Cost-Benefit Analysis in Reasoning: The Value-of-Information Case with Forward-Looking Agent

Larbi Alaoui
Antonio Penta

July 2021

Barcelona GSE Working Paper Series

Working Paper nº 1266

Cost-Benefit Analysis in Reasoning:

The Value-of-Information Case with Forward-Looking Agent*

Larbi Alaoui† Antonio Penta‡

Universitat Pompeu Fabra Universitat Pompeu Fabra

July 1, 2021

Abstract

We extend the baseline setting in Alaoui and Penta (2018) to provide the representation of the cost-benefit analysis in reaoning of a (non-myopic) forward looking agent who anticipates his future steps of reasoning to take the form of a partitional information structure evaluated in a Bayesian-consistent way.

Keywords: cognition and incentives – choice theory – reasoning – fact-free learning – sequential heuristics – value of information – forward looking – Bayesin consistency

JEL Codes: D01; D03; D80; D83.

1 Forward-Looking Value of Information

We consider three-step processes, with $s \in \{0, 1, 2\}$, where $s = 0$ denotes the initial "no information" stage, $s = 1$ denotes the stage after the first step of reasoning, and $s = 2$ the stage after the second step of reasoning.

We provide a representation theorem corresponding to the standard (fully rational) value of information. Relative to the general model, the main restrictions in the representation will thus entail (i) at each step of reasoning, the agent has probabilistic beliefs $\mu^s \in \Delta(\Omega)$ over the state of the world; (ii) from one stage to the next, the agent actually receives *information*, represented as a partition P^s of the set Ω ; (iii) this informational structure is consistent with the axiom of truth and with perfect recall, and hence the collection of partitions $(P^s)_{s\in\{0,1,2\}}$ represents a *filtration*, with P^2 being finer than P^1 and (w.l.o.g.) we set P^0 equal to the trivial coarsest partition $I_o = \{\Omega\}$; (iv) beliefs from one stage to the next follow the chain rule of Bayesian updating, consistent with the information received; (v) at each stage s, the action chosen if the agent stops reasoning at

*Alaoui gratefully acknowledges financial support from the Spanish Ministry of Science and Innovation under project ECO2014-56154-P. Penta gratefully acknowledges the financial support of the ERC Starting Grant 759424.

†Universitat Pompeu Fabra and Barcelona GSE. E-mail: larbi.alaoui@upf.edu

‡Universitat Pompeu Fabra and Barcelona GSE. E-mail: antonio.penta@upf.edu

s maximizes the expected utility given beliefs μ^s; (vi) the agent has a cost of reasoning c^s at every step: $c(1)$ is the cost of performing the first step (i.e. to acquire the irst piece of information), $c(2)$ is cost of undertaking the second, given the first had been undertaken; (vii) the agent values undertaking the next step of reasoning anticipating the expected change in payoff for the future realization of signals, given the future partition and his current understanding, and taking into account the possibility of further postponing his choice in the future, by undertaking an extra step of reasoning if he finds it optimal.

The crucial innovations relative to the framework in the paper are the *information partition* restriction (point (ii)), and the *non-myopic* attitude towards future understanding (point (vii)). Because of the latter, in particular, the correct value function at each step will be identified by the backwards solution to the problem: the agent correctly anticipates what his optimal choice and expected payoffs would be for each possible piece of information $I^2 \in P^2$. For each $\omega \in \Omega$ and s, we let $I^s(\omega) \in P^s$ denote the cell in the partition P^s which contains ω.

1.1 Representation

Letting $\mu^0 \in \Delta(\Omega)$ denote the prior beliefs at the beginning of the reasoning process, which for simplicity we assume has full support, since $(P^s)_{s\in\{0,1,2\}}$ is a *filtration*, Bayesian updating pins down beliefs at any I^s: $\forall s, \forall I^s \in P^s$, for any event $E \subseteq \Omega$, $\mu(E|I^s) = \mu^0(E)/\mu^0(I^s)$.

Given this, we can define the optimal action at each information set (i.e. what would be the optimal choice if the agent stopped reasoning at the information set):

$$\forall s, \forall I^s \in P^s, \text{ let } a(I^s) := \arg\max_{a\in A} \sum_{\omega\in\Omega} \mu(\omega|I^s)\, u(a,\omega)$$
$$\text{and } u(I^s) = \max \sum_{\omega\in\Omega} \mu(\omega|I^s)\, u(a(I^s),\omega)$$

At $s = 1$, the optimal decision on whether to think further or not depends on the information I^1. Being only a one-step ahead valuation, it takes precisely the form of the myopic value of information representation in the paper:

$$\begin{aligned} U\left(wait|I^1\right) &= \sum_{\omega\in\Omega} \mu^1\left(\omega|I^1\right) u\left(a\left(I^2(\omega)\right),\omega\right) \\ &= \sum_{I^2\in P^2} \mu\left(I^2|I^1\right) u\left(I^2\right) \end{aligned}$$
$$\text{and the marginal } V\left(wait|I^1\right) = U\left(wait|I^1\right) - u\left(I^1\right)$$

and the agent decides to think further at I^1 if and only if[1]

$$U\left(wait|I^1\right) - c(2) \geq u(I_1) \text{ or, equivalently, } V\left(wait|I^1\right) > c(2).$$

At $s = 0$, a forward looking agent anticipates the decision rule above. Hence:

$$U\left(wait|I^0\right) = \sum_{I^1 \in P^1} \mu^0\left(I^1\right)\left[\max\left\{U\left(wait|I^1\right) - c(2); u\left(I^1\right)\right\} - u\left(I^0\right)\right]$$
$$\text{and } V\left(wait|I^0\right) = U\left(wait|I^0\right) - u\left(I^0\right)$$

the agent thus decides to think further at I^0 if and only if

$$U\left(wait|I^0\right) - c(1) \geq u\left(I^0\right) \text{ or, equivalently, } V\left(wait|I^0\right) > c(1).$$

(If the cost $c(2)$ is sufficiently high that the agent does not anticipate thinking further after $s = 1$, then this boils down again to a one-step ahead criterion, but not in general; if instead $c(2)$ is sufficiently low to always think, for instance if $c(2) = 0$, then it is as if the criterion is again myopic w.r.t. to the two steps merged into a single step, with total cost $c(1) + c(2)$).

Theorem 1: Under the properties below, there exists a filtration $(P^x)_{x=1,2}$ (where P^x is a partition of Ω, and P^2 is a refinement of P^1), beliefs $\mu^0 \in \Delta(\Omega)$, conditional beliefs $(\mu^x(\cdot|I_x))_{I_x \in P^x}$ s.t. $\mu^x(\cdot|I_x) \in \Delta(\Omega)$ which are consistent with Bayesian updating (i.e., $\mu^0(\omega) = \mu^x(\omega|I_x) \cdot \mu^0(I_x)$) such that:

$$(u + t, 1) \succsim (u, 0)$$
$$\text{if and only if}$$
$$U\left(wait|I^0\right) - c(1) \geq u\left(I^0\right)$$

where:

1. $u\left(I^0\right) = \max_{a \in A} \sum_{\omega \in \Omega} \mu^0(\omega) u(a, \omega)$
2. $U\left(wait|I^0\right) = \sum_{I^1 \in P^1} \mu^0\left(I^1\right)\left[\max\left\{U\left(wait|I^1\right) - c(2); u\left(I^1\right)\right\} - u\left(I^0\right)\right]$

[1] Note that substituting

$$u\left(I^1\right) = \sum_{\omega \in I^1} \mu\left(\omega|I^1\right) u\left(a\left(I^1\right), \omega\right) \text{ and}$$
$$W\left(wait|I^1\right) = \sum_{\omega \in \Omega} \mu\left(\omega|I^1\right) u\left(a\left(I^2(\omega)\right), \omega\right)$$

the marginal $V\left(wait|I^1\right) = W\left(wait|I^1\right) - u\left(I^1\right)$ becomes

$$V\left(wait|I^1\right) = \sum_{\omega \in \Omega} \mu\left(\omega|I^1\right)\left[u\left(a\left(I^2(\omega)\right), \omega\right) - u\left(a\left(I^1\right), \omega\right)\right]$$

which is exactly as our *one-step ahead value of information* representation.

3. $U\left(wait|I^1\right) = \sum_{I^2 \in P^2} \mu\left(I^2|I^1\right) u\left(I^2\right)$ where

4. $u\left(I^2\right) = \max_{a \in A} \sum_{\omega \in I_2} \mu^2\left(\omega|I_2\right) u\left(a, \omega\right)$

$$a^*\left(\cdot\right) \in \arg \max_{\substack{\sigma:\Omega \to A \\ P_2\text{-measurable}}} \sum_{\omega \in \Omega} \mu^1\left(\omega|I^1\right) u\left(a^*\left(I^2\left(\omega\right)\right), \omega\right)$$

1.2 Model of Reasoning

To allow the attitude towards reasoning of a fully rational forward looking agent, we need to enrich the primitives of the baseline myopic/one-step ahead model considered in the paper. In particular, if at $s = 0$ the agent bases his decision on whether or not to move to $s = 1$ depends on whether he anticipates (possibly conditional on what he may learn at $s = 1$) wanting to further continue his reasoning to $s = 2$, then at mental state $s = 0$ we must elicit not only his attitudes towards reasoning at $s = 1$, but also at $s = 2$.

Hence, the *mental preferences* at $s = 0$ now will be resented by a binary relation $\geqslant_0$ over $U \times \{0, 1, 2\}$. Similar to the paper, we will denote the projections over the three components as $\geqslant_0^0$, $\geqslant_0^1$, and $\geqslant_0^2$, respectively. The *behavioral preferences* $\succeq_0$ are about whether to continue reasoning or not (possibly with incentives), and hence formally they still are a binary relation over a subset of $U \times \{0, 1\}$. As we will see, however, the *reasoning consistency* criterion, connecting behavioral and reasoning preferences will be different, to account for the *forward looking* attitude of the decision maker (assumed instead *myopic* in the paper).

In a more general model with more than two periods, the space of binary relations should be extended accordingly, to encompass as many indices as the agent's foresight accommodates in his decision. This would clearly complicate the analysis without adding much to the concepts, and hence here we only focus on this three stage processes, with two-steps ahead reasoning.

Under the maintained assumption that the reasoning ends at $s = 2$ at the latest, when at state $s = 1$ the agent is just facing a one-step ahead problem. Hence, at this state there is no distinction between myopic and forward looking attitude, and the foresight of the agent reaosning is only one. Hence, at $s = 1$, the model of reasoning in the paper needs no change. All the anlaysis that follows will thus focus on the representation of the agent's decision to reason at $s = 0$, which in this setting is the only state in which the forward-looking criterion we want to illustrate kicks in. Having understood that this entails the expansion of the space of preferences described above, in the following we only list the axioms and properties needed on the $\geqslant_0$ preferences to obtain the forward looking representation.

1.3 Properties at $s = 0$

1.3.1 Old Axioms and Necessary Extensions

The *scope, cost-independence* and *continuity* are extended to hold not only between 0 and 1, but also between 1 and 2:

Property 1 (Scope) *1. If u is constant in a, then $(u, 0) \geqslant_0 (u, 1) \geqslant_0 (u, 2)$.*

2. For $x = 0, 1$: If $u - v$ is constant in a, then $(u, x) \geqslant_0 (u, x + 1)$ if and only if $(v, x) \geqslant_x (v, x + 1)$.

Property 2 (Cost-Independence) *For $x = 0, 1$: For any u, v that are constant in a, and for any $t \in R$, $(u + t, x + 1) \doteq_0 (u, x)$ if and only if $(v + t, x + 1) \doteq_0 (v, x)$*

Property 3 (Continuity) *For $x = 0, 1$: For each $u \in U$, if there exists $t \in R$ s.t. $(u + t, x + 1) \geqslant_0 (u, x)$, then there exists $t^* \leq t$ such that $(u + t^*, x + 1) \doteq_0 (u, x)$.*

The *monotonicity* and *archimedean* axioms, which were assumed for both the 0 and 1-relations, now are also extended to the new component:

Property 4 (Monotonicity) *For $x = 0, 1, 2$: If $u \geq v$, then $u \geqslant_0^x v$. If $u >> v$ then $u \gtrdot_0^x v$.*

Property 5 (Archimedean) *For each $x = 0, 1, 2$ and $v, u, w \in U$ such that $u \gtrdot_0^x v \gtrdot_0^x w$, there exist $0 \leq \beta \leq \alpha \leq 1$ such that $\alpha u + (1 - \alpha) w \gtrdot_0^x v \gtrdot_0^x \beta u + (1 - \beta) w$.*

The *consequentialism* axiom concerned the way the agent evaluates the propect of not reasoning further and choosing right away, and it remains unchanged:

Property 6 (Consequentialism) *For any u: $u \doteq_0^0 u^0$.*

The *Independence* and *no improving replacement axioms,* which were only imposed for the 1-preferences, are now also extended to the 2-preferences. This is obviously because we also want to elicit beliefs (expected to be) held at both possible future states, as well as extend the minimal notion of "optimality":

Property 7 (Independence) *For each $x = 1, 2$: For all $u, v, w \in U$, $u \geqslant_0^x v$ if and only if $\alpha u + (1 - \alpha) w \geqslant_0^x \alpha v + (1 - \alpha) w$ for all $\alpha \in (0, 1)$.*

Property 8 (No Improving Replacement) *For each $x = 1, 2$: For any (u, E) and $s = 0 \in S(u, E)$, $u \geqslant_0^x v$ for all a and $v \in R(u, a)$.*

The *1-0 consistency* axiom remains the same. We will need to also introduce a new consistency between 1 and 2, but this first requires setting up the axioms which ensure the reasoning takes the form of an information filtration. We thus list the 1-0 consistency here as it was, and postpone introducing the new one until further below.

Property 9 (1-0 Consistency) *For any (u, E) and for $s = 0 \in S(u, E)$, $u^s \geqslant_0^1 v^s$ for all $v \in R(u, a^s)$.*

1.3.2 Novel Axioms

The next axiom is not inherently required by a standard value of information representation, but we add it nonetheless merely because it simplifies dealing with Bayesian updating. The reason is that, in the representation, it ensures that the beliefs over future understanding don't attach zero ex-ante probability to any state of the world, and hence all conditional probabilities are well-defined.

Property 10 (Full Support) *For $x = 1, 2$: for any $\hat{\omega}$, $\exists u \in \mathcal{U} : u' \gtrdot_0^x u$ whenever u' is s.t. $u'(a, \omega) = u(a, \omega)$ for all $\omega \neq \hat{\omega}$ and $u'(a, \hat{\omega}) > u(a, \hat{\omega})$.*

The next axiom ensures that, in the representation, the information acquisition is partitional with respect to Ω.

Property 11 (Partitional Reasoning) *Let u,u' be s.t. $\exists! (\hat{a}, \hat{\omega}) : u(\hat{a}, \hat{\omega}) \neq u'(\hat{a}, \hat{\omega})$ and v,v' be s.t. $\exists! (a', \hat{\omega}) : v(a', \hat{\omega}) \neq v'(a', \hat{\omega})$ with $a' \neq \hat{a}$. Then:*

$$u' \gtrdot_0^x u \Rightarrow v' \doteq_0^x v.$$

Based on this axiom, for any $x = 1, 2$ and for any ω, there $\exists! a^x(\omega) \in A : u' \gtrdot_0^x u$ for u,u' such that $u(a', \omega') = u'(a', \omega')$ for all $(a', \omega') \neq (a^x(\omega), \omega)$. We can thus define an equivalence relation $\simeq^x$ s.t.$\omega \simeq^x \omega'$ if and only if $a^x(\omega) = a^x(\omega')$, and let P^x be the partition of Ω induced by such an equivalence relation. Also define $\sigma^x : \Omega \to A$ s.t. $\sigma^x(\omega) = a^x(\omega)$ for every ω. By construction, strategy σ^x is P^x-measurable.

Property 12 (Filtration) *For any ω, ω', $\omega \simeq^2 \omega' \Rightarrow \omega \simeq^1 \omega'$.*

This clearly implies that P_2 is a finer partition than P_1. Together with the above, this further implies that if $\omega \in supp\mu^2(\cdot|I_2)$, then $\omega \in supp\mu^1(\cdot|I_1)$ for some $I_1 \in P_1 : I_2 \subseteq I_1$. Or: $supp\mu^2(\cdot|I_2) \subseteq supp\mu^1(\cdot|I_1)$

For any strategy $\sigma : \Omega \to A$, and for any $u \in U$, let

$$u^\sigma \in U : u^\sigma(a, \omega) = u(\sigma(\omega), \omega)$$

Property 13 (Strong 2-1 Consistency) *For any $u, \hat{u} \in U$ and for any P^1-measurable.$\sigma : \Omega \to A$, $u^\sigma \geqslant_0^2 \hat{u}^\sigma$ if and only if $u^\sigma \geqslant_0^1 \hat{u}^\sigma$.*

Finally, to accommodate a non-myopic agent, who takes into account his (expectations about his) future decision on whether to stop reasoning or not, we need to change the *Reasoning Consistency* condition: in the baseline (myopic) model, the agent decides whether or not to think more about u by comparing $(u, 1)$ and $(u, 0)$. A forward looking agent instead would take into account that, if he were to move to $s = 1$, he wouldn't necessarily make a choice at $s = 1$, but could also decide whether to think further or not. Hence, the decision on whether to choose at $s = 0$ or postpone should be based taking into account that moving to $s = 1$ may mean choosing at $s = 1$ in some states, and continue thinking in others. The decision to reason should thus be based on the *forward looking* transformation of u. We next formalize these ideas:

Reasoning Consistency: $(u + t, 1) \succeq_0 (u, 0)$ if and only if $\left(u^F + t, 1\right) \geqslant_0 (u, 0)$, where u^F is the *forward looking* transformation of u, defined as:

$$u^F(a, \omega) = \begin{cases} u\left(\sigma^2(\omega), \omega\right) - c_2 & \text{if } \omega \in Cont_1 \\ u(a, \omega) & \text{otherwise} \end{cases}$$

The set $Cont_1$ in turn is defined as

$$Cont_1 := \left\{\omega \in \Omega : \left(u_{I_1(\omega)} + c_2 \mathbf{1}_{\Omega \setminus I^1(\omega)}, 2\right) \gtrdot_0 \left(u_{I_1(\omega)}, 1\right)\right\},$$

where, for any $I_1 \in P_1$, u_{I_1} is such that, for some $\bar{a} \in A$,

$$u_{I_1}(a, \omega) = \begin{cases} u(a, \omega') & \text{if } \omega \in I_1 \\ u(\bar{a}, \omega') & \text{if } \omega \notin I^1 \end{cases}$$

Theorem 1 *Under the properties above, there exists a filtration $(P^x)_{x=1,2}$ (where P^x is a partition of Ω, and P^2 is a refinement of P^1), beliefs $\mu^0 \in \Delta(\Omega)$, conditional beliefs $(\mu^x(\cdot|I_x))_{I_x \in P^x}$ s.t. $\mu^x(\cdot|I_x) \in \Delta(\Omega)$ which are consistent with Bayesian updating (i.e., $\mu^0(\omega) = \mu^x(\omega|I_x) \cdot \mu^0(I_x)$) such that:*

$$(u + t, 1) \succsim (u, 0)$$

$$\textit{if and only if}$$

$$U\left(wait|I^0\right) - c(1) \geq u\left(I^0\right)$$

where:

1. $u\left(I^0\right) = \max_{a \in A} \sum_{\omega \in \Omega} \mu^0(\omega) u(a, \omega)$
2. $U\left(wait|I^0\right) = \sum_{I^1 \in P^1} \mu^0\left(I^1\right)\left[\max\left\{U\left(wait|I^1\right) - c(2); u\left(I^1\right)\right\} - u\left(I^0\right)\right]$
3. $U\left(wait|I^1\right) = \sum_{I^2 \in P^2} \mu\left(I^2|I^1\right) u\left(I^2\right)$ *where*
4. $u\left(I^2\right) = \max_{a \in A} \sum_{\omega \in I_2} \mu^2(\omega|I_2) u(a, \omega)$

$$a^*(\cdot) \in \arg \max_{\substack{\sigma:\Omega \to A \\ P_2\text{-measurable}}} \sum_{\omega \in \Omega} \mu^1\left(\omega|I^1\right) u\left(a^*\left(I^2(\omega)\right), \omega\right)$$

Appendix

A Proofs

The same arguments as in the baseline proofs, extended to the 2-preferences above, it follows that Old Axioms, adequately extended to the 2-preferences, imply that:

Lemma 1 *For $x = 0, 1, 2$, there exists $W^x(\cdot; 0) : U \to R$ which represents $\succcurlyeq_0^x$, and takes the following form:*

$$W^x(u; 0) = \sum_{a\in A} p^{0,x}(a) \sum_{\omega\in\Omega} \mu^{a,x}(\omega) \cdot u(a^*(\mu), \omega) \text{ for } x = 1, 2 \tag{1}$$

$$W^0(u; 0) = W^0(u^0) \text{ and } a^0 \in \arg\max_{a\in A} \sum_{\omega\in\Omega} \hat{\mu}^0(\omega) \cdot u(a, \omega) \text{ where } \hat{\mu}^0(\omega) = \sum_{a\in A} p^{0,1}(a) \cdot \mu^{a,1}(\omega) \tag{2}$$

Moreover, the costs $c(1)$ and $c(2)$ are also identified by the standard calibration argument, using the the indifference conditions $\left(u^k + c(1), 1\right) \doteq_0 \left(u^k, 0\right)$ and $\left(u^k + c(2), 2\right) \doteq_0 \left(u^k, 1\right)$, with $c(k) = \infty$ whenever the indifference is not satisfied by any finite t.

In the following, we will assume that mental preferences are such that both $c(1)$ and $c(2)$ are finite, since in the opposite case the forward looking representation would trivially coincide with the myopic one in the paper.

Applying the Partition axiom and Property **??** to (**??**) implies that, respectively

$$supp\mu^{a,x} \cap supp\mu^{\hat{a},x} = \emptyset \text{ whenever } a \neq \hat{a}\text{; and}$$
$$\bigcup_{a\in A} supp\mu^{a,x} = \Omega$$

in that representation. Hence, identifying the cells in the partition P^x with the supports of the $(\mu^{a,x})_{a\in A}$, we can write (**??**) and (**??**) as

$$W^x(u; 0) = \sum_{I_x\in P^x} p^{0,x}(I_x) \sum_{\omega\in I_x} \mu^x(\omega|I_x) \cdot u(\sigma^x(\omega), \omega) \text{ for } x = 1, 2 \text{ and} \tag{3}$$

$$\text{where } \forall I_x \in P_x, \forall\hat{\omega} \in I_x,\ \sigma^x(\hat{\omega}) \in \arg\max_{a\in A} \sum_{\omega\in I_x} \mu^x(\omega|I_x) \cdot u(a, \omega) \tag{4}$$

$$\text{and } a^0 \in \arg\max_{a\in A} \sum_{\omega\in\Omega} \hat{\mu}^0(\omega) \cdot u(a, \omega) \text{ where } \hat{\mu}^0(\omega) = \sum_{I_1\in P^1} p^{0,1}(I_1) \cdot \mu^1(\omega|I_1) \tag{5}$$

Lemma 2 *The probabilities defined above are Bayes-consistent in the sense that, for any ω:*

$$\sum_{I_1\in P_1} p^{0,1}(I_1) \mu^1(\omega|I_1) = \sum_{I_1\in P_1} p^{0,2}(I_1) \sum_{I_2\subseteq I_1} p^{0,2}(I_2|I_1) \cdot \mu^2(\omega|I_2),$$

Proof. *Fix u, $\hat{u}$ and P_1-measurable $\sigma : \Omega \to A$ and such that $u^\sigma \doteq_0^1 \hat{u}^\sigma$ (existence of such functions is ensured by the Archimedean axiom). By the consistency axiom, $u^\sigma \doteq_0^2 \hat{u}^\sigma$, and by the representations above we have*

$$\sum_{I_1 \in P_1} p^{0,1}(I_1)\, \mu^1(\omega|I_1)\, u(\sigma(I_1), \omega) = \sum_{I_1 \in P_1} p^{0,1}(I_1)\, \mu^1(\omega|I_1)\, \hat{u}(\sigma(I_1), \omega)$$

$$\text{and} \sum_{I_2 \in P^2} p^{0,2}(I_2) \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot u(\sigma(I_1), \omega) = \sum_{I_2 \in P^2} p^{0,2}(I_2) \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot \hat{u}(\sigma(I_1), \omega)$$

By contradiction, suppose that the equality above fails for some $\hat{\omega}$, i.e. (wlog)

$$\sum_{I_1 \in P_1} p^{0,1}(I_1)\, \mu^1(\hat{\omega}|I_1) > \sum_{I_1 \in P_1} p^{0,2}(I_1) \sum_{I_2 \subseteq I_1} p^{0,2}(I_2|I_1) \cdot \mu^2(\hat{\omega}|I_2).$$

Then, there exists ω' such that

$$\sum_{I_1 \in P_1} p^{0,1}(I_1)\, \mu^1(\omega'|I_1) < \sum_{I_1 \in P_1} p^{0,2}(I_1) \sum_{I_2 \subseteq I_1} p^{0,2}(I_2|I_1) \cdot \mu^2(\omega'|I_2).$$

Let $\tilde{u}$ be such that

$$\tilde{u}(a, \omega) = \begin{cases} u(a, \hat{\omega}) + \varepsilon & \textit{if } \omega = \hat{\omega} \\ u(a, \omega') - \frac{\varepsilon \cdot \sum_{I_1 \in P_1} p^{0,1}(I_1) \mu^1(\hat{\omega}|I_1)}{\sum_{I_1 \in P_1} p^{0,1}(I_1) \mu^1(\omega'|I_1)} & \textit{if } \omega = \omega' \\ u(a, \omega) & \textit{otherwise} \end{cases}$$

By construction,

$$\sum_{I_1 \in P_1} p^{0,1}(I_1)\, \mu^1(\omega|I_1)\, \tilde{u}(\sigma(I_1), \omega) = \sum_{I_1 \in P_1} p^{0,1}(I_1)\, \mu^1(\omega|I_1)\, \hat{u}(\sigma(I_1), \omega)$$

$$\text{and} \sum_{I_2 \in P^2} p^{0,2}(I_2) \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot u(\sigma(I_1), \omega) > \sum_{I_2 \in P^2} p^{0,2}(I_2) \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot \hat{u}(\sigma(I_1), \omega),$$

and hence $\tilde{u} \doteq_0^1 \hat{u}^\sigma$ and $\tilde{u} \gtrdot_0^2 \hat{u}^\sigma$, contradicting the axiom. The other direction is immediate.
■

In light of the previous axiom, from now on we relabel the $p^{0,1}$ and $p^{0,2}$ probabilities in (**??**)-(**??**) so that

$$W^1(u) = \sum_{I^1 \in P^1} \mu^0(I^1) \sum_{\omega \in I_1} \mu^1(\omega|I_1) \cdot u(\sigma^1(I_1), \omega), \text{ and}$$
$$W^2(u) = \sum_{I_2 \in P_2} \mu^0(I_2) \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot u(\sigma^2(I_2), \omega).$$

Towards the representaiton, it will also be useful to relabel terms in W^2 as follows:

$$
\begin{aligned}
W^2(u;0) &= \sum_{I_2\in P^2} \mu^0(I_2) \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot u\left(\sigma^2(\omega),\omega\right) && (6)\\
&= \sum_{I_1\in P_1} \underbrace{\left(\sum_{I_2\subseteq I_1} \mu^0(I_2)\right)}_{=:p^{0,2}(I_1)} \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot u\left(\sigma^2(\omega),\omega\right) && (7)\\
&:= \sum_{I_1\in P_1} p^{0,2}(I_1) \left(\sum_{I_2\subseteq I_1} \underbrace{\frac{p^{0,2}(I_2)}{p^{0,2}(I_1)}}_{=:p^{0,2}(I_2|I_1)}\right) \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot u\left(\sigma^2(\omega),\omega\right) && (8)\\
&:= \sum_{I_1\in P_1} p^{0,2}(I_1) \underbrace{\sum_{I_2\subseteq I_1} p^{0,2}(I_2|I_1) \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot u\left(\sigma^2(\omega),\omega\right)}_{=:W^2(u|I_1)} && (9)\\
&:= \sum_{I_1\in P_1} p^{0,2}(I_1)\cdot W^2(u|I_1) && (10)\\
&\text{where } \forall I_2\in P_2, \forall\hat{\omega}\in I_2,\ \sigma^2(\hat{\omega})\in \arg\max_{a\in A} \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot u(a,\omega) && (11)
\end{aligned}
$$

Lemma 3 *For any $x=0,1$, and for any $u,v\in U$, $(v,x+1)\geqslant_0 (u,x)$ if and only if $W^{x+1}(v)\geq W^x(u)+c(x+1)$.*

Proof. *We start with the case $x=0$. Let v' be such that $(v',1)\doteq_0 (u,0)$, then:*

$$\left(v',1\right)\doteq_0 (u,0)\doteq_0 \left(u^k,0\right)\doteq_0 \left(u^k+c(1),1\right),$$

where the second indifference follows from the consequentialism axiom, and the third from the definition of $c(1)$. By transitivity, $(v',1)\doteq_0 \left(u^k+c(1),1\right)$, and by the representation $W^1(v')=W^1\left(u^k+c(1)\right)$. Then,

$$
\begin{aligned}
W^1\left(v'\right) &= W^1\left(u^k+c(1)\right)\\
&= W^1\left(u^k\right)+c(1)\\
&= W^0\left(u^k\right)+c(1)\\
&= W^0(u)+c(1)
\end{aligned}
$$

*where the second equality follows from (**??**), the third from 1-0-Consistency (which implies $W^0\left(u^k\right)=W^1\left(u^k\right)$), and the fourth again from (**??**). Finally, $(v,1)\geqslant_0 (u,0)$ if and only if $(v,1)\geqslant_0 (v',1)$, and hence if and only if $W^1(v)\geq W^1(v')=W^0(u)+c(1)$.*

For $x=2$, let $(v,2)\geqslant_0 (u,1)$. By the Archimedean property, there exists u' constant *in a such that* $(u',1) \doteq_0 (u,1)$*, and* v' *(if* $c(2)<\infty$*) such that* $(v',2) \doteq_0 (u',1)$*. Then,*

$$\left(v',2\right) \doteq_0 \left(u',1\right) \doteq_0 \left(u'+c(2),2\right),$$

where the second indifference follows from the definition of $c(2)$*. By transitivity,* $(v',2) \doteq_0 (u'+c(2),2)$*, and by the representation* $W^2(v') = W^2(u'+c(2))$*. Then,*

$$\begin{aligned} W^2\left(v'\right) &= W^2\left(u'+c(2)\right) \\ &= W^2\left(u'\right)+c(2) \\ &= W^1\left(u'\right)+c(2) \\ &= W^1(u)+c(2) \end{aligned}$$

*where the second equality follows from (**??**), the the third from 2-1-Consistency, and the fourth from the representation and* $(u',1) \doteq_0 (u,1)$*. Finally,* $(v,2) \geqslant_0 (u,1)$ *if and only if* $(v,2) \geqslant_0 (v',2)$*, and hence if and only if* $W^2(v) \geq W^2(v') = W^1(u)+c(2)$*.* ■

Lemma 4 $I^1 \subseteq Cont_1$ *if and only if* $\sum_{I^2\in P^2} \mu\left(I^2|I^1\right) u\left(I^2\right) - c(2) > u\left(I^1\right)$

Proof. *By definition,* $I_1 \subseteq Cont_1$ *if and only if* $\left(u_{I_1}+c_2\mathbf{1}_{\Omega\setminus I^1},2\right) \succ_0 \left(u_{I_1(\omega)},1\right)$*, and from the lemma above this holds if and only if* $W^2\left(u_{I_1(\omega)}+c_2\mathbf{1}_{\Omega\setminus I^1(\omega)}\right) > W^1(u_{I_1})+c(2)$*, i.e.*

$$\begin{aligned} W^1(u_{I_1}) &= \sum_{I^1\in P^1} \mu^0\left(I^1\right) \sum_{\omega\in I_1} \mu^1(\omega|I_1)\cdot u_{I_1}\left(\sigma^1(I_1),\omega\right) \\ \text{(by def. of } u_{I_1}) &= \mu^0(I_1) \sum_{\omega\in I_1} \mu^1(\omega|I_1)\cdot u\left(\sigma^1(I_1),\omega\right) \\ &\quad + \sum_{I_1'\in P^1} \mu^0\left(I_1'\right) \sum_{\omega\in I_1'} \mu^1\left(\omega|I_1'\right)\cdot u(\bar a,\omega) \\ \text{(by def. of } u(I_1)\,) &= \mu^0(I_1)\cdot u(I_1) \\ &\quad + \sum_{I_1'\in P^1} \mu^0\left(I_1'\right) \sum_{\omega\in I_1'} \mu^1\left(\omega|I_1'\right)\cdot u(\bar a,\omega) \end{aligned}$$

$$\begin{aligned} W^2\left(u_{I_1(\omega)}+c_2\mathbf{1}_{\Omega\setminus I^1(\omega)}\right) &= \sum_{I_2\in P_2} \mu^0(I_2) \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot u_{I_1}\left(\sigma^2(I_2),\omega\right) \\ \text{(by def. of } u_{I_1}) &= \sum_{I_2\subseteq I_1} \mu^0(I_2) \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot u\left(\sigma^2(I_2),\omega\right) \\ &\quad + \sum_{I_2:I_2\cap I_1=\emptyset} \mu^0(I_2) \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot [u(\bar a,\omega)+c(2)] \\ \text{(by def. of } u(I_2)\,) &= \mu^0(I_1) \sum_{I_2\subseteq I_1} \mu^0(I_2|I_1)\cdot u(I_2) \\ &\quad + \sum_{I_1'\neq I_1} \mu^0(I_1) \sum_{I_2\subseteq I_1'} \mu^0(I_2) \sum_{\omega\in I_2} \mu^2(\omega|I_2)\cdot [u(\bar a,\omega)+c(2)] \end{aligned}$$

Hence, $W^2\left(u_{I_1(\omega)}+c_2\mathbf{1}_{\Omega\setminus I^1(\omega)}\right) > W_0^1\left(u_{I_1}\right)+c\left(2\right)$ *if and only if*

$$\mu^0\left(I_1\right)\sum_{I_2\subseteq I_1}\mu^0\left(I_2|I_1\right)\cdot u\left(I_2\right)+\sum_{I_1'\neq I_1}\mu^0\left(I_1\right)\sum_{I_2\subseteq I_1'}\mu^0\left(I_2|I_1\right)\sum_{\omega\in I_2}\mu^2\left(\omega|I_2\right)\cdot\left[u\left(\bar{a},\omega\right)+c\left(2\right)\right]$$

$$>\mu^0\left(I_1\right)\cdot u\left(I_1\right)+\sum_{I_1'\in P^1}\mu^0\left(I_1'\right)\sum_{\omega\in I_1'}\mu^1\left(\omega|I_1'\right)\cdot u\left(\bar{a},\omega\right)+c\left(2\right)$$

$$\Leftrightarrow\mu^0\left(I_1\right)\sum_{I_2\subseteq I_1}\mu^0\left(I_2|I_1\right)\cdot u\left(I_2\right)+\left(1-\mu^0\left(I_1\right)\right)\cdot c\left(2\right)>\mu^0\left(I_1\right)\cdot u\left(I_1\right)+c\left(2\right)$$

$$\Leftrightarrow\mu^0\left(I_1\right)\sum_{I_2\subseteq I_1}\mu^0\left(I_2|I_1\right)\cdot u\left(I_2\right)>\mu^0\left(I_1\right)\cdot\left[u\left(I_1\right)+c\left(2\right)\right]$$

$$\Leftrightarrow\sum_{I_2\subseteq I_1}\mu^0\left(I_2|I_1\right)\cdot u\left(I_2\right)>u\left(I_1\right)+c\left(2\right).$$

■

Lemma 5 $\left(u^F+t,1\right)\succcurlyeq_0\left(u,0\right)$ *if and only if* Representation

Proof. *By the previous Lemma (part* $x=0$*),* $\left(u^F+t,1\right)\ \cdot\!\succcurlyeq_0\ \left(u,0\right)$ *if and only if*

$W^1\left(u^F\right) \geq W^0\left(u\right) + c\left(1\right)$. *Using (**??**),*

$$
\begin{aligned}
W^1\left(u^F\right) &= \sum_{I^1 \in P^1} \mu^0\left(I^1\right) \sum_{\omega \in I_1} \mu^1\left(\omega | I_1\right) \cdot u^F\left(\sigma^1\left(I_1\right), \omega\right) \\
\text{(by def of } u^F\text{)} &= \left\{ \sum_{\substack{I^1 \in P^1: \\ I^1 \subseteq Cont_1}} \mu^0\left(I^1\right) \sum_{\omega \in I_1} \mu^1\left(\omega | I_1\right) \cdot \left[u\left(\sigma^2\left(\omega\right), \omega\right) - c_2\right] \right\} \\
&+ \left\{ \sum_{\substack{I^1 \in P^1: \\ I_1 \subseteq \Omega \setminus Cont_1}} \mu^0\left(I^1\right) \sum_{\omega \in I_1} \mu^1\left(\omega | I_1\right) \cdot u\left(\sigma^1\left(I_1\right), \omega\right) \right\} \\
&= \left\{ \sum_{\substack{I^1 \in P^1: \\ I^1 \subseteq Cont}} \mu^0\left(I^1\right) \left(\sum_{I^2 \in P^2} \mu\left(I^2 | I^1\right) \sum_{\omega \in I_2} \mu^2\left(\omega | I_2\right) \cdot \left[u\left(\sigma^2\left(\omega\right), \omega\right) - c_2\right] \right) \right\} \\
&+ \left\{ \sum_{\substack{I^1 \in P^1: \\ I_1 \subseteq \Omega \setminus Cont_1}} \mu^0\left(I^1\right) \sum_{\omega \in I_1} \mu^1\left(\omega | I_1\right) \cdot u\left(\sigma^1\left(I_1\right), \omega\right) \right\} \\
\text{(by def of } u\left(I^2\right) \text{ and } u\left(I^1\right)\text{)} &= \left\{ \sum_{\substack{I^1 \in P^1: \\ I^1 \subseteq Cont}} \mu^0\left(I^1\right) \left(\sum_{I^2 \in P^2} \mu\left(I^2 | I^1\right) u\left(I^2\right) - c\left(2\right) \right) \right\} \\
&+ \left\{ \sum_{\substack{I^1 \in P^1: \\ I_1 \subseteq \Omega \setminus Cont_1}} \mu^0\left(I^1\right) u\left(I^1\right) \right\} \\
\text{(by previous Lemma)} &= \sum_{I^1 \in P^1} \mu^0\left(I^1\right) \cdot \max\left\{ \sum_{I^2 \in P^2} \mu\left(I^2 | I^1\right) u\left(I^2\right) - c\left(2\right); u\left(I^1\right) \right\}
\end{aligned}
$$

■

Proof.

$$W^2(u;0) = \sum_{I_2 \in P^2} p^{0,2} \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot u\left(\sigma^2(\omega), \omega\right) \tag{12}$$

$$= \sum_{I_1 \in P_1} \underbrace{\left(\sum_{I_2 \subseteq I_1} p^{0,2}(I_2)\right)}_{=:p^{0,2}(I_1)} \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot u\left(\sigma^2(\omega), \omega\right) \tag{13}$$

$$:= \sum_{I_1 \in P_1} p^{0,2}(I_1) \left(\sum_{I_2 \subseteq I_1} \underbrace{\frac{p^{0,2}(I_2)}{p^{0,2}(I_1)}}_{=:p^{0,2}(I_2|I_1)}\right) \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot u\left(\sigma^2(\omega), \omega\right) \tag{14}$$

$$:= \sum_{I_1 \in P_1} p^{0,2}(I_1) \underbrace{\sum_{I_2 \subseteq I_1} p^{0,2}(I_2|I_1) \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot u\left(\sigma^2(\omega), \omega\right)}_{=:W^2(u|I_1)} \tag{15}$$

$$:= \sum_{I_1 \in P_1} p^{0,2}(I_1) \cdot W^2(u|I_1) \tag{16}$$

$$\textit{where } \forall I_2 \in P_2, \forall \hat{\omega} \in I_2,\ \sigma^2(\hat{\omega}) \in \arg\max_{a \in A} \sum_{\omega \in I_2} \mu^2(\omega|I_2) \cdot u(a, \omega) \tag{17}$$

■