reference effect can help to explain empirical deviations from the SEU prediction in sealed bid auctions.

Moreover the AGL model establishes a simple and intuitive link between reference dependence and uncertainty attitudes. Whenever a DM exhibits gain-biased preferences she will be exhibit uncertainty averse preferences, and if she exhibits loss-biased preferences then she must exhibit uncertainty averse preferences. In addition even though the DM in this model has a unique prior over the states, which she uses to assess the reference point, gain-loss asymmetries can contaminate the beliefs in a way that she will behave as if she has multiple priors in mind.

References

F. J. Anscombe and R. J. Aumann. A definition of subjective probablity. *Annals of Mathematical Statistics*, 34:199–205, 1963.

J. Apesteguia and M. Ballester. A theory of reference-dependent behavior. *Economic Theory*, 40(3):427–455, September 2009.

S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio. Uncertainty averse preferences. *Journal of Economic Theory*, 146(4):1275–1330, July 2011.

R. Chambers, S. Grant, B. Polak, and J. Quiggin. A two parameter model of dispersion aversion. Aug. 2012.

D. Ellsberg. Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669, 1961.

T. Gajdos, J.-M. Tallon, and J.-C. Vergnaud. Decision making with imprecise probabilistic information. *Journal of Mathematical Economics*, 40(6):647–681, September 2004.

T. Gajdos, T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud. Attitude toward imprecise information. *Journal of Economic Theory*, 140(1):27–65, May 2008.

P. Ghirardato and M. Marinacci. Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102(2):251–289, February 2002.

P. Ghirardato, F. Maccheroni, and M. Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118(2):133–173, October 2004.

I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, April 1989.

R. Giraud. Reference-dependent preferences : rationality, mechanism and welfare implications. Sept. 2004a.

R. Giraud. Framing under risk : Endogenizing the reference point and separating cognition and decision. Cahiers de la maison des sciences economiques, Sept. 2004b.

S. Grant and B. Polak. Mean-dispersion preferences and constant absolute uncertainty aversion. Cowles Foundation Discussion Papers 1805, Cowles Foundation for Research in Economics, Yale University, June 2011.

F. Gul. A theory of disappointment aversion. *Econometrica*, 59(3):667–86, May 1991.

F. Gul and W. Pesendorfer. The revealed preference implications of reference dependent preferences. Mimeo, 2006.

L. P. Hansen and T. J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91(2):60–66, May 2001.

I. N. Herstein and J. Milnor. An Axiomatic Approach to Measurable Utility. 21(2): 291–297, 1953.

D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–91, March 1979.

B. Kőszegi and M. Rabin. Reference-dependent risk attitudes. *American Economic Review*, 97(4):1047–1073, September 2007.

B. Kőzsegi and M. Rabin. A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165, November 2006.

P. Klibanoff, M. Marinacci, and S. Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892, November 2005.

I. Kopylov. Procedural rationality in the multiple priors model. Mimeo, 2007.

A. Lange and A. Ratan. Multi-dimensional reference-dependent preferences in sealed-bid auctions–how (most) laboratory experiments differ from the field. *Games and Economic Behavior*, 68(2):634–645, 2010.

F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, November 2006.

Y. Masatlioglu and E. Ok. A canonical model of choice with initial endowments. Aug. 2012.

Y. Masatlioglu and E. A. Ok. Rational choice with status quo bias. *Journal of Economic Theory*, 121(1):1–29, March 2005.

P. Ortoleva. Status Quo Bias, Multiple Priors and Uncertainty Aversion. *Games and Economic Behavior*, 69:411–424, 2010.

P. Ortoleva, E. Ok, and G. Riella. Revealed (p)reference theory. *American Economic Review*, 2014. Forthcoming.

J. Quiggin and R. G. Chambers. Risk premiums and benefit measures for generalized-expected-utility theories. *Journal of Risk and Uncertainty*, 17(2):121–37, November 1998.

J. Quiggin and R. G. Chambers. Invariant risk attitudes. *Journal of Economic Theory*, 117(1):96–118, July 2004.

G. Riella and R. Teper. Probabilistic dominance and status quo bias. August 2012. Mimeo.

A. Rubinstein and Y. Salant. (a, f) choice with frames. Technical report, Jul 2007. Mimeo.

J. S. Sagi. Anchored preference relations. *Journal of Economic Theory*, 130(1):283–295, September 2006.

T. Sarver. Optimal anticipation. Mimeo, September 2011.

D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–87, May 1989.

J. Shalev. Loss aversion equilibrium. *International Journal of Game Theory*, 29(2):269–287, 2000.

M. Siniscalchi. Vector expected utility and attitudes toward variation. *Econometrica*, 77 (3):801–855, 05 2009.

E. Stiemke. *Matemathische Annalen*, 76(2-3):340–342, 1915.

T. Strzalecki. Axiomatic foundations of multiplier preferences. *Econometrica*, 79:47–73, 2011.

R. Sugden. Reference-dependent subjective expected utility. *Journal of Economic Theory*, 111(2):172–191, August 2003.

T. Wang. A class of multi-prior preferences. Mimeo, 2003.

A Appendix: Proofs

A.1 Proofs of the Main Results

This section provides proofs for the main results. The proofs for the auxiliary lemmas and propositions are in appendix A.2.

Proof of Lemma 2.2.

This is an obvious consequence of the Herstein and Milnor [1953] Mixture Space Theorem. Fix $E \in \mathcal{E}$. Since $\succsim$ satisfies Alignment Independence, $\mathcal{F}^E$ is convex, and it includes all the constant acts. Therefore $\succsim_E$ and $\mathcal{F}^E$ define a mixture space, so by the Mixture Space Theorem, the conditions for a SEU representation of $\succsim_E$ are satisfied. Therefore there exists a cardinally unique expected utility function $U_E : \mathcal{L} \to \mathbb{R}$, and an unique probability distribution $\mu_E : 2^S \to [0,1]$, such that for any $f, g \in \mathcal{F}^E$,

$$f \succsim g \iff V_E(f) \geq V_E(g)$$

Where

$$V_E(f) = \int_S (U_E \circ f)\mu_E(ds)$$

By strict monotonicity, $\mu_E(s) > 0$ for all s, so every state is non-null.

$\square$

Proof of Lemma 2.3.

From Lemma 2.2, for every $E \in \mathcal{E}$, $\mathcal{F}^E$ admits a subjective expected utility (SEU) representation given by (U_E, μ_E) [Anscombe and Aumann, 1963]. Since any constant $c \in \mathcal{F}_c$ is in $\mathcal{F}^E$ for all $E \in \mathcal{E}$, every U_E induces the same cardinal ranking of X. So it is possible to normalize all representations to have the same utility function U. Therefore for every $E \in \mathcal{E}$, preferences over $\mathcal{F}^E$, are represented by (U, μ_E) for any $E \in \mathcal{E}$. Let $V_E : \mathcal{F}^E \to \mathbb{R}$ be given as $V_E(f) = \int_S (U \circ f)\mu_E(ds)$. For $E \in \mathcal{E}$, and any $c \in \mathcal{F}_c$, $c \in \mathcal{F}^E$ $f \succsim c$ if and only if $V_E(f) = V_E(c) = U(c)$. Since every $f \in \mathcal{F}$ has a certainty equivalence c_f, and $\succsim$ is complete and transitive, for any $f \in \mathcal{F}^E$, $g \in \mathcal{F}^{E'}$, $f \succsim g$, if and only if $c_f \succsim c_g$, hence

$$f \succsim g \iff \int_S (U \circ f)\mu_E(ds) \geq \int_S (U \circ g)\mu_{E'}(ds)$$

proving the result.

$\square$

Proof of Theorem 2.1.

Start with the representation of Lemma 2.2 and Lemma 2.3, which is guaranteed by A1-A5. Hence there is a utility function $U : \mathcal{L} \to \mathbb{R}$ and a set of probability distributions over S, indexed by $\mathcal{E}$, $\{\mu_E\}_{E\in\mathcal{E}}$.

STEP 1: Show that for every E, E' the conditional distributions of μ_E and $\mu_{E'}$- conditional on any event F which is non-overlapping for E and E', are the same. And show that there is a unique distribution μ over S, that generates all the conditionals. In other words, there exists a unique μ, such that $\mu_E(\cdot|F) = \mu(\cdot|F)$ for all $E \in \mathcal{E}$ and F as long as F is non-overlapping for E. This is achieved by the addition of Local Mixture Consistency (A6).

Proposition A.1. *Suppose (U, μ_E) is a SEU representation $\succsim_E$ on $\mathcal{F}^E$ for all $E \in \mathcal{E}$. If $\succsim$ satisfies A1-A7, then there exists a unique distribution $\mu : 2^S \to [0, 1]$, such that for any $F, E \in \mathcal{E}$ such that $F \subseteq E$ or $F \cap E = \emptyset$, $\mu_E(\cdot|F) = \mu(\cdot|F)$, where for all $s \in F$, $\mu(s|F) = \frac{\mu(s)}{\mu(F)}$.*

Proof. In appendix A.2. □

By Proposition A.1, given $E \in \mathcal{E}$, for any $s, s' \in E$, $\frac{\mu(s)}{\mu(s')} = \frac{\mu_E(s)}{\mu_E(s')}$. This holds if for all $s \in E$, $\mu(s) = \gamma\mu_E(s)$, where $\gamma \in \mathbb{R}_{++}$. So on E, μ_E is just a constant perturbation of the original prior μ. Then the distribution μ_E can be written in the following way

$$\mu_E(s) = \begin{cases} \gamma_E^+\mu(s) & \text{if} \quad s \in E \\ \gamma_E^-\mu(s) & \text{if} \quad s \in E^c \end{cases}$$

where γ_E^+ represents how the original prior is perturbed on the positive states (i.e. E), and γ_E^- represents how the original prior is modified on the negative states. Both γ_E^+ and γ_E^- are positive numbers from monotonicity of $\succsim$, where $\gamma_E^+ \geq 1 \Leftrightarrow \gamma_E^- \leq 1$ for every E because μ_E is a probability distribution and the sum need to add up to 1. Therefore distributions if μ_E and $\mu_{E'}$ agree on the probability assigned to a state they are exactly the same.

Lemma A.2. *Given $E \neq E' \in \mathcal{E}$. Suppose $\mu_E(s) = \mu_{E'}(s)$ for some s, then $\mu_E = \mu_{E'}$.*

Proof. In appendix A.2. □

STEP 2: Adding Antisymmetry, implies some consistency between the distributions induced on $\mathcal{F}^E$ and $\mathcal{F}^{E^c}$. In which the average probability attached to each s is always the same for the pair μ_E, μ_{E^c}, or in other words the distortions on E and E^c exactly balance out.

Proposition A.3. *Let $\succsim$ satisfy A1-A7, then for any $E, F \in \mathcal{E}$,*

$$\mu_E + \mu_{E^c} = \mu_F + \mu_{F^c}$$

Proof. In appendix A.2. □

Lemma A.4. *Let $\succsim$ satisfy A1-A7, then for all $E \in \mathcal{E}$,*

$$\frac{\mu_E + \mu_{E^c}}{2} = \mu$$

Where μ_E is the distribution from Theorem 2.3 that represents preferences over $\mathcal{F}^E$.

Proof. This is an immediate consequence of Propositions A.1 and A.3. $\square$

From Lemma A.4, further conclude that $\gamma_E^+ + \gamma_{E^c}^- = 2$ for all $E \in \mathcal{E}$. A more relevant implication is that μ, uniquely defines e_f for all $f \in \mathcal{F}$. Recall that e_f is defined as the constant where $e_f = \frac{1}{2}f(s) + \frac{1}{2}\bar{f}(s)$ for all $s \in S$. Let $f \in \mathcal{F}^E$ and hence $\bar{f} \in \mathcal{F}^{E^c}$.

Proposition A.5. *Let $\succsim$ satisfy A1-A7, then for every $f \in \mathcal{F}$, $e_f \in \mathcal{F}_c$ is an act such that $U(e_f) = \int_S (U \circ f)\mu(ds)$.*

Proof. In appendix A.2. $\square$

Step 3: Show that the distribution induced on $\mathcal{F}^E$ and $\mathcal{F}^{E\setminus s}$ does not depend on E. Hence, the difference in measures only depend on the states they do not have in common.

Proposition A.6. *Let $\succsim$ satisfy A1-A7, then for any $E, F \in \mathcal{E}$ and $s \in S$, such that $S \in E \cap F$,*

$$\mu_E - \mu_{E\setminus s} = \mu_F - \mu_{F\setminus s} \tag{A.1}$$

Proof. In appendix A.2 $\square$

An immediate consequence of Proposition A.6, is the following extension of the result.

Lemma A.7. *Let $\succsim$ satisfy A1-A7, then for any $T \subsetneq E \cap F$,*

$$\mu_E - \mu_{E\setminus T} = \mu_F - \mu_{F\setminus T}$$

Step 4: Based on the previous results, provide a characterization of the distortions γ_E^+ and γ_E^- (3.1), as functions of μ and μ_E. And show that for any particular $E \in \mathcal{E}$, the difference between the negative and the positive distortion is always constant.

Proposition A.8. *If $\succsim$ satisfies A1-A7, then for any $E, F \in \mathcal{E}$, $\gamma_E^- - \gamma_E^+ = \gamma_F^- - \gamma_F^+$. Where γ_E^+, γ_E^- are defined as*

$$\gamma_E^+ = \frac{\mu_E(s)}{\mu(s)} \qquad \textit{for } s \in E$$
$$\gamma_E^- = \frac{\mu_E(s)}{\mu(s)} \qquad \textit{for } s \in E^c$$

Proof. In appendix A.2. $\square$

Therefore $\gamma_E^- - \gamma_E^+$ is a constant (independence of E), which can de defined as

$$\gamma_E^- - \gamma_E^+ \equiv \lambda$$

The next step is to characterize λ.

Proposition A.9. *If $\gamma_E^- - \gamma_E^+ = \lambda$ for all $E \in \mathcal{E}$, then for any $E \in \mathcal{E}$,*

$$\begin{aligned}
&\mu_E(s) = \mu(s)\,(1 - \lambda\,(1 - \mu(E))) && \text{if } s \in E && \text{and}\\
&\mu_E(s) = \mu(s)\,(1 + \lambda\,(\mu(E))) && \text{if } s \in E^c &&
\end{aligned} \tag{A.2}$$

Proof. In appendix A.2. □

Step 5: Use the definition of μ_E from Proposition A.9 into the representation from Lemma 2.3.

For for any $f \in \mathcal{F}^E$, then $V(f) = \int_S (u \circ f)\mu_E(s)$ which is equivalent to

$$\begin{aligned}
\int_S (U \circ f)\mu_E(s) &= \int_E (U \circ f)\mu(s)\,(1 - \lambda(1 - \mu(E))) + \int_{E^c} (U \circ f)\mu(s)\,(1 + \lambda\mu(E))\\
&= \int_E (U \circ f)\mu(s) - \lambda(1 - \mu(E)) \int_E (U \circ f)\mu(s)\\
&\quad + \int_{E^c} (U \circ f)\mu(s) + \lambda(\mu(E)) \int_{E^c} (U \circ f)\mu(ds)
\end{aligned}$$

$$\begin{aligned}
&= \int_S (U \circ f)\mu(ds) + \lambda \left(\mu(E) \int_E (U \circ f)\mu(ds) - \int_E (U \circ f)\mu(ds) \right)\\
&+ \lambda \left(\int_{E^c} (U \circ f)\mu(ds) - \mu(E^c) \int_{E^c} (U \circ f)\mu(ds) \right)\\
&= \mathrm{E}_\mu[U \circ f] + \lambda \left(\mu(E) \left(\mathrm{E}_\mu[U \circ f] - \int_{E^c} (U \circ f)\mu(ds) \right) - \int_E (U \circ f)\mu(ds) \right)\\
&+ \lambda \left(\int_{E^c} (U \circ f)\mu(ds) - \mu(E^c) \int_{E^c} (U \circ f)\mu(ds) \right)\\
&= \mathrm{E}_\mu[U \circ f] + \lambda \left(\mu(E)\mathrm{E}_\mu[U \circ f] - \int_E (U \circ f)\mu(ds) + \int_{E^c} (U \circ f)\mu(ds) - \mu(E^c)\mathrm{E}_\mu[U \circ f] \right)\\
&= \mathrm{E}_\mu[U \circ f] + \lambda \left(\int_E (\mathrm{E}_\mu[U \circ f] - (U \circ f))\,\mu(ds) + \int_{E^c} ((U \circ f) - \mathrm{E}_\mu[U \circ f])\,\mu(ds) \right)
\end{aligned} \tag{A.3}$$

Step 6: Show that λ is unique, but the specific weights on gains and losses are not unique. Since representation is the same for any λ_l, λ_g and λ_l', λ_g' such that $\lambda_g + \lambda_l = \lambda_g' + \lambda_l'$. So λ captures everything that can be identified about gain-loss asymmetry.

Let the function $\eta_\lambda' : U(\mathcal{L})^{|S|} \times \Delta(S) \to [0,1]^{|S|}$ be defined as

$$\eta_\lambda'(s)(v, \mu) = -\lambda |v_s - \mathrm{E}_\mu[v]|$$

For any $f \in \mathcal{F}^E$, for all $s \in E$, $\mathrm{E}_\mu[U \circ f] - U(f(s)) \leq 0$ and for all $s \in E^c$, $U(f(s)) -$

$\mathrm{E}_\mu[U \circ f] \leq 0$. Moreover,

$$\int_S \left((U \circ f) - \left(\int_S (U \circ f) \mu(ds) \right) \right) \mu(ds) = 0$$

Hence (A.3) can be written as,

$$\int_S (U \circ f) \mu_E(ds) = \mathrm{E}_\mu \left[U \circ f + \frac{\lambda}{2} \left| (U \circ f) - \mathrm{E}_\mu[U \circ f] \right| \right] \tag{A.4}$$

Moreover, it follows from the representation that η'_λ can be written as η_μ as defined in Theorem 2.1 for any pair (λ_g, λ_l), such that $\lambda_g + \lambda_l = \lambda$. To see this, note that

$$\int_E \left((U \circ f) - \mathrm{E}_\mu[U \circ f] \right) \mu(ds) = \int_{E^c} \left(\mathrm{E}_\mu[U \circ f] - (U \circ f) \right) \mu(ds)$$

Therefore for any $\lambda = \lambda_g + \lambda_l$, whenever $f \in \mathcal{F}^E$,

$$\begin{aligned}
\frac{\lambda}{2} \mathrm{E}_\mu[|(U \circ f) - \mathrm{E}_\mu(U \circ f)|] &= \left(\frac{\lambda_g + \lambda_l}{2} \right) \left(\int_E \left((U \circ f) - \mathrm{E}_\mu[U \circ f] \right) \mu(ds) \right) \\
&\quad + \left(\frac{\lambda_g + \lambda_l}{2} \right) \left(\int_{E^c} \left(E_\mu[U \circ f] - (U \circ f) \right) \mu(ds) \right) \\
&= (\lambda_g + \lambda_l) \int_E \left((U \circ f) - \mathrm{E}_\mu[U \circ f] \right) \mu(ds) \\
&= \lambda_g \int_E \left((U \circ f) - \mathrm{E}_\mu[U \circ f] \right) \mu(ds) + \lambda_l \int_{E^c} \left(\mathrm{E}_\mu[U \circ f] - (U \circ f) \right) \mu(ds) \\
&= \mathrm{E}_\mu[\eta_{\lambda_l}^{\lambda_g}(U \circ f)]
\end{aligned}$$

U is cardinally unique since it represents preferences over constant acts as a result of the Mixture Space Theorem. μ is unique, and moreover every μ_E is unique as well. This implies that $\lambda = \gamma_E^- - \gamma_E^+$ is unique as well from the definition of γ's from (A.28) and (A.29).

□

Proof of Theorem 3.1.

It suffices to show that for any $f \in \mathcal{F}^E$, if $\lambda < 0$,

$$\mu_E = \arg \min_{\nu \in \{\mu_E\}_{E \in \mathcal{E}}} \left(\int_S (U \circ f) \nu(ds) \right)$$

whereas if $\lambda > 0$,

$$\mu_E = \arg \max_{\nu \in \{\mu_E\}_{E \in \mathcal{E}}} \left(\int_S (U \circ f) \nu(ds) \right)$$

If this is true, then the result follows from Lemma 2.3.

Consider the case where $\lambda > 0$, let $f \in \mathcal{F}^E$. Suppose $v \in \mathbb{R}^{|S|}$ is the utility vector assigned to f, i.e. $U \circ f = v$. Since μ is the unique distribution that determines e_f as shown in the proof of Theorem 2.1, then for all $s \in E$, $v_s \geq \int_S v\mu(ds)$, and for all $s \in E^c$, $v_s \leq \int_S v\mu(ds)$. From Theorem 2.1, for any $E \in \mathcal{E}$,

$$\mu_E(s) = \begin{cases} \mu(s)\,(1 - \lambda\mu(E^c)) & \text{if} \quad s \in E \\ \mu(s)\,(1 + \lambda\mu(E)) & \text{if} \quad s \in E^c \end{cases} \tag{A.5}$$

Want to show that $\int_S v\mu_E(ds) \leq \int_S v\mu_F(ds)$, for any $F \in \mathcal{E}$, $F \neq E$. This is equivalent to showing that for all F, $\int_S v(\mu_E - \mu_F)(ds) \leq 0$. For any E, F, $\mu_E(s) - \mu_F(s)$ depends on whether $s \in E \cap F$, $E^c \cap F$, $E \cap F^c$, or $E^c \cap F^c$. To calculate $\mu_E - \mu_F$, use the following table

s	$\frac{\mu_E(s)}{\mu(s)}$	$\frac{\mu_F(s)}{\mu(s)}$
$E \cap F$	$(1 - \lambda\mu(E^c))$	$(1 - \lambda\mu(F^c))$
$E \cap F^c$	$(1 - \lambda\mu(E^c))$	$(1 + \lambda\mu(F))$
$E^c \cap F$	$(1 + \lambda\mu(E))$	$(1 - \lambda\mu(F^c))$
$E^c \cap F^c$	$(1 + \lambda\mu(E))$	$(1 + \lambda\mu(F))$

Hence, $\mu_E - \mu_F$ can be calculated by simplifying the expressions above.

s	$\frac{\mu_E(s)}{\mu(s)} - \frac{\mu_F(s)}{\mu(s)}$
$E \cap F$	$\lambda(\mu(E) - \mu(F))$
$E \cap F^c$	$\lambda(\mu(E) - \mu(F) - 1)$
$E^c \cap F$	$\lambda(\mu(E) - \mu(F) + 1)$
$E^c \cap F^c$	$\lambda(\mu(E) - \mu(F))$

So for $v = U \circ f$, if $\lambda > 0$ (assume $\lambda = 1$ for notational convenience).

$$\begin{aligned}
&\int_S v(\mu_E - \mu_F)(ds) = \sum_{E\cap F} v_t\mu(t)\,(\mu(E) - \mu(F)) + \sum_{E^c\cap F^c} v_t\mu(t)\,(\mu(E) - \mu(F)) \\
&+ \sum_{E\cap F^c} v_t\mu(t)\,(\mu(E) - \mu(F) - 1) + \sum_{E^c\cap F} v_t\mu(t)\,(\mu(E) - \mu(F) + 1) \\
&= \sum_S v_t\mu(t)\,(\mu(E) - \mu(F)) - \sum_{E\cap F^c} v_t\mu(t) + \sum_{E^c\cap F} v_t\mu(t) \\
&\leq \sum_S v_t\mu(t)\,(\mu(E) - \mu(F)) - \sum_{E\cap F^c}\left(\sum_S v_t\mu(t)\right)\mu(t) + \sum_{E^c\cap F}\left(\sum_S v_t\mu(t)\right)\mu(t) \\
&= \sum_S v_t\mu(t)\,(\mu(E) - \mu(F)) - \left(\sum_S v_t\mu(t)\right)(\mu(E\cap F^c)) + \left(\sum_S v_t\mu(t)\right)(\mu(E^c\cap F)) = 0
\end{aligned}$$

This holds because $\mu(E) - \mu(F) = \mu(E \cap F^c) - \mu(E^c \cap F)$. For the case where $\lambda < 0$ the inequality are just reversed, hence μ_E is the maximizer rather than the minimizer. It follows that for $f \in \mathcal{F}^E$, for any $\nu \in \text{conv}\,(\{\mu_E\}_{E\in\mathcal{E}})$, $\int_S(U\circ f)\nu(ds) \geq (\leq) \int_S(U\circ f)\mu_E(ds)$

for $\lambda > 0 (\lambda < 0)$. Then representation from Lemma 2.3, implies that for $\lambda > 0$,

$$f \succsim g \iff \min_{\nu \in C} \int_S (U \circ f)\nu(ds) \geq \min_{\nu \in C} \int_S (U \circ g)\nu(ds) \tag{A.6}$$

and for $\lambda < 0$,

$$f \succsim g \iff \max_{\nu \in C} \int_S (U \circ f)\nu(ds) \geq \max_{\nu \in C} \int_S (U \circ g)\nu(ds) \tag{A.7}$$

Where $C = \text{conv}(\{\mu_E\}_{E \in \mathcal{E}})$, and μ_E is defined as in (A.5). Proving the result.

□

Proof of Proposition 3.2.

We will Show only the first set of equivalences. The proof for the second set is exactly the same, where the inequalities are switched. Note that from the proof of Theorem 2.1, equation (A.4) implies that the functional $V(f)$ from the AGL representation can be rewritten as

$$V(f) = \int_S (U \circ f)\mu(ds) + \frac{\lambda}{2} \int_S |(U \circ f) - (U \circ e_f)|\mu(ds) \tag{A.8}$$

(i) $\Rightarrow$ (iii). Let $\succsim$ be gain-biased, then for any $f \in \mathcal{F}$, $c_f \succsim e_f$. From the AGL representation and equation (A.8) for all $f \in \mathcal{F}$,

$$\int_S (U \circ f)\mu(ds) + \frac{\lambda}{2} \int_S |(U \circ f) - (U \circ e_f)|\mu(ds) \geq U(e_f) = \int_S (U \circ f)\mu(ds)$$

Since $U(e_f) = \int_S (U \circ f)\mu(ds)$, this implies $\frac{\lambda}{2} \int_S |U \circ f - U \circ e_f|\mu(ds) \geq 0$, which holds only if $\lambda \geq 0$.

(iii) $\Rightarrow$ (ii). Let $\lambda \geq 0$. From the equivalent representation in Theorem 3.1, $\succsim$ also admits a Maxmax representation. Therefore from $\succsim$ is uncertainty seeking, since it is a necessary condition for this representation.

(ii) $\Rightarrow$ (i). Let $\succsim$ be uncertainty seeking. Then for all f,

$$e_f = \frac{1}{2}f + \frac{1}{2}\bar{f} \precsim f \sim c_f$$

since $f \sim \bar{f}$ by definition. Hence $e_f \precsim c_f$, which implies that $\succsim$ is gain-biased.

□

Proof of Theorem 3.3.

Given that $V(f)$ from the AGL representation can be written as

$$V(f) = \int_S (U \circ f)\mu(ds) + \frac{\lambda}{2}\int_S |(U \circ f) - (U \circ e_f)|\mu(ds)$$

For any $s \in S$, $f \vee \bar{f}(s) \succsim e_f$, hence $U(f \vee \bar{f}(s)) = U(e_f) + d_s$, since $\frac{1}{2}f(s) + \frac{1}{2}\bar{f}(s) = e_f$. Also $d_s \geq 0$, where $d_s = |U(f(s)) - U(e_f)| = |U(\bar{f}(s) - U(e_f)|$. Use the notation that superscripts denote the DM, e.g, e_f^i is the hedge of f for DM i. Moreover since U_1 and U_2 are cardinally equivalent, can assume $U_1 = U_2 = U$.

(i) $\Rightarrow$ (ii). Let $\succsim_1$ be more gain-biased than $\succsim_2$. Consider any $f, g \in \mathcal{F}$ such that $e_f^1 = e_g^2$.[16] Suppose in addition that $e^1_{f\vee\bar{f}} \succsim_i e^2_{g\vee\bar{g}}$ for $i = 1, 2$. By the AGL representation result, μ_i determines $U(e_f^i)$ for all $f \in \mathcal{F}$. Since $U_1 = U_2 = U$, then $e^1_{f\vee\bar{f}} \succsim_i e^2_{g\vee\bar{g}}$ for $i = 1, 2$ implies

$$\int_S (U \circ e^1_{f\vee\bar{f}})\mu_1(ds) \geq \int_S (U \circ e^2_{g\vee\bar{g}})\mu_2(ds)$$

From the previous observation about d_s, for all $s \in S$, $U(f \vee \bar{f}) = U(e_f) + |U(f(s) - U(e_f)|$, hence

$$\begin{aligned}
&\int_S ((U \circ e_f^1) + |U(f(s)) - U(e_f^1)|)\mu_1(ds) && \geq \int_S ((U \circ e_g^2) + |U(f(s)) - U(e_g^2)|)\mu_2(ds)\\
\Rightarrow\ & U(e_f^1) + \int_S |U(f(s)) - U(e_f^1)|\mu_1(ds) && \geq U(e_g^2) + \int_S |U(f(s)) - U(e_g^2)|\mu_2(ds)\\
\Rightarrow\ & \int_S |U(f(s)) - U(e_f^1)|\mu_1(ds) && \geq \int_S |U(f(s)) - U(e_g^2)|\mu_2(ds)
\end{aligned}$$

Now suppose that $g \succsim_2 c$ for some $c \in \mathcal{F}_c$, implies that $f \succsim_1 c$. Using equation (A.8)

$$V(g) = \int_S (U \circ g)\mu_2(ds) + \frac{\lambda_2}{2}\int_S |U(g(s)) - U(e_g^2)|\mu_2(ds) \geq U(c)$$

implies

$$V(f) = \int_S (U \circ g)\mu_1(ds) + \frac{\lambda_1}{2}\int_S |U(g(s)) - U(e_g^2)|\mu_1(ds) \geq U(c)$$

This implies that $V(f) \geq V(g)$ for all such f and g, where $e_f^1 = e_g^2$ and $e^1_{f\vee\bar{f}} \succsim_i e^2_{g\vee\bar{g}}$.

[16] By monotonicity and U_1, U_2 being cardinally equivalent we can simplify $e_f^i \sim_i e_g^2$ for $i = 1, 2$.

Since $e^1_f = e^2_g$, for any $f, g \in \mathcal{F}$,

$$\begin{aligned}
V(g) &= \int_S (U \circ g)\mu_2(ds) + \frac{\lambda_2}{2}\int_S |U(g(s)) - U(e^2_g)|\mu_2(ds) \\
&= U(e^2_g) + \frac{\lambda_2}{2}\int_S |U(g(s)) - U(e^2_g)|\mu_2(ds) \\
&= U(e^1_f) + \frac{\lambda_2}{2}\int_S |U(g(s)) - U(e^2_g)|\mu_2(ds) \\
&\leq U(e^1_f) + \frac{\lambda_2}{2}\int_S |U(f(s)) - U(e^1_f)|\mu_1(ds) \\
&\leq U(e^1_f) + \frac{\lambda_1}{2}\int_S |U(f(s)) - U(e^1_f)|\mu_1(ds) \ = \ V(f)
\end{aligned}$$

This holds even in the case of $e_{f\vee\bar{f}} = e_{g\vee\bar{g}}$, which given the previous result implies that $\int_S |U(g(s)) - U(e^2_g)|\mu_2(ds) = \int_S |U(f(s)) - U(e^1_f)|\mu_1(ds)$. Hence $\lambda_1 \geq \lambda_2$.

(ii) $\Rightarrow$ (i). Let $\lambda_1 \geq \lambda_2$. Let $f, g \in \mathcal{F}$ be such that $e^1_f = e^2_g$, and $e^1_{f\vee\bar{f}} \succsim e^2_{g\vee\bar{g}}$. Suppose for some $c \in \mathcal{F}_c$, $g \succsim_2 c$. Then by the AGL representation and the equivalent representation of (A.8),

$$\begin{aligned}
V(g) &= \int_S (U \circ g)\mu_2(ds) + \frac{\lambda_2}{2}\int_S |U(g(s)) - U(e^2_g)|\mu_2(ds) \\
&= U(e^2_g) + \frac{\lambda_2}{2}\int_S |U(g(s)) - U(e^2_g)|\mu_2(ds) \ \geq \ V(c)
\end{aligned}$$

Since $e^1_f = e^2_g$, then $U(e^2_g) = U(e^1_f)$, and $e^1_{f\vee\bar{f}} \succsim_i e^2_{g\vee\bar{g}}$,

$$\int_S |U(g(s)) - U(e^2_g)|\mu_2(ds) \leq \int_S |U(f(s)) - U(e^1_f)|\mu_1(ds)$$

Then

$$\begin{aligned}
V(g) &= U(e^2_g) + \frac{\lambda_2}{2}\int_S |U(g(s)) - U(e^2_g)|\mu_2(ds) \\
&\leq U(e^1_f) + \frac{\lambda_2}{2}\int_S |U(f(s)) - U(e^1_f)|\mu_1(ds) \\
&\leq U(e^1_f) + \frac{\lambda_1}{2}\int_S |U(f(s)) - U(e^1_f)|\mu_1(ds) \ = \ V(f)
\end{aligned}$$

Therefore $V(f) \geq U(c)$, showing that $f \succsim_1 c$.

$\square$

Proof of Theorem 4.1.

Each player of type θ_i solves the following maximization problem:

$$\max_{b_i} V_i(b_i) \equiv \int_{\theta_{-i}} U_i(b_i)dF(\theta_{-i}) + \int_{\theta_{-i}} \eta \left(U_i(b_i) - \left(\left(\int_{\theta_{-i}} U_i(b_i)dF(x_{-i}) \right) \right) \right) dF(\theta_{-i})$$

We utilize the fact that we are searching for a symmetric and monotone equilibrium. Symmetry dictates that each bid has a one-to-one mapping with a type. Monotonicity dictates that, assuming all other bidders $j \neq i$ bid according to the equilibrium strategy, the probability of bidder i winning the auction is equal to the probability that the virtual type (the type associated with i's bid, b_i) is the larger than $max_{j \neq i}\theta_j$, that is, $G(b^{-1}(b_i))$.

A loss-biased bidder will never bid above her type. This can be seen by noting that a bid, $b_i > \theta_i$ is strictly dominated by $b_i - \epsilon$ for some small enough ϵ. This is because lowering the bid reduces the likelihood of a losing state (winning the auction and paying more than her type –a strictly negative payoff) and weakly increase the payoff in all states.

Therefore, the reference point for any equilibrium bidding strategy is bounded weakly between 0 and the difference between the bidders type and her bid. Thus, we know that states in which the bidder loses the auction are considered losses and states where the she wins the auction are considered gains. Using these facts we can re-write the above maximization problem

$$\begin{gathered} \max_{b_i} \;\; G(b^{-1}(b_i))\left(\theta_i - b_i\right) + \\ G(b^{-1}(b_i))\lambda_g \left(\theta_i - b_i - G(b^{-1}(b_i))\left(\theta_i - b_i\right)\right) + \left(1 - G(b^{-1}(b_i))\right)\lambda_l \left(G(b^{-1}(b_i))\left(\theta_i - b_i\right)\right) \end{gathered}$$

The first order condition is:

$$\begin{gathered} -G(b^{-1}(b_i)) + G(b^{-1}(b_i))(\lambda_g + \lambda_l)\left(-1 + G(b^{-1}(b_i)) - \tfrac{g(b^{-1}(b_i))}{b'(b^{-1}(b_i))}\left(\theta_i - b_i\right)\right)\dots \\ +(\lambda_g + \lambda_l)\left(\left(1 - G(b^{-1}(b_i))\right)\left(\theta_i - b_i\right)\right)\tfrac{g(b^{-1}(b_i))}{b'(b^{-1}(b_i))} + \tfrac{g(b^{-1}(b_i))}{b'(b^{-1}(b_i))}\left(\theta_i - b_i\right) = 0 \end{gathered}$$

Which is equivalent to (noting that $\lambda = \lambda_g + \lambda_l$)

$$G(b^{-1}(b_i)) + G(b^{-1}(b_i))\lambda\left(1 - G(b^{-1}(b_i))\right) = \left(1 + \left[\lambda\left(1 - 2G(b^{-1}(b_i))\right)\right]\right)\frac{g(b^{-1}(b_i))}{b'(b^{-1}(b_i))}\left(\theta_i - b_i\right)$$

In equilibrium, we know that players will play monotonic and symmetric strategies. Therefore, each player's bid is a function of his type: $b_i = b(\theta_i)$. Further, $b^{-1}(b_i) = \theta_i$. Finally, we can drop the subscript i:

$$b'(\theta)G(\theta) + b'(\theta)G(\theta)\lambda\left(1 - G(\theta)\right) = \left(1 + \left[\lambda\left(1 - 2G(\theta)\right)\right]\right)g(\theta)\left(\theta - b(\theta)\right)$$

Subtracting the last term from both sides and rearranging

$$\begin{gathered} b'(\theta)G(\theta) + b(\theta)g(\theta) + b'(\theta)G(\theta)\lambda\left(1 - G(\theta)\right) + b(\theta)\left[\lambda\left(1 - 2G(\theta)\right)\right]g(\theta) \\ = \\ \left(1 + \left[\lambda\left(1 - 2G(\theta)\right)\right]\right)g(\theta)\theta \end{gathered}$$

We can rewrite the top line by noting that it is, in fact, a differential equation

$$\frac{d}{d\theta}\left[G(\theta)b(\theta) + G(\theta)b(\theta)(1 - G(\theta))\lambda\right] = \left(1 + \left[\lambda\left(1 - 2G(\theta)\right)\right]\right) g(\theta)\theta$$

Integrating both sides:

$$\int_{\underline{\theta}}^{\theta} \frac{d}{dx}\left[G(x)b(x) + G(x)b(x)(1 - G(x))\lambda\right] dx = \int_{\underline{\theta}}^{\theta} \left(1 + \left[\lambda\left(1 - 2G(x)\right)\right]\right) g(x)x dx$$

Using the boundary condition given that $G(\underline{\theta}) = 0$,

$$\left[G(\theta)b(\theta) + G(\theta)b(\theta)(1 - G(\theta))\right]\lambda = \int_{\underline{\theta}}^{\theta} \left(1 + \left[\lambda\left(1 - 2G(x)\right)\right]\right) g(x)x dx$$

Finally, we can solve for the bid function

$$b(\theta) = \frac{\int_{\underline{\theta}}^{\theta} \left(1 + \left[\lambda\left(1 - 2G(x)\right)\right]\right) g(x)x dx}{G(\theta)\left[1 + (1 - G(\theta))\lambda\right]}$$

We can rewrite this as the desired function.

$$b^I(\theta) = \frac{\int_{\underline{\theta}}^{\theta} g(x)x dx}{G(\theta)\left[1 + (1 - G(\theta))\lambda\right]} + \frac{\int_{\underline{\theta}}^{\theta} g(x)x\left(\lambda\left(1 - 2G(x)\right)\right) dx}{G(\theta)\left[1 + (1 - G(\theta))\lambda\right]}$$

It remains to show that there is a $\hat{\theta}$ such that for all $\theta \geq \hat{\theta}$, θ will overbid. Since the first term of b^I is always weakly larger than b^{RN} is suffices to show that there is $\hat{\theta}$ such that for all $\theta \geq \hat{\theta}$

$$\frac{\int_{\underline{\theta}}^{\theta} g(x)x\left(\lambda\left(1 - 2G(x)\right)\right) dx}{G(\theta)\left[1 + (1 - G(\theta))\lambda\right]} \geq 0 \tag{A.9}$$

We will show that for $\bar{\theta}$ the above holds strictly: the result then follows from the continuity of the bidding function. At $\bar{\theta}$, the denominator of (A.9) is 1 we need only show the numerator is strictly positive.

Through integration by parts, we obtain the following identity:

$$G^2(x)x|_{\underline{\theta}}^{\bar{\theta}} - \int_{\underline{\theta}}^{\bar{\theta}} G^2(x)dx - \int_{\underline{\theta}}^{\bar{\theta}} xg(x)G(x)dx = \int_{\underline{\theta}}^{\bar{\theta}} xg(x)G(x)dx$$

which allows us to rewrite the numerator of (A.9) as

$$\lambda\Big[\int_{\underline{\theta}}^{\bar{\theta}} g(x)x dx - G^2(x)x|_{\underline{\theta}}^{\bar{\theta}} + \int_{\underline{\theta}}^{\bar{\theta}} G^2(x)dx\Big]$$

Or, denoting μ_G as the expectation according to the distribution G,

$$\lambda\big[\mu_G - \bar{\theta} + \int_{\underline{\theta}}^{\bar{\theta}} G^2(x)dx\big]$$

And so, under the assumption that $\lambda < 0$, it suffices to show that

$$\int_{\underline{\theta}}^{\bar{\theta}} G^2(x)dx < \bar{\theta} - \mu_G \tag{A.10}$$

It is a well known consequence of Tonelli's Theorem that we can write the expectation of a non-negative random variable in terms of its CDF as

$$\mu_G = \int_0^\infty 1 - G(x)dx$$

which, given the support of $F(x)$, can be re-written as

$$\begin{aligned}\mu_G &= \int_0^{\bar{\theta}} 1 - G(x)dx \\ &= \bar{\theta} - \int_{\underline{\theta}}^{\bar{\theta}} G(x)dx\end{aligned}$$

So, since $G(\theta) \leq 1$ for all θ we know

$$\begin{aligned}\int_{\underline{\theta}}^{\bar{\theta}} G^2(x)dx &< \int_{\underline{\theta}}^{\bar{\theta}} G(x)dx \\ &= \bar{\theta} - \mu_G\end{aligned}$$

satisfying (A.10) and thus, proving the claim.

□

A.2 Proofs of Lemmas and Propositions

Proof of Proposition A.1.

Consider any $E, E' \in \mathcal{E}$, and single alignment acts $f \in \mathcal{F}^E$, $g \in \mathcal{F}^{E'}$. Let F be a non-overlapping event for both E and E'. Consider two distinct $h, h' \in \mathcal{F}$, such that $h(s) \sim \bar{h}(s)$ and $h'(s) \sim \bar{h}'(s)$ for all $s \notin F$, which means that their alignment is neutral in F^c. Since the range of U is $(-\infty, \infty)$ by unboundedness, without loss of generality let $U(h'_s) = U(h_s) = 0$ for every $s \notin F$. Moreover suppose that for every $s \in F$, $h(s) \succ \bar{h}(s)$ or $h(s) \prec \bar{h}(s)$, and $h'(s) \succ \bar{h}'(s)$ or $h'(s) \prec \bar{h}'(s)$, that means that on every state in F, h and h' are either strictly considered positive or negative. In addition, let $h \sim h'$. Note that the alignment condition of $s \in F$, and strict monotonicity imply that $U(h(s)) \neq 0$ and $U(h'(s)) \neq 0$ for any $s \in F$, further assume that for all $s \in F$, $h(s) \succ h'(s)$ or $h'(s) \succ h(s)$ (hence on F the acts are always different in terms of preferences).[17]

Local Mixture Consistency guarantees that for f, g, h exists some α_{fgh} such that for any $\alpha \in (\alpha_{fgh}, 1)$, $f \succsim \alpha f + (1-\alpha)h$ if and only if $g \succsim \alpha g + (1-\alpha)h$. Then for any $\alpha \in (\min\{\alpha_{fgh}, \alpha_{fgh'}\}, 1)$

$$\alpha f + (1-\alpha)h \succsim \alpha f + (1-\alpha)h' \iff \alpha g + (1-\alpha)h \succsim \alpha g + (1-\alpha)h' \tag{A.11}$$

Moreover, since f and g are single-aligned continuity of preferences imply that alignment does not change for small perturbations around f and g. Hence, the for α close to one, $\alpha f + (1-\alpha)h \in \mathcal{F}^E$, and $\alpha g + (1-\alpha)h \in \mathcal{F}^{E'}$. From the representation of 2.3, (A.11) implies

$$\begin{aligned}
&\int_S U \circ (\alpha f + (1-\alpha)h)\mu_E(ds) = \int_S U \circ (\alpha f + (1-\alpha)h')\mu_E(ds) \qquad \text{if and only if} \\
&\int_S U \circ (\alpha f + (1-\alpha)h)\mu_{E'}(ds) = \int_S U \circ (\alpha g + (1-\alpha)h')\mu_{E'}(ds)
\end{aligned}$$

By linearity of U (from 2.3) this is equivalent to

$$\begin{aligned}
& & \alpha \int_S (U \circ f)\mu_E(ds) + (1-\alpha)\int_S (U \circ h)\mu_E(ds) \\
& = & \alpha \int_S (U \circ f)\mu_E(ds) + (1-\alpha)\int_S (U \circ h')\mu_E(ds) \\
\iff & & \alpha \int_S (U \circ g)\mu_{E'}(ds) + (1-\alpha)\int_S (U \circ h)\mu_{E'}(ds) \\
& = & \alpha \int_S (U \circ g)\mu_{E'}(ds) + (1-\alpha)\int_S (U \circ h')\mu_{E'}(ds)
\end{aligned} \tag{A.12}$$

Since $U(h_s) = U(h'_s) = 0$ for all $s \notin F$, (A.12) reduces to

$$\sum_F (U \circ h - U \circ h')\mu_E(s) = 0 \iff \sum_F (U \circ h - U \circ h')\mu_{E'}(s) = 0 \tag{A.13}$$

Since S is finite, we can consider $(U \circ h)|F$, the restriction of $U \circ h$ to states in F, as a vector in $\mathbb{R}^{|F|}$.[18] Normalize μ_E and $\mu_{E'}$ conditional on F to be a probability distribution

[17] By Strong Monotonicity and Continuity there is always possible to find such acts h and h'.

[18] More precisely to the $|F|$-dimensional vector subspace of $\mathbb{R}^{|S|}$ where any coordinate in $S \setminus F$ is 0.

over F, then equation (A.13) becomes

$$\sum_F (U \circ h - U \circ h') \underbrace{\frac{\mu_E(ds)}{\mu_E(F)}}_{\mu_E(s|F)} = 0 \iff \sum_F (U \circ h - U \circ h') \underbrace{\frac{\mu_{E'}(ds)}{\mu_{E'}(F)}}_{\mu_{E'}(s|F)} = 0$$

Since all states are non-null, $\mu_E(\cdot|F)$ and $\mu_{E'}(\cdot|F)$ are strictly positive $|F|$-dimensional vectors. These vectors are normal to $(U \circ h - U \circ h') \in \mathbb{R}^{|F|}$, which consists of non-zero elements by the assumption that h and h' are different for all $s \in F$, $(U \circ h - U \circ h')_s \neq 0$ for any $s \in F$. Therefore $\mu_E(\cdot|F)$ and $\mu_{E'}(\cdot|F)$ are colinear as vectors in $\mathbb{R}^{|F|}$, and moreover both are probability distributions which means $\sum_F \mu_E(s|F) = \sum_F \mu_{E'}(s|F) = 1$. Then for all $s \in F$,

$$\frac{\mu_E(s)}{\mu_E(F)} = \frac{\mu_{E'}(s)}{\mu_{E'}(F)} \quad \text{or equivalently} \quad \mu_E(s|F) = \mu_{E'}(s|F) \tag{A.14}$$

For any $E, E' \in \mathcal{E}$, and any F which is non-overlapping with E and E', the conditional distribution $\mu_E(\cdot|F)$ and $\mu_{E'}(\cdot|F)$ which are generated by the set of priors $\{\mu_E\}_{E\in\mathcal{E}}$ from Lemma 2.2. It follows from (A.14) that for any E, E' where $s, s' \in E \cap E'$,

$$\frac{\mu_E(s)}{\mu_E(s')} = \frac{\mu_{E'}(s)}{\mu_{E'}(s')}$$

It remains to show that if (A.14) holds there exists a unique distribution μ that generates the conditional distributions, i.e. such that for all $E \in \mathcal{E}$ and non-overlapping F,

$$\frac{\mu_E(s)}{\mu_E(F)} = \frac{\mu(s)}{\mu(F)} \tag{A.15}$$

Consider the events $E_i = \{s_i, s_{i+1}\}$ for $i = 1, ..., n-1$.[19] If such a μ exists it follows from (A.15) that for E_i,

$$\frac{\mu(s_i)}{\mu(s_{i+1})} = \frac{\mu_{E_i}(s_i)}{\mu_{E_i}(s_{i+1})} \tag{A.16}$$

If such a μ exists and it is unique for the set $\{E_i\}_{i=1,\dots,n-1}$, then by (A.2), it will hold for any $E \in \mathcal{E}$. This is because the set $\{E_i\}_{i=1,..n}$ can be used to find the conditional distributions for any other $E \in \mathcal{E}$ using the results from the first par of the proof. To see this suppose that there is a unique μ such that for all $\{E_i\}_{i=1,\dots,n-1}$, (A.15) holds. Now consider some $E \notin \{E_i\}_{i=1,\dots,n-1}$. $E \subset \bigcup_{i=m}^M E_i := E_m^M$ for some $m < M$, where $m \geq 1$ and $M \leq n$. To see this note that given a numbering of states, there is a maximal s_M, and a minimal s_m states in E, hence $E \subseteq \{s_m, s_{m+1}, ..., s_{M-1}, s_M\} = E_m^M$. Then for any

[19] This is a set that "spans" S, and it is minimal in the sense that there are no overlapping events to E_i which are not singleton states.

$s_t, s_{t+k} \in E$, $t < k$ and $t, k \in \{m, ..., M\}$, $\frac{\mu_E(s_t)}{\mu_E(s_{t+k})} = \frac{\mu_{E_m^M}(s_t)}{\mu_{E_m^M}(s_{t+k})}$ from (A.14). Moreover,

$$\begin{aligned}
\frac{\mu_{E_m^M}(s_t)}{\mu_{E_m^M}(s_{t+k})} &= \left(\frac{\mu_{E_m^M}(s_t)}{\mu_{E_m^M}(s_{t+1})}\right)\left(\frac{\mu_{E_m^M}(s_{t+1})}{\mu_{E_m^M}(s_{t+2})}\right)...\left(\frac{\mu_{E_m^M}(s_{t+k-1})}{\mu_{E_m^M}(s_{t+k})}\right) \\
&= \left(\frac{\mu_{E_t}(s_t)}{\mu_{E_t}(s_{t+1})}\right)\left(\frac{\mu_{E_{t+1}}(s_{t+1})}{\mu_{E_{t+1}}(s_{t+2})}\right)...\left(\frac{\mu_{E_{t+k-1}}(s_{t+k-1})}{\mu_{E_{t+k-1}(s_{t+k})}}\right) \\
&= \left(\frac{\mu(s_t)}{\mu(s_{t+1})}\right)\left(\frac{\mu(s_{t+1})}{\mu(s_{t+2})}\right)...\left(\frac{\mu(s_{t+k-1})}{\mu(s_{t+k})}\right) \\
&= \frac{\mu(s_t)}{\mu(s_{t+k})}
\end{aligned}$$

Hence it suffices to show that for the set $\{E_i\}_{i=1,...,n}$, a unique distribution exists such that (A.15) holds. Note that (A.16) implies that for any $i = 1, 2, ...n-1$,

$$\mu(s_i) = \left(\frac{\mu_{E_i}(s_i)}{\mu_{E_i}(s_{i+1})}\right)\mu(s_{i+1})$$

These $n-1$ equations, along with the necessary condition to be a probability distribution:

$$\sum_{i=1}^{n} \mu(s_1) = 1,$$

gives n equations and n unknowns (the $\mu(s_i)$'s), which can be written in the following form:

$$\underbrace{\begin{bmatrix}
1 & -\frac{\mu_{E_1}(s_1)}{\mu_{E_1}(s_2)} & 0 & 0 & \dots & \dots & 0 \\
0 & 1 & -\frac{\mu_{E_2}(s_2)}{\mu_{E_2}(s_3)} & 0 & \dots & \dots & 0 \\
0 & 0 & 1 & -\frac{\mu_{E_3}(s_3)}{\mu_{E_3}(s_4)} & 0 & \dots & 0 \\
\vdots & \vdots & & \ddots & & & \vdots \\
0 & 0 & \dots & & \dots & 1 & -\frac{\mu_{E_{n-1}}(s_{n-1})}{\mu_{E_{n-1}}(s_n)} \\
1 & 1 & \dots & & & \dots & 1
\end{bmatrix}}_{A_n} \mu = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{A.17}$$

Equation (A.17) has a unique solution if and only if the matrix A_n is invertible. This is proved by induction on $|S|$. Let $|S| = 3$, then need to show that $\det(A_3) \neq 0$.[20] Let $a_{ij} = \frac{\mu_{E_i}(s_i)}{\mu_{E_i}(s_j)} > 0$. $a_{ij} \in (0, \infty)$ because every state is non-null so $\mu_{E_i}(s_j) > 0$ for all i, j.

[20] $|S| = 2$ it is vacuously true, since $E_1 = S$ by definition of E_i, hence $E_1 \notin \mathcal{E}$.

Prove the stronger condition that $\det(A_k) > 0$ instead.

$$\det\left(\begin{bmatrix} 1 & -a_{12} & 0 \\ 0 & 1 & -a_{23} \\ 1 & 1 & 1 \end{bmatrix}\right) = 1\left(1 + a_{23}\right) - \left(-a_{12}\right)\left(a_{23}\right) = 1 + a_{23}(1 + a_{12}) > 0$$

Suppose now that $\det(A_k) > 0$ for all $k < m$. Then

$$A_m = \left[\begin{array}{c|cccccc} 1 & -a_{12} & 0 & 0 & \dots & \dots & 0 \\ \hline 0 & & & & & & \\ 0 & & & & & & \\ \vdots & & & A_{m-1} & & & \\ 0 & & & & & & \\ 1 & & & & & & \end{array}\right]$$

Hence $\det(A_k) = \det(A_{m-1}) + a_{12}\det(B_k)$ where B_k is defined as

$$B_k = \begin{bmatrix} 0 & -a_{23} & 0 & \dots & \dots & & 0 \\ 0 & 1 & -a_{34} & \dots & \dots & & 0 \\ \vdots & \vdots & & & \ddots & & \vdots \\ 0 & 0 & & \dots & \dots & 1 & -a_{n(n-1)} \\ 1 & 1 & & \dots & & \dots & 1 \end{bmatrix} = \left[\begin{array}{c|ccccc} 0 & -a_{23} & 0 & \dots & \dots & 0 \\ \hline 0 & & & & & \\ \vdots & & A_{m-3} & & & \\ 0 & & & & & \\ 1 & & & & & \end{array}\right]$$

Therefore $\det(A_k) = \det(A_{m-1}) + a_{12}(a_{23}\det(A_{m-3})) > 0$, from the induction hypothesis that for all $k < m$, $\det(A_k) > 0$. Hence the system from equation (A.17) has a unique solution, μ. From the previous result, for any $E \in \mathcal{E}$ such that $|E| > 2$, μ_E is also generated by μ. Hence there exists a unique $\mu : 2^S \to [0, 1]$ such that every conditional distribution of μ (conditional on event F), is the same as the conditional distribution of μ_E provided that F and E are non-overlapping.

□

Proof of Lemma A.2.

Suppose there exists $s \in S$ such that $\mu_E(s) = \mu_{E'}(s)$ for some $E \neq E'$. Then by Proposition A.1, there are two cases:

(i.) $s \in E \cap E'$ (or $s \in E^c \cap E'^c$).

(ii.) $s \in E \cap (E')^c$ (or $s \in E' \cap E^c$).

For case (i) by Proposition A.1, for every $s' \in E \cup E'$, $\mu_E(s') = \mu_{E'}(s)$ since $\mu_E(s) = \gamma_E^+ \mu(s)$ and $\mu_E(s) = \gamma_{E'}^+ \mu(s)$. In addition, for any $t \in E \cap (E')^c$, $\mu_E(t) = \mu_{E'}(t)$ by the same argument, which implies that for all $s \in (E')^c$, $\mu_{E'}(t) = \gamma_{E'}^+ \mu(t)$, which can only be true if $\gamma_E^+ = \gamma_{E'}^+ = 1$. Hence $\mu_E = \mu_{E'} = \mu$. For the second case the argument is symmetric (replacing $\gamma^+_{E'}$ for $\gamma^-_{E'}$).

$\square$

Proof of Proposition A.3.

Consider some single alignment $f \in \mathcal{F}^E$, and $g \in \mathcal{F}^F$ such that $f \sim g$, where $\bar{f} \in \mathcal{F}^{E^c}$ and $\bar{g} \in \mathcal{F}^{F^c}$ are the respective balancing acts. Given $s \in S$, consider some $h \in \mathcal{F}$ such that $U(h_t) = 0$ for all $t \neq s$ and $U(h_s) > 0$.

Suppose $\alpha f + (1-\alpha)h \succ \alpha g + (1-\alpha)h$, then by Antisymmetry $\alpha\bar{f} + (1-\alpha)h \prec \alpha\bar{g} + (1-\alpha)h$. From the definition of alignment and continuity of $\succsim$, for α close to 1, then $\alpha f + (1-\alpha)h \in \mathcal{F}^E$ and $\alpha g + (1-\alpha)h \in \mathcal{F}^F$; likewise $\alpha\bar{f} + (1-\alpha)h \in \mathcal{F}^{E^c}$ and $\alpha\bar{g} + (1-\alpha)h \in \mathcal{F}^{F^c}$. Therefore by the representation result from Lemma 2.3, from Antisymmetry $\mu_E(s) > \mu_F(s)$ if and only if $\mu_{F^c}(s) > \mu_{E^c}(s)$.

Suppose that $\mu_E + \mu_{E^c} \neq \mu_F + \mu_{F^c}$. Then for some $s \in S$, $\mu_E(s) + \mu_{E^c}(s) > \mu_F(s) + \mu_{F^c}(s)$, this in turn implies that there exists some $s' \in S$ where $\mu_E(s') + \mu_{E^c}(s') < \mu_F(s') + \mu_{F^c}(s')$ because all μ_E, μ_{E^c}, μ_F, and μ_{F^c} are probability distributions so $\sum_{t\in S}(\mu_E(t) + \mu_{E^c}(t)) = \sum_{s\in S}(\mu_F(t) + \mu_{F^c}(t)) = 2$. Then

$$\begin{aligned} \mu_E(s) - \mu_F(s) &> \mu_{F^c}(s) - \mu_{E^c}(s) \qquad \text{and} \\ \mu_E(s') - \mu_F(s') &< \mu_{F^c}(s') - \mu_{E^c}(s') \end{aligned} \tag{A.18}$$

It must be the case that the following two conditions hold:

$$\begin{aligned} \mu_E(s) - \mu_F(s) &= \theta\left(\mu_{F^c}(s) - \mu_{E^c}(s)\right) \quad \text{for some} \quad \theta < 1 \\ \mu_E(s') - \mu_F(s') &= \theta'\left(\mu_{F^c}(s') - \mu_{E^c}(s')\right) \quad \text{for some} \quad \theta' < 1 \end{aligned} \tag{A.19}$$

For any $v \in \mathbb{R}^{|S|}$ such that $v_t = 0$ for all $t \neq s, s'$, and $v_s, v_{s'} \neq 0$, there exists some $h \in \mathcal{F}$ associated with v, i.e v is the utility vector associated with h_v. For this $h_v \in \mathcal{F}$, $U(h_t) = 0$ for all $t \neq s, s'$ and $U(h_s) = v_s \neq 0$ and $U(h_{s'}) = v_{s'} \neq 0$, this is guaranteed by strong monotonicity and continuity of $\succsim$. By the same argument previously mentioned, for single alignment $f \in \mathcal{F}^E$ and $g \in \mathcal{F}^F$ where $f \sim g$, for α close to 1, then $\alpha f + (1-\alpha)h \in \mathcal{F}^E$ and $\alpha g + (1-\alpha)h \in \mathcal{F}^F$; likewise $\alpha\bar{f} + (1-\alpha)h \in \mathcal{F}^{E^c}$ and $\alpha\bar{g} + (1-\alpha)h \in \mathcal{F}^{F^c}$. Hence for h_v, where $U \circ h_v = v$, representation from proportion 2.3 and Antisymmetry, which requires $\alpha f + (1-\alpha)h_v \succ \alpha g + (1-\alpha)h_v$ if and only if $\alpha\bar{f} + (1-\alpha)h_v \prec \alpha\bar{g} + (1-\alpha)h_v$. This reduces to

$$\begin{aligned} & \mu_E(s)v_s + \mu_E(s')v_{s'} > \mu_F(s)v_s + \mu_F(s')v_{s'} \\ \Leftrightarrow \qquad & \mu_{E^c}(s)v_s + \mu_{E^c}(s')v_{s'} > \mu_{F^c}(s)v_s + \mu_{F^c}(s')v_{s'} \end{aligned}$$

In other words, there is no $v_s, v'_s \in \mathbb{R}$ that solve the system

$$\begin{aligned} \left(\mu_E(s) - \mu_F(s)\right)v_s + \left(\mu_E(s') - \mu_F(s')\right)v_{s'} &> 0 \\ \left(\mu_{E^c}(s) - \mu_{F^c}(s)\right)v_s + \left(\mu_{E^c}(s') - \mu_{F^c}(s')\right)v_{s'} &> 0 \end{aligned}$$

Using the observation (A.19), this is equivalent to

$$\theta\left(\mu_{F^c}(s)-\mu_{E^c}(s)\right)v_s+\theta'\left(\mu_{F^c}(s')-\mu_{E^c}(s')\right)v_{s'}>0$$
$$-\left(\mu_{F^c}(s)-\mu_{E^c}(s)\right)v_s+-\left(\mu_{F^c}(s')-\mu_{E^c}(s')\right)v_{s'}>0$$

$$\begin{bmatrix}\theta\left(\mu_{F^c}(s)-\mu_{E^c}(s)\right) & \theta'\left(\mu_{F^c}(s')-\mu_{E^c}(s')\right)\\ -\left(\mu_{F^c}(s)-\mu_{E^c}(s)\right) & -\left(\mu_{F^c}(s')-\mu_{E^c}(s')\right)\end{bmatrix}\begin{bmatrix}v_s\\ v_{s'}\end{bmatrix}>\begin{bmatrix}0\\0\end{bmatrix} \tag{A.20}$$

Since there is no solution to (A.20), there exists some $p>0$ (Stiemke's Alternative [Stiemke, 1915]) such that

$$\begin{bmatrix}p_1 & p_2\end{bmatrix}\underbrace{\begin{bmatrix}\theta\left(\mu_{F^c}(s)-\mu_{E^c}(s)\right) & \theta'\left(\mu_{F^c}(s')-\mu_{E^c}(s')\right)\\ -\left(\mu_{F^c}(s)-\mu_{E^c}(s)\right) & -\left(\mu_{F^c}(s')-\mu_{E^c}(s')\right)\end{bmatrix}}_{A}=\begin{bmatrix}0\\0\end{bmatrix}$$

Which implies that $\left(\mu_{F^c}(s)-\mu_{E^c}(s)\right)(p_1\theta-p_2)=0$ and $\left(\mu_{F^c}(s')-\mu_{E^c}(s')\right)(p_1\theta'-p_2)=0$, where $p>0$. Since $\left(\mu_{F^c}(s)-\mu_{E^c}(s)\right)\neq 0$ and $\left(\mu_{F^c}(s')-\mu_{E^c}(s')\right)\neq 0$, it must be the case that $(p_1\theta'-p_2)=(p_1\theta-p_2)=0$, which never holds when at least p_1 or p_2 are non-zero, and $\theta>1>\theta'$. A contradiction. Therefore for any $E,F\in\mathcal{E}$,

$$\mu_E+\mu_{E^c}=\mu_F+\mu_{F^c}$$

$\square$

Proof of Proposition A.5.

Recall e_f is defined as the constant such that for a balanced pair, $(f,\bar{f})$, $\frac{1}{2}f(s)+\frac{1}{2}\bar{f}(s)=e_f$ for all $s\in S$. From lemma A.4, the representation from Lemma 2.3, and linearity of U,

$$\begin{aligned}\int_S(U\circ f)\mu(ds)&=\frac{1}{2}\int_S(U\circ f)\mu_E(ds)+\frac{1}{2}\int_S(U\circ f)\mu_{E^c}(ds)\\&=\frac{1}{2}\int_S(U\circ\bar{f})\mu_{E^c}(ds)+\frac{1}{2}\int_S(U\circ f)\mu_{E^c}(ds)\\&=\int_S\left(\frac{1}{2}(U\circ f)+\frac{1}{2}(U\circ\bar{f})\right)\mu_{E^c}(ds)\\&=\int_S(U\circ e_f)\mu_{E^c}(ds)=U\circ e_f\end{aligned}$$

The same holds for $\int_S(U\circ\bar{f})\mu(ds)$, since for all $s,s'\in S$, $U(f(s))+U(\bar{f}(s))=U(f(s'))+U(\bar{f}(s'))=2U(e_f)$, since U represents preferences over constant acts, and the definition of e_f. Hence for all f, e_f is the constant that is equivalent to the expected utility of $f\in\mathcal{F}$ with the prior μ.

$\square$

Proof of Proposition A.6.

Prove by induction on $|E \cap F|$. Let $E \cap F = s$. Then $F \setminus s = E^c$ and $E \setminus s = F^c$. Hence

$$\begin{aligned} \mu_E - \mu_{E\setminus s} &= \mu_E - \mu_F^c \qquad \text{and} \\ \mu_F - \mu_{F\setminus s} &= \mu_F - \mu_E^c \end{aligned}$$

Hence from Proposition A.3, $\mu_E + \mu_{E^c} = \mu_F + \mu_{F^c}$ for all $E, f \in \mathcal{E}$ and the above observation imply $\mu_E - \mu_{E\setminus s} = \mu_F - \mu_{F\setminus s}$.

Suppose (A.1) holds for all E, F with $|E \cap F| = m - 1$ for some $m \in \mathbb{N}$, with $m < n$. Let $|E \cap F| = m$, such that $E \cap F = \{T, s\}$, where $|T| = m - 1$, and $s \notin T$. Consider any $t \in T$, hence $|\{E \setminus s\} \cap \{F \setminus s\}| = |T| = m - 1$ then by the induction hypothesis for any $t \in T$

$$\mu_{E\setminus s} - \mu_{E\setminus\{s,t\}} = \mu_{F\setminus s} - \mu_{F\setminus\{s,t\}} \tag{A.21}$$

Likewise, following the same argument and by the induction hypothesis, numbering $T = \{t_1, ..., t_k\}$, we have that

$$\mu_{E\setminus\{s,t_1,..,t_i\}} - \mu_{E\setminus\{s,...,t_{i+1}\}} = \mu_{F\setminus\{s,t_1,..,t_i\}} - \mu_{F\setminus\{s,...,t_{i+1}\}} \tag{A.22}$$

And finally for $E \setminus \{T, s\}$ and $F \setminus \{T, s\}$, $E \setminus \{T, s\} = F^c$ and $F \setminus \{T, s\} = E^c$ Proposition A.3 implies that

$$\mu_E - \mu_{E\setminus\{T,s\}} = \mu_F - \mu_{F\setminus\{T,s\}} \tag{A.23}$$

Hence, subtracting (A.22) for $i = 2$ from (A.21), yields

$$\mu_{E\setminus s} - \mu_{E\setminus\{s,t_1,t_2\}} = \mu_{F\setminus s} - \mu_{F\setminus\{s,t_1,t_2\}}$$

Continuing this procedure for $i = 3, ..., k$, along with (A.23) yields that (A.1) must hold, hence proving the result.

$\square$

Proof of Proposition A.8.

Consider 3 different cases, (i) $E = F^c$, (ii) $E \cap F \neq \emptyset$ and $E^c \cap F \neq \emptyset$, and $E \cap F^c \neq \emptyset$, and (iii) $F \subset E$. It suffices to consider these three conditions since $E \cap F = \emptyset$, and $E^c \cap F^c \neq \emptyset$, and lemma A.4 will get the result for E and F, from E^c and F^c.

First note that the case where $F = E^c$ the result follows straightforwardly from Proposition A.3. Now consider the cases where $E \cap F \neq \emptyset$ and $E^c \cap F \neq \emptyset$, and $E \cap F^c \neq \emptyset$. From Proposition A.1, $E \in \mathcal{E}$, μ_E is obtained by distorting μ by the same amount on all the positive states (γ_E^+), and by the same amount (γ_E^-) on the negative states. Hence for all $t \in S$,

$$\frac{\mu_E(t)}{\mu(t)} - \frac{\mu_{E\setminus s}(t)}{\mu(t)} = \frac{\mu_F(t)}{\mu(t)} - \frac{\mu_{F\setminus s}(t)}{\mu(t)} \tag{A.24}$$

Since it is possible to divide by $\mu(t) > 0$, and the equality still holds. By the definition of $\gamma_E^+ = \frac{\mu_E(s)}{\mu(s)}$ for $s \in E$, and $\gamma_E^+ = \frac{\mu_E(s)}{\mu(s)}$ for $s \in E^c$. Suppose $s \in E \cap F$, then using (A.24) using the definition of states as positive or negative (when viewed from E or F),

$$\text{for } s = E \cap F: \qquad \gamma_E^+ - \gamma_{E\setminus s}^- = \gamma_F^+ - \gamma_{F\setminus s}^- \tag{A.25a}$$

$$\text{for } t \in E \cap F^c: \qquad \gamma_E^+ - \gamma_{E\setminus s}^+ = \gamma_F^- - \gamma_{F\setminus s}^- \tag{A.25b}$$

$$\text{for } t' \in E^c \cap F: \qquad \gamma_E^- - \gamma_{E\setminus s}^- = \gamma_F^+ - \gamma_{F\setminus s}^+ \tag{A.25c}$$

$$(\text{A.25c}) - (\text{A.25a}): \qquad \gamma_E^- - \gamma_E^+ = \gamma_{F\setminus s}^- - \gamma_{F\setminus s}^+ \tag{A.26a}$$

$$(\text{A.25b}) - (\text{A.25a}): \qquad \gamma_{E\setminus s}^- - \gamma_{E\setminus s}^+ = \gamma_F^- - \gamma_F^+ \tag{A.26b}$$

$$(\text{A.25b}) - (\text{A.25c}): \qquad (\gamma_E^- - \gamma_E^+) - (\gamma_{E\setminus s}^- - \gamma_{E\setminus s}^+) = (\gamma_{F\setminus s}^- - \gamma_{F\setminus s}^+) - (\gamma_F^- - \gamma_F^+) \tag{A.26c}$$

It is easy to see that substituting (A.26a) and (A.26b) into (A.26c) gets $\gamma_E^- - \gamma_E^+ = \gamma_F^- - \gamma_F^+$.

The other case, where $F \subsetneq E$, is simpler. There exist $t \in E^c$, and $s \in E \cap F$. Applying the same procedure as before yields

$$s \in E \cap F: \qquad \gamma_E^+ - \gamma_{E\setminus s}^- = \gamma_F^+ - \gamma_{F\setminus s}^- \tag{A.27a}$$

$$t \in E \cap F^c: \qquad \gamma_E^+ - \gamma_{E\setminus s}^+ = \gamma_F^- - \gamma_{F\setminus s}^- \tag{A.27b}$$

$$\tag{A.27c}$$

Subtracting (A.27b) from (A.27a), yields $\gamma_E^- - \gamma_E^+ = \gamma_F^- - \gamma_F^+$.[21]

$\square$

Proof of Proposition A.9.

By definition $\mu_E(s) = \gamma_E^+\mu(s)$ if $s \in E$ and $\mu_E(s) = \gamma_E^-\mu(s)$ if $s \in E^c$. Since γ_E^+ is the same for all $s \in E$, it follows that for any $E' \subseteq E$, then $\mu_E(E') = \gamma_E^+\mu(E')$ as well. Then

[21] Note that if $|F| = 1$, $\gamma_\emptyset$ would not be defined, but in that case if $|S| \geq 3$, $|F^c| = n - 1$ and the result can follow from reversing the roles of F and E, with E^c and F^c and Proposition A.3.

$\gamma_E^- - \gamma_E^+ = \lambda$ implies that

$$
\begin{aligned}
& \frac{\mu_E(E)}{\mu(E)} - \frac{\mu_E(E^c)}{\mu(E^c)} = \lambda \\
\Rightarrow \quad & \frac{\mu_E(E)}{\mu(E)} - \frac{1-\mu_E(E)}{1-\mu(E)} = \lambda \\
\Rightarrow \quad & \mu_E(E)(1-\mu_E(E)) - (1-\mu_E(E))\mu(E) = \lambda(1-\mu(E))\mu(E)
\end{aligned}
$$

Solving for $\mu_E(E)$ yields $\mu_E(E) = \mu(E)\,(1-\lambda(1-\mu(E)))$, hence

$$
\gamma_E^+ = \frac{\mu_E(E)}{\mu(E)} = (1-\lambda(1-\mu(E))) \tag{A.28}
$$

Likewise we can solve for γ_E^- to get

$$
\gamma_E^- = \frac{\mu_E(E^c)}{\mu(E^c)} = (1+\lambda\mu(E)) \tag{A.29}
$$

The fact that $\frac{\mu_E(E)}{\mu(E)} = \frac{\mu_E(s)}{\mu(s)}$ for all $s \in E$, and $\frac{\mu_E(E^c)}{\mu(E^c)} = \frac{\mu_E(s)}{\mu(s)}$ for all $s \in E^c$ follows from Proposition A.1. Therefore

$$
\begin{aligned}
\mu_E(s) &= \mu(s)\,(1-\lambda(1-\mu(E))) && \text{for } s \in E \\
\mu_E(s) &= \mu(s)\,(1+\lambda\mu(E)) && \text{for } s \in E^c
\end{aligned}
$$

This proves the result.

$\square$