Dynamic Semi-Consistency
Very Preliminary and Incomplete

Sophie Bade*

March 24, 2016

Abstract

The behavior of dynamically consistent agents who follow through with any ex ante optimal plan cannot be distinguished from the behavior of Bayesians. The notion of dynamic semi-consistency allows for models in which different ambiguity attitudes generate different predictions while at the same time keeping the normatively appealing tenet that agents follow through with ex ante optimal pure strategies. Semi-consistent agents do, however, not follow through with all ex ante optimal mixed strategies, as they do not update their preferences upon learning independent randomization outcomes. If players are dynamically semi-consistent, the equilibria of games with different ambiguity attitudes may differ substantially. If we complement the assumption of ambiguity aversion with the assumption of dynamic consistency, the equilibrium sets of games with ambiguity averse agents coincide with the equilibrium sets of Bayesian games. While the standard revelation principle for mechanism design applies in full force if agents are dynamically consistent a modification is required for semi-consistent agents.

Keywords: Ambiguity Aversion, Dynamic Consistency, Games with Incomplete Information. *JEL Classification Numbers*: .

*Royal Holloway College, Universiy of London and Max Planck Institute for Research on Collective Goods, Bonn sophie.bade@rhul.ac.uk

1 Introduction

Do models with different assumptions on ambiguity attitudes generate different predictions? An event is ambiguous for an agent he does not know its probability. To describe the agent's ambiguity attitude fix a set of indifferent bets on such ambiguous events. The agent is ambiguity neutral if he is indifferent between all objective mixtures over the bets in this set. He exhibits ambiguity aversion if he strictly prefers to hedge over some of these bets. Conversely he exhibits a liking of ambiguity if he strictly prefers the original bets to some of mixture over them.

If agents are dynamically consistent in the sense that they follow through with all their ex ante optimal plans then ambiguity attitudes are irrelevant. Fix prediction of a model with dynamically consistent agents, who may show any attitude towards ambiguity. Then this prediction can be replicated by an alternative model with expected utility maximizing agents that is otherwise identical to the original model. All dynamically consistent behavior looks like that of expected utility maximizers.

However, dynamic consistency without ambiguity neutrality only holds if the agent overturns his preferences upon learning some outcomes of the randomization device. So dynamic consistency clashes with the assumption that the randomization device is independent. Typically learning some event E does not affect an agent's preferences over acts that are conditioned on events which are independent on E. In line with this interpretation of the independence of the randomization device, dynamically semi-consistent agents never overturn their original preferences upon learning independent randomization events. Except for this difference dynamic consistency and semi-consistency impose the same requirement that agents follow through with their plans. I show that and how different ambiguity attitudes generate different predictions if agents are semi-consistent.

Agents learn signals and set out plans $\mathfrak{a}$ that specify and action for each signal. Agents may, in addition, use independent and objective randomization devices to decide which plan to adopt. A complete plan (c-plan) maps any outcome of his randomization device to a plan that then maps signals to actions. With these two layers of uncertainty the agent faces two opportunities to deviate from his c-plan. At any randomization outcome the agent has to decide whether to stick with the plan set out by his c-plan for the given randomization outcome. Upon learning his signal the agent then has to decide whether to choose the action his c-plan foresees for the given signal and randomization outcome.

Preferences govern all of the agent's choices. The agent's ex ante preferences guide his choice of a c-plan. Upon learning the randomization outcome E the agent chooses an optimal plan $\mathfrak{a}$ according to this E-conditional preferences. His conditional preferences given E and his signal determine which action to choose.

Different normative principles imply different relations between ex ante and conditional preferences. The principle of dynamic consistency requires that agents follows through with ex ante optimal plans. The agent is dynamically consistent if the prescription of an ex ante optimal c-plan for some event is optimal according to the agents conditional preferences for that event. I impose throughout that agents are dynamically consistent with respect to the signals they learn. However there is a second appealing normative principle for conditional preferences given outcomes of the randomization device: independence. The independence of the randomization device of all payoff relevant events (including the agent's signals) entails a different relation between ex ante and conditional preferences. Learning the outcome of the independent randomization device should not overturn the agent's ex ante preference over any two plans (that may in turn be conditioned on his signals).

If the agent is Bayesian the two principles do not conflict. In fact both generate the same conditional preferences for randomization events: there is no outcome of the (independent) randomization device for which the agent's prior on all payoff relevant events differs from his posterior. If the agent is not a Bayesian the two principles may generate two different families of conditional preferences. To see this consider the Ellsberg paradox. An agent may bet on the color of a ball drawn from an urn that contains 10 red and black balls in unknown proportion. The agent wins 1 if a ball of the named color is drawn, otherwise he gets nothing. The Ellsberg paradox consists in the agent's strict preference for a lottery m according to which he wins $1-\epsilon$

with probability one half and otherwise gets nothing. No expected utility maximizer would make this choice. Raiffa [16] then argues that we should not see anyone make this choice. Instead of choosing m any agent should use a fair coin to decide which of the two bets to play. By doing the agent wins 1 with probability of one half - no matter the proportion of black and red balls in the urn.

In terms of the present terminology Raiffa [16] suggests that the c-plan for which the agent chooses to bet on black (b) if the coin comes up heads (H) and to bet on red (r) if the coin comes of tails (T) is ex ante optimal in the set of all lotteries on m, b, and r. To follow through with his ex ante optimal c-plan, b needs to be ranked (weakly) above m according to the agent's conditional preference given heads. So dynamic consistency requires the agent's ex ante preference of m over b to be overturned upon learning heads. The independence of the coin, in contrast, requires that the agent keeps his ex ante preferences over m, b, and r.

To define the notion of dynamic semi-consistency we need to consider updating with respect to randomization events. To facilitate this analysis the agent's randomization devices have to be modelled via an algebra on the state space. Modelling the agent's choice space as a set of Anscombe-Aumann that use objective randomization devices would impose by fiat that agent's may commit to any randomization. In Theorem **??** I show that the choices of a dynamically semi-consistent agent can be represented as the optimal choice from a subset of all Anscome Aumann acts. This subset is such that the agent only assings positive probability to optimal pure plans. Theorem **??** shows how to represent the choices of dynamically semi-consistent agents via restricted choices sets of Anscombe-Aumann acts.

The second part of the paper shows how the behavior of dynamically semi-consistent agents differs from that of Bayesians. Following Theorem **??** I represent the

Theorems 2 shows that the behavior of dynamically consistent agent cannot be distinguished that of a Bayesian. this observational equivalence does not apply to the case of weakly dynamically consistent agents. and 3 Section 8 embeds Theorems 2 and 3 into a game theoretic context. Games with incomplete information are observationally equivalent to Bayesian games with non-common priors if all players are strongly dynamically consistent (Theorem 4). Conversely, the game theoretic context may amplify the substantial difference between the behavior of weakly dynamically consistent agents and expected utility maximizers. Considering mechanism design I compare the outcomes that are implementable with dynamically consistent agents and with expected utility maximizers. I first show (Theorem 6) that the introduction of ambiguous communication into standard mechanism design problems, keeping all else fixed, does not change the set of implementable social choice functions. The preceding result holds for strong as well as for weakly dynamically consistent behavior. If we additionally modify the standard setup to allow for social choice functions that not only depend on the agent's types but also on the outcomes of ambiguous randomization devices, then ambiguity aversion makes a difference. Theorem 7 shows that ambiguous communication can in this case be used to increase the set of implementable choice functions if agents are weakly dynamically consistent.

The examples that demonstrate an observational difference between weakly dynamically consistent agents that are or are not ambiguity neutral all involve the most well-studied model of preferences in which agents violate ambiguity neutrality. All claimed observational differences hold for the case that the agent has a maxmin expected utility representation following Gilboa and Schmeidler [6].

The results contrast sharply with the existing literature on applied game theoretic models involving ambiguous signals or ambiguous communication which often draws surprising an interesting conclusions from the assumption of dynamic inconsistencies. In the context of mechanism design Bose and Renou [3] show that the use of ambiguous communication devices vastly increases the set of implementable choice functions when the agents responses to such communication are dynamically inconsistent. Ellis [7] shows that dynamically inconsistent choices in Condorcet model with ambiguity averse citizens may overturn the classical result that information aggregates correctly in the limit. Kellner and Le Quemet [10] show that strict Pareto improvements can be obtained in cheap talk games if the receiver's responses to the sender's messages are dynamically inconsistent. In contrast I show that even the assumption of dynamically consistent ambiguity averse agents may lead to novel predictions in games with incomplete information and in mechanism design.

2 Decision Problems

There is a finite set of outcomes X and state space Ω which is endowed with an algebra Σ. The agent's preference, defined over Σ-measurable Anscombe-Aumann acts $F : \Omega \to \Delta X$,[1] can be decomposed into two parts: an expected utility preference over constant acts and a preference $\succsim$ over utility valued acts. The agent weakly prefers the act $F : \Omega \to \Delta X$ to the act $F' : \Omega \to \Delta X$ if $u \circ F \succsim u \circ F'$, where the expected utility $u : \Delta X \to \mathbb{R}$ represents the agent's preferences over constant acts. To save on notation, I directly consider utility valued Σ-measurable acts $f : \Omega \to \mathbb{R}$ with $f : = u \circ F$ for some $F : \Omega \to \Delta X$. The agent's preference $\succsim$ is transitive, complete and monotonic in the sense that $f \gg f'$ implies $f \succ f'$ for any two utility valued acts f, f, where $f \gg f'$ holds if and only if $f(\omega) > f'(\omega)$ for all $\omega \in \Omega$.[2] If $\succsim$ satisfies some further axioms[3], it has an expected utility representation with $U(f) = \int_{\omega\in\Omega} f(\omega)d\pi(\omega)$ where π is a Σ-measurable probability on Ω. In this case the agent is a **Bayesian** with prior π.

The agent has to choose an action a from a finite action set A. The function $G : A \times \Omega \to \Delta X$ maps any action together with a state to a lottery over outcomes. The corresponding utility valued function is $g : = u \circ G : A \times \Omega \to \mathbb{R}$ with $g(a, \omega) = u(G(a, \omega))$ for all $(a, \omega) \in A \times \Omega$. Any action

[1]The set ΔS is the set of all lotteries on the finite set S.

[2]A preference R on acts $F : \Omega \to \Delta X$ can represented via such a $\succsim$ on utility valued acts if R is complete, transitive, monotonic, risk independent, and risk continuous. The preference R is monotonic if $F(\omega)RF'(\omega)$ for all $\omega \in \Omega$ implies FRF' for any $F, F' : \Omega \to \Delta X$. It is risk independent if pRq implies $\alpha p + (1-\alpha)rR\alpha q + (1-\alpha)r$ for all $\alpha \in (0, 1)$ and all constant acts $p, q, r \in \Delta X$. It is risk continuous if for any sequences $(p_n)_{n=1}^{\infty}$ and $(q_n)_{n=1}^{\infty}$ of constant acts in ΔX with $\lim p_n = p$ and $\lim q_n = q$ and p_nRq_n for all n we have pRq. Since R is transitive, complete, risk independent and risk continuous, then the restriction of R to constant acts ΔX has by the von-Neumann-Morgenstern Theorem an expected utility representation $u : \Delta X \to \mathbb{R}$. Since R is monotone there exists a preference $\succsim$ on the set of all utility valued acts such that FRF' holds if and only if $u \circ F \succsim u \circ F'$.

[3]Anscombe and Aumann [] show that the preference has an expected utility representation if the preference is in addition assumed to be continuous and independent.

a induces a $\mathcal{R}$-measurable act $g(a,\cdot):\Omega\to\mathbb{R}$, where $\mathcal{R}\subset\Sigma$ is the finite **partition of payoff relevant events** on Ω. So $g(a,\cdot)$ maps two states ω,ω' in the same event $\sigma\in\mathcal{R}$ to $g(a,\omega)=g(a,\omega'):=g(a,\sigma)$.[4]

The agent may condition his choice of an action a on his information which is described by a finite **information partition** $\mathcal{Q}\subset\Sigma$ on Ω. At state ω the agent learns the signal (type) θ with $\omega\in\theta\in\mathcal{Q}$. Letting his action depend on his information, the agent chooses a $\mathcal{Q}$-measurable plan $\mathfrak{a}:\Omega\to A$. The set of all such (pure) plans is $\mathcal{A}$. Since A and $\mathcal{Q}$ are both finite, the set of plans $\mathcal{A}$ is finite as well. Any plan $\mathfrak{a}\in\mathcal{A}$ induces a $\mathcal{Q}\wedge\mathcal{R}$-measurable act $g(\mathfrak{a}(\cdot),\cdot):\Omega\to\mathbb{R}$ that maps any ω to $g(\mathfrak{a}(\theta),\sigma)$ for $\omega\in\theta\cap\sigma$.

The agent may also condition his choice on the outcome of an objective, rich, and independent **universal randomization device** which is modelled via the algebra $\Sigma^r\subset\Sigma$. This randomization device is objective in the sense that restricted to Σ^r-measurable acts the agent's preference has an expected utility representation with prior π^r. It is rich in the sense that any lottery p on the set of all possible plans $\mathcal{A}$ can be represented via a set of sets $\{E(\mathfrak{a})\mid\mathfrak{a}\in\mathcal{A}\}\subset\Sigma^r$ with $\pi^r(E(\mathfrak{a}))=p(\mathfrak{a})$ for all $\mathfrak{a}\in\mathcal{A}$ and $\{E(\mathfrak{a})\mid E(\mathfrak{a})\neq\emptyset\}$ a partition of Ω.

By choosing a subset of pure plans $\overline{\mathcal{A}}\subset\mathcal{A}$ and a partition $\mathcal{P}^{\mathfrak{p}}:=\{E(\mathfrak{a})\mid\mathfrak{a}\in\overline{\mathcal{A}}\}\subset\Sigma^r$ the agent induces a $\Sigma^r\times\mathcal{Q}$-measurable complete plan (**c-plan**) $\mathfrak{p}:\Omega\to A$ with $\mathfrak{p}(\omega)=\mathfrak{a}(\omega)$ for $\omega\in E(\mathfrak{a})$ for all $\omega\in\Omega$. In this case $\mathcal{P}^{\mathfrak{p}}$ is the **(particular) randomization device** used the generate the c-plan $\mathfrak{p}$, which in turn induces a $\Sigma^r\wedge\mathcal{Q}\wedge\mathcal{R}$-measurable act $g(\mathfrak{p}(\cdot),\cdot):\Omega\to A$.

The universal randomization device Σ^r is **weakly independent** in the sense that for any c-pan $\mathfrak{p}$ the act $g(\mathfrak{p}(\cdot),\cdot):\Omega\to A$ is indifferent to the $\mathcal{Q}\wedge\mathcal{R}$-measurable act $f:\Omega\to\mathbb{R}$ with $f(\omega):=\sum_{E(\mathfrak{a})\in\mathcal{P}^{\mathfrak{p}}}\pi^r(E(\mathfrak{a}))g(\mathfrak{a}(\omega),\omega)$ for

[4]If $F(\omega)=F(\omega')$ holds for all ω,ω' in some event E, then $F(E)$ denotes $F(\omega)$ for $\omega\in E$. Any function that maps the entire domain to the same constant x and any lottery that assigns probability 1 to the outcome x are denoted x. A partition $\mathcal{Q}'$ is finer than another partition $\mathcal{Q}$ if any event $\mathcal{Q}$ is the union of some events in $\mathcal{Q}'$, in this case $\mathcal{Q}$ is coarser than $\mathcal{Q}'$. The meet $\mathcal{Q}\wedge\mathcal{Q}'$ of two partitions $\mathcal{Q}$ and $\mathcal{Q}'$ is the coarsest partition that is finer than both $\mathcal{Q}$ and $\mathcal{Q}'$.

all $\omega \in \Omega$ (where $\mathcal{P}^{\mathfrak{p}}$ is the partition associated with $\mathfrak{p}$).[5] This condition reflects the indifference between a compound act that determines the plan $\mathfrak{a}$ to be followed in the event $E(\mathfrak{a})$ and a compound act which maps any event $\theta \cap \sigma \in \mathcal{Q} \wedge \mathcal{R}$ to the expectation of $g(\mathfrak{a}(\theta), \sigma)$ given that $\mathfrak{a}$ occurs with probability $\pi^r(E(\mathfrak{a}))$. This indifference holds if no event $E(\mathfrak{a})$ is correlated with any signal θ or payoff relevant event σ. Weak independence also entails (a form of) continuity: the agent is indifferent between $g(\mathfrak{p}(\cdot), \cdot)$ and $g(\mathfrak{p}'(\cdot), \cdot)$ if the randomization devices $\mathcal{P}^{\mathfrak{p}}$ and $\mathcal{P}^{\mathfrak{p}'}$ used to generate these acts only differ on events that have zero probability according to π^r.

If the agent is a Bayesian then π^r is the Σ^r-marginal of his prior π and the randomization device is weakly independent if and only if any two events $E \in \Sigma^r$ and $\theta \cap \sigma \in \mathcal{Q} \wedge \mathcal{R}$ are independent according π

To maximally avoid the issue of conditional preferences and plans given null events, I assume that the agent only uses randomization devices $\mathcal{P}^{\mathfrak{p}}$ that contain no null events. So $E(\mathfrak{a}) \in \mathcal{P}^{\mathfrak{p}}$ implies $\pi^r(E(\mathfrak{a})) > 0$; the agent either never adopts $\mathfrak{a}$ or he adopts it with positive probability. The set of all c-plans that satisfy this criterion is $R(\mathcal{A})$. This restriction is without loss of generality in the sense that for any partition $\mathcal{P}^{\mathfrak{p}} : = \{E(\mathfrak{a}) \mid \mathfrak{a} \in \overline{\mathcal{A}}\}$ with $\pi^r(E(\mathfrak{a})) = 0$ for some $\mathfrak{a} \in \overline{\mathcal{A}}$ there exists another partition $\mathcal{P}^{\mathfrak{p}'} : = \{E'(\mathfrak{a}) \mid \mathfrak{a} \in \overline{\mathcal{A}}'\}$ such that $\pi^r(E'(\mathfrak{a})) > 0$ for all $\mathfrak{a} \in \overline{\mathcal{A}}'$ and $g(\mathfrak{p}(\cdot), \cdot) \sim g(\mathfrak{p}'(\cdot), \cdot)$.[6] To further restrict the scope of null events, assume that the agent is not indifferent between two $\mathcal{Q}$-measurable acts $f, f' : \Omega \to \mathbb{R}$ with $f(\theta^*) \neq f'(\theta^*)$ for some $\theta^* \in \mathcal{Q}$ and $f(\theta) = f'(\theta)$ for all other $\theta \in \mathcal{Q}$. So the agent considers it possible that any of the signals $\theta \in \mathcal{Q}$ may arise.

[5]Recall that acts f are derived from more basic acts F that map to lotteries ΔX. The agent is an expected utility maximizer with ΔX. In the present case we can calculate the utility of the compound lottery in which $G(\mathfrak{a}(\cdot), \cdot)$ is played with probability $\pi^r(E(\mathfrak{a}))$ as the expectation of the utility $g(\mathfrak{a}(\cdot), \cdot) = u(G(\mathfrak{a}(\cdot), \cdot))$ given that the probability of each $g(\mathfrak{a}(\cdot), \cdot)$ is $\pi^r(E(\mathfrak{a}))$.

[6]To see this fix some $\mathfrak{a}^*$ for which $\pi^r(E(\mathfrak{a}^*)) > 0$, define $E'(\mathfrak{a}^*)$ as the union of $E(\mathfrak{a}^*)$ and all $E(\mathfrak{a})$ with $\pi^r(E(\mathfrak{a})) = 0$. For all $\mathfrak{a} \neq \mathfrak{a}^*$ with $\pi^r(E(\mathfrak{a})) > 0$ let $E'(\mathfrak{a}) : = E(\mathfrak{a})$. Note that the set of all such $E'(\mathfrak{a})$ partitions Ω. The acts $g(\mathfrak{p}'(\cdot), \cdot)$ and $g(\mathfrak{p}(\cdot), \cdot)$ are respectively indifferent to the $\mathcal{Q} \wedge \mathcal{R}$-measurable plans f and f' with $f(\omega) = \sum_{\mathfrak{a} \in \mathcal{A}} \pi^r(E(\mathfrak{a})) g(\mathfrak{a}(\omega), \omega)$ and $f'(\omega) = \sum_{\mathfrak{a} \in \mathcal{A}} \pi^r(E'(\mathfrak{a})) g(\mathfrak{a}(\omega), \omega)$ for all $\omega \in \Omega$. Since $\pi^r(E(\mathfrak{a})) = \pi^r(E'(\mathfrak{a}))$ holds for each $\mathfrak{a} \in \mathcal{A}$, the acts f and f^* are identical. By transitivity $g(\mathfrak{p}'(\cdot), \cdot)$ and $g(\mathfrak{p}(\cdot), \cdot)$ are indifferent.

Before learning the agent chooses a randomization device $\mathcal{P}^{\mathfrak{p}}$ to generate a c-plan $\mathfrak{p} \in R(\mathcal{A})$. Observing the randomization outcome E the adopts a plan $\mathfrak{a} \in \mathcal{A}$. Upon learning his type θ the agent then chooses an action $a \in A$. The agent is sophisticated in the sense that he correctly predicts his own conditional preferences and choices. Having observed his signal θ and the randomization outcome E the agent chooses an action a that is optimal with respect to his conditional preferences given $\theta \cap E$. Knowing the outcome of the randomization device but before learning θ the agent chooses a plan $\mathfrak{a}$. Being sophisticated the agent will only consider the set of plans $\mathfrak{a}$ that set out actions a that he will indeed adopt upon learning θ. The agent will choose a plan in this set that is optimal according to his E-conditional preferences. Before learning anything the agent chooses an **optimal c-plan**. For $\overline{\mathfrak{p}}$ to be an optimal c-plan it has to be a $\succsim$-best c-plan among all c-plans $\mathfrak{p}$ that set out plans $\mathfrak{p}(E)$ and actions $\mathfrak{p}(E \cap \theta)$ that the agent will indeed adopt upon learning the randomization outcome E and the signal θ.

3 Semi-Consistency

The agent is dynamically consistent if all his conditional preferences prescribe for him to stay with any ex ante optimal c-plan $\mathfrak{p} \in R(\mathcal{A})$. For a dynamically semi-consistent agent this condition only applies to pure plans: all his conditional preferences prescribe for him to stay with an ex ante optimal pure plan $\mathfrak{a} \in \mathcal{A}$. However, a dynamically semi-consistent agent's preferences over pure plans $\mathcal{A}$ do not change with with his observation of any randomization outcome $E \in \Sigma^r$. The present article only refers to one type of consistency: dynamic consistency. I therefore drop the modifier "dynamic" in the sequel.

To explicitly define consistent conditional preferences one has to be careful to leave room for different ambiguity attitudes. Machina [14] and McClennen [15] showed that dynamic consistency implies ambiguity neutrality if conditional preferences are defined for all possible events and if conditional preferences only depend on the conditioning events. There are two paths out of this dilemma. One is to let conditional preferences not only depend on the conditioning event but also on the ex ante plan. The other is to require dynamic consistency only on particular families of events. To allow for both, I only define conditional preferences for the events the agent may learn and let these conditional preferences depend on the agent's ex ante plan. The agent's **conditional preference** given $E \cap \theta$ and given that he set out with the plan $\mathfrak{p}$ is denoted $\succsim^{\mathfrak{p}}_{E\cap\theta}$. Conditional preferences are defined for all events $E \cap \theta$ with $E \in \Sigma^r$ and $\theta \in \mathcal{Q}$.

Fix any ex ante optimal $\overline{\mathfrak{p}} \in R(\mathcal{A})$ together with the corresponding particular randomization device $\mathcal{P}^{\overline{\mathfrak{p}}}$. The conditional preferences of a consistent agent are then such that $\overline{\mathfrak{p}}$ remains optimal upon learning any $E(\overline{\mathfrak{a}}) \in \mathcal{P}^{\overline{\mathfrak{p}}}$:

$$g(\overline{\mathfrak{p}}(\cdot), \cdot) \succsim g(\mathfrak{p}(\cdot), \cdot) \text{ for all } \mathfrak{p} \in R(\mathcal{A}) \text{ and some } \overline{\mathfrak{p}} \in R(\mathcal{A}) \Rightarrow$$
$$g(\overline{\mathfrak{a}}(\cdot), \cdot) \succsim^{\overline{\mathfrak{p}}}_{E(\overline{\mathfrak{a}})} g(\mathfrak{a}(\cdot), \cdot) \text{ for all } E(\overline{\mathfrak{a}}) \in \mathcal{P}^{\overline{\mathfrak{p}}} \text{ and all } \mathfrak{a} \in \mathcal{A}$$

Conversely a dynamically semi-consistent agent does not update his conditional preferences over pure acts $\mathcal{A}$ upon learning any outcome of the randomization device:

$$g(\overline{\mathfrak{a}}(\cdot), \cdot) \succsim g(\mathfrak{a}(\cdot), \cdot) \text{ for some } \overline{\mathfrak{a}} \text{ and all } \mathfrak{a} \in \mathcal{A} \Leftrightarrow$$
$$g(\overline{\mathfrak{a}}(\cdot), \cdot) \succsim^{\mathfrak{p}}_{E} g(\mathfrak{a}(\cdot), \cdot) \text{ for all } E \in \Sigma^r \text{ with } \pi^r(E) > 0 \text{ and all } \mathfrak{p} \in R(\mathcal{A})$$

Consistent and semi-consistent agents both follow through with any optimal plan $\overline{\mathfrak{a}}$ they adopt upon learning the outcome of the randomization device.

$$g(\overline{\mathfrak{a}}(\cdot), \cdot) \succsim^{\overline{\mathfrak{p}}}_{E(\overline{\mathfrak{a}})} g(\mathfrak{a}(\cdot), \cdot) \text{ for some } E(\overline{\mathfrak{a}}) \in \mathcal{P}^{\overline{\mathfrak{p}}} \text{ and all } \mathfrak{a} \in \mathcal{A} \Rightarrow$$
$$g(\overline{\mathfrak{a}}(\theta), \cdot) \succsim^{\overline{\mathfrak{p}}}_{E(\overline{\mathfrak{a}})\cap\theta} g(a, \cdot) \text{ for all } \theta \in \mathcal{Q} \text{ and all } a \in A.$$

This condition states that if $\overline{\mathfrak{a}}$ is optimal according to the agent's conditional preference given the agent's ex ante c-plan $\overline{\mathfrak{a}}$ and the event $E(\overline{\mathfrak{a}})$ that this c-plan prescribes the choice of $\overline{\mathfrak{a}}$, then the agent prefers the action $\overline{\mathfrak{a}}(\theta) \in A$ to all other actions $a \in A$ upon learning θ.

The order of the resolution of uncertainty does not matter: if $\overline{\mathfrak{p}}$ is an optimal c-plan for a semi-consistent agent who learns the outcome of the randomization device and his signal in one order, then $\overline{\mathfrak{p}}$ is an optimal choice for a semi-consistent agent who learns in the other order. To see this consider an optimal c-plan $\overline{\mathfrak{p}}$ for a semi-consistent agent who learns his signal θ before the outcome of the randomization device. The agent must then be willing to follow through with $\overline{\mathfrak{p}}$ and he must be $\succsim$-prefer $g(\overline{\mathfrak{p}}(\cdot),\cdot)$ to all $g(\mathfrak{p}(\cdot),\cdot)$ that are induced by a c-plan $\mathfrak{p}$ he is willing to carry out. Since the agent is consistent with respect to θ he in turn follows through with any c-plan $\mathfrak{p}$ that only prescribes actions $\mathfrak{p}(\omega)$ with $g(\mathfrak{p}(\theta \cap E(\mathfrak{a})),\cdot) \succsim_{\theta}^{\mathfrak{p}} g(a,\cdot)$ for all $a \in A$, $\theta \in \mathcal{Q}$ and $E(\mathfrak{a}) \in \mathcal{P}^{\mathfrak{p}}$. The latter statement holds since any c-plan $\mathfrak{p}$ that foresees for a given θ actions that are optimal according the the agent's conditional preference $\succsim_{\theta}^{\mathfrak{p}}$ will be followed through given that learning $E(\mathfrak{a})$ entails no further updating of the agent's preference. In sum $\mathfrak{p}$ must be $\succsim$-optimal among all acts $\mathfrak{p}$ with $g(\mathfrak{p}(\theta \cap E(\mathfrak{a})),\cdot) \succsim_{\theta}^{\mathfrak{p}} g(a,\cdot)$ for all $a \in A$, $\theta \in \mathcal{Q}$ and $E(\mathfrak{a}) \in \mathcal{P}^{\mathfrak{p}}$.

But the same condition holds for a semi-consistent agent who first learns the outcome of the randomization device and then his signal. We saw above that $g(\mathfrak{p}(\theta \cap E(\mathfrak{a})),\cdot) \succsim_{\theta \cap E(\mathfrak{a})}^{\mathfrak{p}} g(a,\cdot)$ must hold for all $a \in A$, $\theta \in \mathcal{Q}$ and $E(\mathfrak{a}) \in \mathcal{P}^{\mathfrak{p}}$ for any c-plan $\mathfrak{p}$ such an agent chooses. Since $\succsim_{\theta \cap E(\mathfrak{a})}^{\mathfrak{p}}$ and $\succsim_{\theta}^{\mathfrak{p}}$ describe the same preference, the same necessary condition has to hold for $\overline{\mathfrak{p}}$ to be an optimal choice for a semi-consistent agent. For both orders a semi-consistent agent may choose the $\succsim$-best plan that satisfies this condition.

4 A simplified representation

Universal objective randomization devices are typically not modelled via a partition Σ^r but rather as set of lotteries $\Delta\mathcal{A}$ over pure plans $\mathcal{A}$. A lottery q the induces the $\mathcal{Q}\wedge\mathcal{R}$-measurable utility valued act $\sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})g(\mathfrak{a}(\cdot),\cdot) : \Omega \to \mathbb{R}$. A lottery $q \in \Delta\mathcal{A}$ is **equivalent** to a c-plan $\mathfrak{p} \in R(\mathcal{A})$ if $q(\mathfrak{a}) = \pi^r(E(\mathfrak{a}))$ holds for all $E(\mathfrak{a}) \in \mathcal{P}^{\mathfrak{p}}$, the particular randomization device associated with $\mathfrak{p}$. The lottery $\overline{q}$ is ex ante optimal in some set S if $\sum_{\mathfrak{a}\in\mathcal{A}} \overline{q}(\mathfrak{a})g(\mathfrak{a}(\cdot),\cdot) \succsim \sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})g(\mathfrak{a}(\cdot),\cdot)$ holds for all $q \in S$.

Theorem 1 shows that the choice set $R(\mathcal{A})$ can without loss of generality be replaced with $\Delta\mathcal{A}$ if the agent consistent. The definition of $\Delta\mathcal{A}$ also leads to a simpler representation of the choice problem of a semi consistent agent.

A weakly consistent agent chooses the best lottery in $\Delta\mathcal{A}$ that only has ex ante optimal pure plans in its support.

Theorem 1 *a) A c-plan $\overline{\mathfrak{p}} \in R(\mathcal{A})$ is optimal for a consistent agent if and only if it is equivalent to an ex ante optimal lottery $\overline{q}$ in $\Delta\mathcal{A}$.*

b) A plan $\overline{\mathfrak{p}} \in R(\mathcal{A})$ is optimal for a semi-consistent agent if and only if it is equivalent to an ex ante optimal lottery $\overline{q}$ in $\Delta\overline{\mathcal{A}}$, where $\overline{\mathcal{A}} \subset \mathcal{A}$ is the set of all acts $\overline{\mathfrak{a}}$ with $g(\overline{\mathfrak{a}}(\cdot), \cdot) \succsim g(\mathfrak{a}(\cdot), \cdot)$ for all $\mathfrak{a} \in \mathcal{A}$.

Proof First note that the weak independence of Σ^r implies that $g(\mathfrak{p}(\cdot), \cdot)$ is indifferent to $\sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})g(\mathfrak{a}(\cdot), \cdot) : \Omega \to \mathbb{R}$ if $\mathfrak{p}$ is equivalent to q. The richness of Σ^r implies there is an equivalent $\mathfrak{p} \in R(\mathcal{A})$ for every $q \in \Delta\mathcal{A}$.

a) Fix any ex ante optimal plan $\overline{\mathfrak{p}} \in R(\mathcal{A})$, so $g(\overline{\mathfrak{p}}(\cdot), \cdot) \succsim g(\mathfrak{p}(\cdot), \cdot)$ holds for all $\mathfrak{p} \in R(\mathcal{A})$. Since the agent is consistent, we have $g(\overline{\mathfrak{a}}(\cdot), \cdot) \succsim^{\overline{\mathfrak{p}}}_{E(\overline{\mathfrak{a}})} g(\mathfrak{a}(\cdot), \cdot)$ and $g(\overline{\mathfrak{a}}(\theta), \cdot) \succsim^{\overline{\mathfrak{p}}}_{E(\overline{\mathfrak{a}})\cap\theta} g(a, \cdot)$ for all $E(\overline{\mathfrak{a}}) \in \mathcal{P}^{\overline{\mathfrak{a}}}$, all $\mathfrak{a} \in \mathcal{A}$, all $\theta \in \mathcal{Q}$ and all $a \in A$. So the agent is willing to carry out all choices prescribed by $\overline{\mathfrak{p}}$. As a sophisticated agent he knows this, and $\overline{\mathfrak{p}} \in R(\mathcal{A})$ is optimal for the consistent agent if and only if it is ex ante optimal. Now let $\overline{q} \in \Delta\mathcal{A}$ be equivalent to $\overline{\mathfrak{p}}$. To see that $\overline{q}$ is an ex ante optimal lottery suppose there was another $q \in \Delta\mathcal{A}$ that induced a strictly preferred act $\sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})g(\mathfrak{a}(\cdot), \cdot)$. Letting $\mathfrak{p}$ be equivalent to q we then obtain the contradiction

$$g(\mathfrak{p}(\cdot), \cdot) \sim \sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})g(\mathfrak{a}(\cdot), \cdot) \succ \sum_{\mathfrak{a}\in\mathcal{A}} \overline{q}(\mathfrak{a})g(\mathfrak{a}(\cdot), \cdot) \sim g(\overline{\mathfrak{p}}(\cdot), \cdot).$$

b) Let $\overline{\mathfrak{p}} \in R(\mathcal{A})$ be an optimal choice for a semi-consistent agent. Let $\overline{q} \in \Delta\mathcal{A}$ be equivalent to $\overline{\mathfrak{p}}$. Since $\pi^r(E(\overline{\mathfrak{a}})) > 0$ implies that $g(\overline{\mathfrak{a}}(\cdot), \cdot) \succsim g(\mathfrak{a}(\cdot), \cdot)$ for all $\mathfrak{a} \in \mathcal{A}$ and since $q(\mathfrak{a}) = \pi^r(E(\mathfrak{a}))$ holds for all $\mathfrak{a} \in \mathcal{A}$, we have $\overline{q} \in \Delta\overline{\mathcal{A}}$. To see that $\overline{q}$ is ex ante optimal in $\Delta\overline{\mathcal{A}}$ suppose it was not, so suppose that there exists a $q \in \Delta\overline{\mathcal{A}}$ such that $\sum_{\mathfrak{a}\in\overline{\mathcal{A}}} q(\mathfrak{a})g(\mathfrak{a}(\cdot), \cdot)$ is strictly preferred to $\sum_{\mathfrak{a}\in\overline{\mathcal{A}}} \overline{q}(\mathfrak{a})g(\mathfrak{a}(\cdot), \cdot)$. For $\mathfrak{p}$ equivalent to q we obtain

$$g(\mathfrak{p}(\cdot), \cdot) \sim \sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})g(\mathfrak{a}(\cdot), \cdot) \succ$$
$$\sum_{\mathfrak{a}\in\mathcal{A}} \overline{q}(\mathfrak{a})g(\mathfrak{a}(\cdot), \cdot) \sim g(\overline{\mathfrak{p}}(\cdot), \cdot).$$

Since $q \in \Delta\overline{\mathcal{A}}$ $\pi^r(E(\mathfrak{a})) = q(\mathfrak{a}) > 0$ only holds for $\mathfrak{a} \in \overline{\mathcal{A}}$. Therefore $\mathfrak{p}$ is among the possible choices for a semi-consistent agent and we obtain a contradiction. The proof that any ex ante optimal lottery $\overline{q} \in \Delta\overline{\mathcal{A}}$ is equivalent to the optimal choice $\overline{\mathfrak{p}}$ of a semi-consistent agent is very similar and therefore omitted. □

Theorem 1 shows that modelling the agent's choice set as $\Delta\mathcal{A}$ is not without loss of generality. The assumption that the agent may choose any lottery over actions upon learning his type θ hides an assumption of dynamic consistency. To excavate this assumption, I did model the agents choice set as $R(\mathcal{A})$ which describes the universal randomization device as an explicit partition Σ^r. In part b) of Theorem 1 I show that one can model the choice set of a semi-consistent agent as a subset of $\Delta\mathcal{A}$ that depends on the agent's preferences over pure acts $\mathcal{A}$. By the arguments in the preceding section Theorem 1 applies whether the agent learns his signal before or after he learns the outcome of the randomization device.

5 Consistency and Observational Equivalence

Bayesian behavior is indistinguishable from any other type of consistent behavior. In Theorem 2 I fix an arbitrary decision problem and show that any optimal c-plan of a consistent agent is optimal for some Bayesian with the same utility u for constant acts.

Theorem 2 *Let $\overline{q} \in \Delta\mathcal{A}$ be an optimal lottery for a consistent agent, then $\overline{q}$ is optimal for some Baysian with the same utility u over constant acts.*

Proof Since the partitions $\mathcal{Q}$ and $\mathcal{R}$ are both finite we can represent any act $\sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})g(\mathfrak{a}(\cdot),\cdot)$ as the vector $\left(\sum_{\mathfrak{a}\in A} q(\mathfrak{a})g(\mathfrak{a}(\theta),\sigma)\right)_{\theta\in\mathcal{Q},\sigma\in\mathcal{R}}$ in $\mathbb{R}^m$ with $m\colon =\mid \mathcal{Q} \mid \times \mid \mathcal{R} \mid$. The set $S\colon = \{\sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})g(\mathfrak{a}(\cdot),\cdot) : q \in \Delta\mathcal{A}\}$ is convex hull of all vectors $\{g(\mathfrak{a}(\theta),\sigma)\}_{\theta\in\mathcal{Q},\sigma\in\mathcal{R}}$. Therefore S is convex and compact.

Let $\overline{f}\colon = \sum_{\mathfrak{a}\in\mathcal{A}} \overline{q}(\mathfrak{a})g(\mathfrak{a}(\cdot),\cdot)$. The set $\{f \mid f \gg \overline{f}\}$ is convex and has $\overline{f}$ is on its boundary. The monotonicity of $\succsim$ implies that $\{f \mid f \gg \overline{f}\}$ is disjoint from S. Since $\overline{f}$ is also on the boundary of S there exists a separating hyperplane $C\colon = \{f \mid f\pi = \overline{f}\pi^*\}$ with $f'\pi^* \leq \overline{f}\pi^* < f\pi^*$ for all $f' \in S$

and all $f \gg \overline{f}$. Since $C \cap \{f \mid f \gg \overline{f}\} = \emptyset$, $\pi^*(\theta \cap \sigma) \geq 0$ holds for all $\theta \cap \sigma \in \mathcal{Q} \wedge \mathcal{R}$ and $\pi^*(\theta^* \cap \sigma^*) > 0$ for some $\theta^* \cap \sigma^* \in \mathcal{Q} \wedge \mathcal{R}$ and π^* can be normalized to be a $\mathcal{Q} \wedge \mathcal{R}$-measurable probability. A Bayesian with prior π^* assigns the expected utility $f\pi^*$ to the utility valued act $f \in S$ and is maxmizing his expected utility by choosing $\overline{q}$. □

Theorem 2 extends the Raiffa [16] critique of the Ellsberg paradox to the case in which agents may learn some information before taking actions. The separating hyperplane argument is, of course, not new. Most recently Kuzmics [13] extended this argument to a wide range of decision problems involving ambiguity. Kuzmics [13] reviews the long pedigree of this type of argument.

6 Maxmin expected utilities

All of the upcoming constructive results on semi-consistent behavior are couched in the most widely used decision theoretic model of ambiguity averse preferences. A representation U of $\succsim$ with $U(f) = \min_{\pi \in C} \sum_{\omega \in \Omega} f(\omega)\pi(\omega)$ for C some convex and compact set of Σ-measurable priors π on Ω is a maxmin expected utility (**MMEU**) following Gilboa and Schmeidler [6]. Recalling that $\succsim$ is defined over utility valued acts andy MMEU preference is summarized by a set of beliefs C. Preferences $\succsim$ with a MMEU representation are **ambiguity averse** in the sense that $f \succsim f'$ implies $\alpha f + (1-\alpha)f' \succsim f'$ for any two Σ-measurable acts $f, f' : \Omega \to \mathbb{R}$ and all $\alpha \in (0,1)$. A preference $\succsim$ with a MMEU representation is Bayesian if C is a singleton. This case obtains if and only if $\succsim$ is ambiguity neutral which in turns holds if $f \sim f'$ implies $\alpha f + (1-\alpha)f' \sim f'$ for any two Σ-measurable acts $f, f' : \Omega \to \mathbb{R}$ and all $\alpha \in (0,1)$.

The conflict between ambiguity aversion and the existence of consistent conditional preferences that only depend on conditioning events for all possible events, that was discussed in Section 3 is relevant for MMEU preferences: If C represents $\succsim$, and if consistent conditional preferences $\succsim_E$ are defined for all non-null $E \in \Sigma$, then C must be a singleton $\{\pi\}$. Epstein and Schneider [5] mapped a first path out of this dilemma. They showed that MMEU preferences $\succsim$ permit a set of consistent conditional preferences $\{\succsim_E\}_{E\in\mathcal{P}}$ for all E in some partition $\mathcal{P}$ if C is **rectangular** with respect to $\mathcal{P}$. This latter condition in turn holds if C can be represented as

$$C = \{\pi : \pi\mid_{\mathcal{P}} \in C\mid_{\mathcal{P}} \text{ and } \pi(\cdot \mid E) \in C(\cdot \mid E) \text{ for all } E \in \mathcal{P}\}$$

where $C\mid_{\mathcal{P}}$ is the set of all of $\mathcal{P}$-marginals $\pi\mid_{\mathcal{P}}$ of some prior in C ($C\mid_{\mathcal{P}} := \{\pi\mid_{\mathcal{P}} \mid \pi \in C\}$) and a for each $E \in \mathcal{P}$ $C(\cdot \mid E)$ is the set of E-updates $\pi(\cdot \mid E)$ for some $\pi \in C$ ($C(\cdot \mid E) := \{\pi(\cdot \mid E) \mid \pi \in C\}$). In that case **full Bayesian updating** of C yields a family of consistent conditional preferences, where each of these preferences $\succsim_E$ has a maxmin expected utility representation with $C(\cdot \mid E)$ the set of all Bayesian updates of C with respect to E as the set of beliefs.

Hanany and Klibanoff [8] laid out the second path out of the dilemma. Letting conditional preferences $\succsim_E^{\mathfrak{p}}$ not only depend on the conditioning event E but also on the ex ante choice $\mathfrak{p}$, they defined families of consistent conditional maxmin expected utility-preferences $\succsim_E^{\mathfrak{p}}$ for all non-null events $E \in \Sigma$.[7]

The assumption that the agent is not indifferent between two $\mathcal{Q}$-measurable acts $f, f' : \Omega \to \mathbb{R}$ with $f(\theta^*) \neq f'(\theta^*)$ for some $\theta^* \in \mathcal{Q}$ and $f(\theta) = f'(\theta)$ for all other $\theta \in \mathcal{Q}$, made in Section 2, translates to $0 < \pi(\theta)$ for all $\theta \in \mathcal{Q}$ and all $\pi \in C$ the set of beliefs defining the MMEU representation of $\succsim$.

7 Semi-Consistency: Observational Difference

We already know from the discussion of the Raiffa critique in the Introduction, that semi-consistent behavior of ambiguity averse agents differs from consistent behavior. In Theorem 3 I show that the integration of semi-consistent behavior into models where agents may indeed learn something ($\mathcal{Q}$ not a singleton) requires no mathematical acrobatics. Even in simple decision problems with learning the semi-consistent behavior is observationally different from consistent behavior, which is in turn by Theorem **??** indistinguishable from that of Bayesians.

[7] In Hanany and Klibanoff [8] the conditional updates also depend on the agent's ex ante choice set. Since this choice set is fixed here, I omit this dependence here.

A decision problem is **simple** if it has the following three features. The agent has a MMEU preference, ambiguity only enters the decision problem via the agent's signals and his dynamically consistent preferences for all possible signals depend only on his signals. So the decision problem is simple if $\succsim$ is represented by a set of beliefs C. All priors $\pi \in C$ share the same $\mathcal{R}$-marginal $\pi^* \mid_{\mathcal{R}}$. The agent knows the probability distribution $\pi^* \mid_{\mathcal{R}}$ of all payoff relevant events, without any information he is an expected utility maximizer. Finally C is rectangular with respect to $\mathcal{Q}$. The family of consistent conditional preferences $(\succsim_\theta)_{\theta \in \mathcal{Q}}$ can be derived via full Bayesian updating of C.

Theorem 3 *The set of optimal choices of a semi-consistent agent in a simple decision problem may be disjoint from the set of optimal choices of a consistent agent.*

Example 1 proves Theorem 3.

Example 1 Fix a state space Ω^*. Say $\mathcal{R} = \{\lambda, \rho\}$ is the partition of payoff relevant events, $\mathcal{Q} = \{L, M, R\}$ is the information partition and $A\colon = \{l, m, r\}$ the set of actions. Define the set of beliefs $C^*\colon = \{\pi \mid \pi(\lambda \cap L) = .5 - \alpha, \pi(\lambda \cap M) = \alpha, \pi(\rho \cap M) = .3 - \alpha, \pi(\rho \cap R) = .2 + \alpha, \alpha \in [.1, .2]\}$. and note that $\pi(\lambda) = \pi(\rho) = \frac{1}{2}$ and $\pi(M) = .3$ hold for each $\pi \in C^*$. The following three tables represent the set of beliefs C^* as parametrised by $\alpha \in [.1, .2]$ and the agent's utility $g(a, \sigma)$ for all $a \in \{l, m, r\}$ and $\sigma \in \{\lambda, \rho\}$.

	L	M	R
λ	$.5 - \alpha$	α	0
ρ	0	$.3 - \alpha$	$.2 + \alpha$

the set of priors C^*

	l	m	r
λ	21	10	0
ρ	0	10	21

utilities $u(g(a, \omega))$

Note that C^* is rectangular with respect to $\mathcal{Q}$: we can represent C^* as $\{\pi \mid \pi \mid_{\mathcal{Q}} \in C^* \mid_{\mathcal{Q}}$ and $\pi(\cdot \mid \theta) \in C^*(\cdot \mid \theta)$ for all $\theta \in \mathcal{Q}\}$ with $C^* \mid_{\mathcal{Q}} = \{(.5 - \alpha, .3, .2 + \alpha) \mid \alpha \in [.1, .3]\}$ where the three components of any vector $(.5 - \alpha, .3, .2 + \alpha)$ stand for the probabilities of the events L, M, and R, $C^*(\cdot \mid L) = \{1\}$, $C^*(\cdot \mid M) = [\frac{1}{3}, \frac{2}{3}]$, and $C^*(\cdot \mid R) = \{0\}$ as the sets of conditional probabilities of λ given the signal $\theta \in \{L, M, R\}$. A unique plan in $\mathcal{A}$ maximizes $\min_{\pi \in C} \sum_{\theta \in \mathcal{Q}, \sigma \in \mathcal{R}} \pi(\theta \cap \sigma) g(\mathfrak{a}(\theta), \sigma)$:

$$\max_{\mathfrak{a} \in \mathcal{A}} \min_{\pi \in C} \sum_{\theta \in \mathcal{Q}, \sigma \in \mathcal{R}} \pi(\theta \cap \sigma) g(\mathfrak{a}(\theta), \sigma) =$$

$$\max_{\mathfrak{a}(M) \in A} \min_{\alpha \in [.1, .2]} \Big((.5 - \alpha) g(l, \lambda) + \alpha g(\mathfrak{a}(M), \lambda) +$$

$$(.3 - \alpha) g(\mathfrak{a}(M), \rho) + (.2 + \alpha) g(r, \rho) \Big) =$$

$$.7 \times 21 + .3 \Big(\max_{\mathfrak{a}(M) \in A} \min_{\beta \in [\frac{1}{3}, \frac{2}{3}]} (\beta g(\mathfrak{a}(M), \lambda) + (1 - \beta) g(\mathfrak{a}(M), \rho) \Big) =$$

$$.7 \times 21 + .3 \Big(\min_{\beta \in [\frac{1}{3}, \frac{2}{3}]} (\beta g(m, \lambda) + (1 - \beta) g(m, \rho) \Big).$$

The first equality follows from the definition of C^* together with the optimality of l and respectively r in the events L and R (no matter which action is chosen in the event M). The second equality recognizes that the agent obtains utility $21 = g(l, \lambda) = g(r, \rho)$ if either L or R occurs which happens with probability $.7 = .5 - \alpha + .2 + \alpha$. With the complementary probability .3 the agent faces a basic Ellsberg urn type problem, where $\beta = \frac{\alpha}{\alpha + .3 - \alpha}$. In this problem the agent is best off choosing m which yields the same utility in the two payoff relevant events. So $\bar{\mathfrak{a}}$ with $\bar{\mathfrak{a}}(L) = l$, $\bar{\mathfrak{a}}(M) = m$, and $\bar{\mathfrak{a}}(R) = r$ is the agent's unique most preferred pure plan and we have $\overline{\mathcal{A}} = \{\bar{\mathfrak{a}}\}$. The lottery that assigns probability 1 to $\bar{\mathfrak{a}}$ is consequently the unique optimal lottery in $\Delta\overline{\mathcal{A}}$. In sum $\bar{\mathfrak{a}}$ is the unique optimal choice of a semi-consistent agent.

However, for any Bayesian with the same utility over constant acts ΔX either $\mathfrak{a}'$ or $\mathfrak{a}''$ with $\mathfrak{a}'(M) = l$, $\mathfrak{a}''(M) = r$, and $\mathfrak{a}'(\theta) = \mathfrak{a}''(\theta) = \bar{\mathfrak{a}}(\theta)$ for $\theta \in \{L, R\}$ is strictly preferred to $\bar{\mathfrak{a}}$.

8 Games with incomplete information

There is a set of n players N. As above there is a state space Ω endowed with an algebra Σ and a set outcomes is X. All players preferences are defined over Σ-measurable act $F : \Omega \to \mathbb{R}$. Each player has an expected utility

$u_i : \Delta X \to \mathbb{R}$ for constant acts and complete, transitive, and monotonic preference $\succsim_i$ over utility valued acts $u_i \circ F : \Omega \to \mathbb{R}$. Since different players i and j may derive different utilities $u_i(p)$ and $u_j(p)$ from the same lottery $p \in \Delta X$ over outcomes, I do not directly consider utility valued acts in this section. Instead each agent's preference over acts $F : \Omega \to \Delta X$ is represented as a tuple $(u_i, \succsim_i)$.

The action set of player i is A_i. The function $G : A \times \Omega \to \Delta X$ maps action profiles $a \in A := A_1 \times \cdots \times A_n$ and states to lotteries over outcomes. Each action profile a induces a $\mathcal{R}$-measurable act $G(a, \cdot) : \Omega \to \Delta X$, where $\mathcal{R} \subset \Sigma$ is the finite partition of payoff relevant events. Player i may condition his choice of an action $a_i \in A_i$ on his signal which is modelled via player i's information partition $\mathcal{Q}_i$. The meet of all players' information partitions is $\mathcal{Q} := \mathcal{Q}_1 \wedge \cdots \wedge \mathcal{Q}_n$. A pure strategy $\mathfrak{a}_i : \Omega \to A_i$ for player i is a $\mathcal{Q}_i$-measurable function that determines an action a_i for each possible type of player i. The set of player i's pure strategies is $\mathcal{A}_i$. The set of all pure strategy profiles $\mathfrak{a}$ is $\mathcal{A}$. The players may also condition their choices on the outcomes of objective, rich and independent universal randomization devices. Following Theorem 1 player i's set of mixed strategies is represented as $\Delta\mathcal{A}_i$. A profile of mixed strategies $q = (q_1, \ldots, q_n)$ induces the act $\sum_{\mathfrak{a} \in \mathcal{A}} q(\mathfrak{a}) G(\mathfrak{a}(\cdot), \cdot)$ where $q(\mathfrak{a}) = q_1(\mathfrak{a}_1) \times \ldots q_n(\mathfrak{a}_n)$ for all $\mathfrak{a} \in \mathcal{A}$.[8]

To summarize a **game** $\Gamma := \big((\mathcal{Q}_i, A_i, u_i, \succsim_i)_{i \in N}, G\big)$ consists of a set of players N, an information partition $\mathcal{Q}_i$, and a set of actions A_i for each player $i \in N$. Player i's preferences $(u_i, \succsim_i)$ are defined over acts $F : \Omega \to \Delta X$. The function $G : A \times \Omega \to \Delta X$ maps action profiles $a \in A$ and states $\omega \in \Omega$

[8] To extend Theorem 1 to games we need to additionally assume that the player's universal randomization devices are independent of each other. Say the algebra $\Sigma_i^r \subset \Sigma$ models player i's rich, weakly independent, and objective universal randomization device and say agents may choose $\Sigma_i^r \wedge \mathcal{Q}_i$-measurable acts $\mathfrak{p}_i : \Omega \to A_i$ associated with the particular randomization devices $\mathcal{P}^{\mathfrak{p}_i} \subset \Sigma_i^r$. A profile of such strategies $\mathfrak{p} := (\mathfrak{p}_1, \ldots, \mathfrak{p}_n)$ induces the act $G(\mathfrak{p}(\cdot), \cdot) : \Omega \to \Delta X$. If we complement the assumptions made in Theorem 1 by the assumption that the agent's randomization devices are independent of each other, this act is indifferent to the act $\sum_{\mathfrak{a} \in \mathcal{A}} q(\mathfrak{a}) G(\mathfrak{a}(\cdot), \cdot)$ with q_i being equivalent to $\mathcal{P}^{\mathfrak{p}_i}$ for each i. Given that all players' randomization devices are objective, all players share the same prior π^r on all randomization device events. The players' randomization devices are the independent of each other if $\pi^r(E_1 \cap \cdots \cap E_n) = \pi^r(E_1) \times \cdots \times \pi^r(E_n)$ holds for any collection of events $E_i \in \Sigma_i^r$.

to lotteries over outcomes ΔX.

A strategy profile q is a **consistent equilibrium** if for each agent i q_i is consistent optimal choice given q_{-i}, so

$$\sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})u_i(G(\mathfrak{a}(\cdot),\cdot)) \succsim_i \sum_{\mathfrak{a}\in\mathcal{A}} q_i'(\mathfrak{a}_i)q_{-i}(\mathfrak{a}_{-i})u_i(G(\mathfrak{a}(\cdot),\cdot))$$

for all $q_i \in \Delta A_i$ and all $i \in N$. The notions of games Γ and of consistent equilibrium proposed here are identical to the notions proposed by Azrieli and Teper [2]. The existence result of Azrieli and Teper [2] therefore applies to consistent equilibria: If each preference $\succsim_i$ over $\mathcal{Q}\wedge\mathcal{R}$-measurable utility valued acts is continuous and ambiguity averse then Γ has a consistent equilibrium.

Similarly the profile q is a **semi-consistent equilibrium** if for each agent i q_i is a semi-consistent optimal choice given q_{-i}. For any q^*_{-i} define $\overline{\mathcal{A}}_i \subset \mathcal{A}_i$ as the set of optimal pure best replies for agent i, so $\overline{\mathfrak{a}}_i \in \overline{\mathcal{A}}_i$ if and only if

$$\sum_{\mathfrak{a}_{-i}\in\mathcal{A}_{-i}} q_{-i}(\mathfrak{a}_{-i})u_i(G((\overline{\mathfrak{a}}_i,\mathfrak{a}_{-i})(\cdot),\cdot)) \succsim_i \sum_{\mathfrak{a}_{-i}\in\mathcal{A}_{-i}} q_{-i}(\mathfrak{a}_{-i})u_i(G((\mathfrak{a}_i,\mathfrak{a}_{-i})(\cdot),\cdot))$$

holds for all $\mathfrak{a}_i \in \mathcal{A}_i$. Then q_i is a semi-consistent optimal plan given q_{-i} if $q_i \in \Delta\overline{\mathcal{A}}_i$ and if

$$\sum_{\mathfrak{a}\in\mathcal{A}} q(\mathfrak{a})u_i(G(\mathfrak{a}(\cdot),\cdot)) \succsim_i \sum_{\mathfrak{a}\in\mathcal{A}} q_i'(\mathfrak{a}_i)q_{-i}(\mathfrak{a}_{-i})u_i(G(\mathfrak{a}(\cdot),\cdot))$$

holds for all $q_i' \in \Delta\overline{\mathcal{A}}_i$.

If all players are Bayesians then $\Gamma = \big((\mathcal{Q}_i, A_i, u_i, \succsim_i)_{i\in N}, G\big)$ is a **Bayesian game**. Taking into account the agent's priors π_i a Bayesian game Γ is denoted $\big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in N}, G\big)$. If in addition all players have the same prior $\pi_i = \pi$ for all i and some π, then Γ is a **Bayesian game with a common prior**. Since a mixed strategy q_i is best reply for a Bayesian if and only if each pure strategy in its support is a best reply, the sets sets of consistent and semi-consistent equilibria in any Bayesian game coincide. Any such equilibrium is a **Bayes Nash equilibrium** - whether the players share a common prior or not.

Theorem 4 *Fix a game* $\Gamma = \big((\mathcal{Q}_i, A_i, u_i, \succsim_i)_{i\in N}, G\big)$ *If* q *is a consistent equilibrium in* Γ*, then* q *is a Bayes Nash equilibrium in some* $\big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in N}, G\big)$*. Let* q, q^* *be two consistent equilibria, then there may be no* $\big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in N}, G\big)$ *Bayesian game in which both* q *and* q^* *are equilibria.*

Proof Fix a consistent equilibrium q and an agent $i^* \in N$. Consider agent the decision problem of agent i^* given by his action set A_{i^*}, his information partition $\mathcal{Q}_{i^*}$ his preferences $\succsim_{i^*}$ over utility valued acts and the function $g : A_i \times \Omega \to \mathbb{R}$ and $g(a_i, \omega){:}\; = \sum_{\mathfrak{a}_{-i}\in\mathcal{A}_{-i}} u_i(G((a_i, \mathfrak{a}_{-i}(\omega)), \omega))$ for all $\omega \in \Omega$. Since q is a consistent equilibrium q_i is an optimal choice for a consistent agent in the problem P. By Theorem 2 there exists a Σ-measurable prior π_{i^*} on Ω such that $q_{i^*} \in \Delta A_{i^*}$ is an optimal choice for i^*. Since agent i^* was chosen arbitrarily we can construct such a prior π_i for every agent i and q is a Bayes Nash equilibrium of the game $\big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in N}, G\big)$.

Now consider the game $\Gamma = \big((\mathcal{Q}_i, A_i, u_i, C)_{i\in N}, G\big)$ with $N = \{1, 2\}$, $A_1 = \{t, b\}$, $A_2 = \{l, r\}$, $\mathcal{Q}_1 = \mathcal{Q}_2$ the trivial partition. There are two payoff relevant λ and ρ. Players are maxmin expected utility maximizers. Both hold the same set $[.1, .9]$ of beliefs that that λ occurs. The payoffs $u_i(G(a, \sigma))$ are given as follows:

	l	r
t	$2,2$	$0,5$
b	$5,0$	$2,2$

states $\omega \in \lambda$

	l	r
t	$2,2$	$5,0$
b	$0,5$	$2,2$

states in $\omega \in \rho$

To see that (t, l) and (b, r) are both consistent equilibria of Γ, note that either agent obtains a utility of 2 in either one of these profiles. If agent 1 deviates from t to q_1 with $q_1(t) < 1$ while agent 2 plays l agent 1 obtains the utility $\min_{\pi(\lambda)\in[.1,.9]} q_1(t)2 + (1 - q_1(t))(\pi(\lambda) \times 5 + (1 - \pi(\lambda)) \times 0) < 2$. By symmetric arguments neither player has a profitable deviation in neither one of the two profiles. Now suppose (t, l) and (b, r) were also equilibria in some Bayesian game $\big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in N}, G\big)$ (which may not have a common prior). For (t, l) to be an equilibrium $\pi_1(\lambda)$ cannot exceed $\frac{2}{5}$ But for (b, r) to be an equilibrium $\pi_1(\lambda)$ must be at least $\frac{3}{5}$, a contradiction. □

While Theorem 2 shows that any consistent equilibrium q of some game is observationally equivalent to an equilibrium of a Bayesian game (with non-common priors) that is otherwise identical to the original game, the stronger result that q is the Bayes Nash equilibrium of a Bayesian game with a common prior does not hold. In fact this stronger result would imply that any Bayes Nash equilibrium of a game with non-common priors is the equilibrium of a Bayesian game with a common prior. Aumann's [1] Example 2.3 shows that this implication does not hold: Consider the game $\Gamma = \big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in\{1,2\}}, G\big)$ where $\mathcal{R} = \{\rho, \lambda\}$, $N = \{1, 2\}$, neither player obtains any information ($\mathcal{Q}_1 = \mathcal{Q}_2 = \{\Omega\}$), $A_1 = \{t, b\}$, and $A_2 = \{l, r\}$, agents 1 and 2 respectively assign probabilities $\frac{1}{4}$ and $\frac{3}{4}$ to state λ ($\pi_1(\lambda) = \frac{1}{4}$ and $\pi_2(\lambda) = \frac{3}{4}$) and payoffs $u_i(G(a, \omega))$ are given by the following table.

	l	r
t	$0, 8$	$3, 3$
b	$1, 1$	$0, 0$

states $\omega \in \lambda$

	l	r
t	$0, 0$	$3, 3$
b	$1, 1$	$8, 0$

states in $\omega \in \rho$

Aumann [1] shows that while (t, r) is an equilibrium in G there exists no game $\big((\mathcal{Q}_i, A_i, u_i, \pi)_{i\in\{1,2\}}, G\big)$ that is identical to Γ, except for the common prior π in which (t, r) is an equilibrium.[9]

Since the single agent decision problems studied in Section 5 can be interpreted as games with only one agent $| N |= 1$, Theorem 3 immediately yields the following corollary pertaining to the observational difference between semi-consistent equilibria and Bayes Nash equilibria.

Corollary 1 *Fix a game* $\big((\mathcal{Q}_i, A_i, u_i, \succsim_i)_{i\in N}, G\big)$ *together with a semi-consistent equilibrium* q. *Then* q *need not be a Bayes Nash equilibrium in any game* $\big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in N}, G\big)$.

To see how a strategic context may amplify the observational difference between Bayesians and semi-consistent agents consider the following Example 2 which incorporates the decision problem defined in Example 1 into a game with more than 1 player.

[9] Interestingly there are games $\big((\mathcal{Q}_i, A_i, u_i, C)_{i\in\{1,2\}}, G\big)$ in which all are maxmin expected utility maximizers with the same set of beliefs C that have (t, r) as a (strongly and weakly consistent) equilibrium: $C := [\frac{1}{4}, \frac{3}{4}]$ is an example.

Example 2 Building on Example 1 define a game $\Gamma = \big((\mathcal{Q}_i, A_i, u_i, C^*)_{i\in N}, G\big)$ with $N = \{1,2\}$, $\mathcal{R} = \mathcal{Q}_1 := \{\lambda, \rho\}$, and $\mathcal{Q}_2 := \{L, M, R\}$. Both agents hold the set of beliefs C^*. Let $A_1 = \{t, b, Safe\}$ and $A_2 = \{l, m, r\}$. The player's utilities $u_i(G(a,\lambda))$ and $u_i(G(a,\rho))$ for all $a \in A$ and the two payoff relevant events λ and ρ are as follows.

	l	m	r
t	$1,21$	$10,10$	$1,0$
b	$0,0$	$10,10$	$0,21$
$Safe$	$2,0$	$2,0$	$2,0$

states $\omega \in \lambda$

	l	m	r
t	$0,21$	$10,10$	$0,0$
b	$1,0$	$10,10$	$1,21$
$Safe$	$2,0$	$2,0$	$2,0$

states $\omega \in \rho$

The pure strategy profile $\mathfrak{a}$ with $\mathfrak{a}_1(\lambda) = t$, $\mathfrak{a}_1(\rho) = b$, $\mathfrak{a}_2(L) = l$, $\mathfrak{a}_2(M) = m$, and $\mathfrak{a}_2(R) = r$ is a semi-consistent equilibrium. Given $\mathfrak{a}_1$ player 2's optimization problem is identical to the one studied in Example 1, and $\mathfrak{a}_2$ is a semi-consistent best reply to $\mathfrak{a}_1$.

To see that $\mathfrak{a}_1$ is a best reply to $\mathfrak{a}_2$ first note that agent 1 obtains the utility 3.7 for the given strategy profile:

$$\min_{\alpha\in[.1,.2]} \big(u_1((t,l),\lambda)(.5-\alpha) + u_1((t,m),\lambda)\alpha + u_1((b,m),\rho)(.3-\alpha) + u_1((b,r),\rho)(.2+\alpha)\big) = \min_{\alpha\in[.1,.2]} \big(1(.5-\alpha) + 10\alpha + 10(.3-\alpha) + 1(.2+\alpha)\big) = 3.7$$

If agent 1 plays $Safe$ when he sees the signal λ his utility is at most $\min_{\alpha\in[.1,.2]} 2(.5-\alpha) + 2\alpha + 10(.3-\alpha) + 1(.2+\alpha) = 2.4$ given that his payoff $u_1((Safe, a_2), \sigma)$ equals 2 and given that 10 and 1 are his maximal payoffs when agent 2 respectively plays m and r. So agent 1 should not play $Safe$ if he is of type λ, by symmetry he should not play $Safe$ if he is of type ρ either. Fix a strategy $\mathfrak{a}'_1$ with $\mathfrak{a}'_1(\omega) \neq Safe$ for all ω. Given that $u_1((a_1, m), \omega) = 10$ holds if $a_1 \in \{t, b\}$ agent 1's maxmin utility at the profile $(\mathfrak{a}'_1, \mathfrak{a}_2)$ can be calculated as

$$\min_{\alpha\in[.1,.2]} \big(u_1((\mathfrak{a}_1(\lambda), l), \lambda)(.5-\alpha) + 10 \times .3 + u_1((\mathfrak{a}'_1(\rho), r), \rho)(.2+\alpha)\big) = \min_{\alpha\in[.1,.2]} u_1\big(((\mathfrak{a}_1(\lambda), l), \lambda)(.5-\alpha) + 3 + u_1((\mathfrak{a}'_1(\rho), r), \rho)(.2+\alpha)\big).$$

Since $u_1((t,l),\lambda) = u_1((b,r),\rho) = 1 > u_1((b,l),\lambda) = u_1((t,r),\rho) = 0$ this expression is uniquely maximized at $\mathfrak{a}_1$. Since $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are unique pure best replies against each other, $\mathfrak{a}$ is a semi-consistent equilibrium.

To see that $\mathfrak{a}$ is not a Bayes Nash equilibrium in any $\big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in N}, G\big)$ consider player 2's action $\mathfrak{a}_2(M) = m$ in the case that he observes M. If $\pi_2(\lambda \mid M) \geq \frac{1}{2}$ he is better off playing l; If $\pi_2(\lambda \mid M) \leq \frac{1}{2}$ he is better off playing r. So there is no prior π_2 for which $\mathfrak{a}_2$ is a best reply to $\mathfrak{a}_1$.

The above example uses the strategic context to amplify the impact of ambiguity aversion already identified Example 1. Player 1 is happiest if player 2 chooses m while he himself chooses t or b. If player 1, does not go for his safe option, but instead plays the the only other undominated actions given his signals, then player 2 faces the decision problem defined in Example 1. Given his ambiguity aversion and semi-conistency he sometimes chooses m, the action which levels out all ambiguity. If it is not sufficiently likely for player 2 to choose m then player 1 withdraws to his safe action. The ambiguity levelling choice m has no appeal to player 2 if he is Bayesian. Given that a Bayesian player 2 will never play m, $Safe$ is player 1's unique best reply to any optimal action of a Bayesian player 2. While players 1 and 2 respectively obtain maxmin utilities of 3.7 and $.7 \times 21 + .3 \times 10 = 14.7$ under q in $\big((\mathcal{Q}_i, A_i, u_i, C^*)_{i\in N}, G\big)$, they do obtain the utilities $(2, 0)$ in any equilibrium of any Bayesian game $\big((\mathcal{Q}_i, A_i, u_i, \pi_i)_{i\in N}, G\big)$.

9 Mechanism Design

Keeping the state space Ω, the algebra Σ, the partition of payoff relevant events $\mathcal{R}$, and X fixed, a **mechanism design problem** D is defined as a collection $\big((\mathcal{Q}_i, u_i, \succsim_i)_{i\in N}, H\big)$. As in the definition of a game, there is a set N of agents. Agent i has the information partition $\mathcal{Q}_i$ on Ω and a preference $(u_i, \succsim_i)$ on Σ-measurable acts $F : \Omega \to \Delta X$. There is a finite set of social choices Y and the function $H : Y \times \Omega \to \Delta X$ maps each social choice y and state ω to a lottery $H(y, \omega)$ over outcomes X. The social choice y induces the $\mathcal{R}$-measurable act $H(y, \cdot) : \Omega \to \Delta X$. A social choice y^b is **very bad** if each agent prefers $H(y(\cdot), \cdot)$ to $H(y^b, \cdot)$. A **social choice function**

$scf : \Omega \to \Delta Y$ maps the state space Ω to the set of all lotteries over social choices. The social choice function is deterministic if it maps each ω to a social choice y (degenerate lottery).

A **mechanism** $M : A \times \Omega \to \Delta Y$ sets out an action set A_i for each agent and maps any action profile $a \in A$ and state $\omega \in \Omega$ to a lottery over social choices $M(a, \omega) \in \Delta Y$. The design problem D and the mechanism M induce the game $\Gamma = ((\mathcal{Q}_i, A_i, u_i, \precsim_i)_{i \in N}, G)$ where the function $G : A \times \Omega \to \Delta X$ is defined such that $G(a, \omega) = \sum_{y \in Y}(M(a, \omega)(y))H(y, \omega)$. So $G(a, \omega)(x)$ the probability of outcome x at state ω at the strategy profile a is the expected probability of $H(y, \omega)(x)$ given that the social choice y is drawn from the distribution $M(a, \omega)$.

The mixed strategy profile $q \in \Delta A$ **induces** the social choice function $scf : \Omega \to \Delta Y$ with $scf(\omega) = \sum_{\mathfrak{a} \in \mathcal{A}} q(\mathfrak{a}) M(\mathfrak{a}(\omega), \omega)$ for all $\omega \in \Omega$. The mechanism M **consistently (semi-consistently) implements** the social choice function scf in D if the game has a consistent (semi-consistent) equilibrium q that induces scf. A mechanism $\hat{M} : B \times \Omega \to \Delta Y$ is a direct revelation mechanism if B_i is the set $\mathcal{Q}_i$ of agent i's types. Pure strategies in the game induced by $\hat{M}$ in D are denoted $\mathfrak{b} \in \mathcal{B}$, the set of mixed strategies is $\Delta\mathcal{B}$. The truthtelling strategy $t_i \in \Delta\mathcal{B}_i$ assigns probability 1 to the pure truthtelling strategy $\theta_i \in \mathcal{B}$ that maps any ω to the agents type θ_i at ω (so $\theta_i(\omega) = \theta_i$ if $\omega \in \theta_i$). The revelation principle, which states that a social choice function is implementable if and only if it is implementable via a direct revelation mechanism only holds with some modifications:

Theorem 5 *Fix a mechanism design problem* $D = \big((\mathcal{Q}_i, u_i, \precsim_i)_{i \in N}, H\big)$ *and a social choice function* $scf : \Omega \to \Delta Y$.

a) If scf *is consistently implementable, then* scf *is consistently implementable via a direct revelation mechanism.*

b) If scf *is semi-consistently implementable, then* scf *need not be semi-consistently implementable via a direct revelation mechanism.*

c) If a pure strategy profile semi-consistently implements scf*, then* scf *is semi-consistently implementable via a direct revelation mechanism.*

The proof of parts a) and c) adapts the from the proof for the revelation principle with expected utility maximizers to the case that does not require ambiguity neutrality. This part of the proof is in the Appendix. Part b) is shown by the following variation of the Ellsberg paradox.

Example 3 Let $X = \{w, l\}$, where w stands for winning and l for loosing, let $Y = \{b, r, mid\}$ representing the choices of betting on black, red or a known odds bet. Let there be just one agent, so the design problem with a maxmin expected utility can be represented as $D\colon = (\mathcal{Q}, u, C, H)$. Let $\mathcal{Q}\colon = \{\nu, \beta\}$ describe the agent's information and let there be two payoff relevant events so $\mathcal{R} = \{B, R\}$ where B and R respectively stand for a black and a red ball being drawn. The agent's set of beliefs $C\colon = \{(\frac{1}{4}, \frac{1}{4}, \frac{1}{4} - \alpha, \frac{1}{4} + \alpha) \mid \alpha \in [.1, .2]\}$, where any vector of probabilities $\pi \in C$ stands for $(\pi(\nu \cap B), \pi(\nu \cap R), \pi(\beta \cap L), \pi(\beta \cap R))$. So under C the agent of type ν assigns probability $\frac{1}{2}$ to B and R. The agent of type β is not sure with which probability B ad R occur. The agent's utility over lotteries over winning and loosing (ΔX) is such that $u(w) = 1$ and $u(l) = 0$ and I represent any $p \in \Delta X$ by $p(w) \in [0, 1]$. The function $H : Y \times \Omega \to \Delta X$ is such that $H(b, B) = 1$, $H(r, R) = 1$, $H(b, R) = 0$, $H(r, B) = 0$, and $H(mid, \omega) = .49$ for all $\omega \in \Omega$.

Denoting the elements of ΔY as two component vectors, with the understanding that these two components represent the probability of b and r, the $scf : \Omega \to \Delta Y$ with $scf(\nu) = (\frac{1}{2}, \frac{1}{2})$ and $scf(\beta) = (0, 0)$ is weakly consistently implementable in D. To see this consider the mechanism $M : A \times \Omega \to \Delta Y$ that directly lets the agent choose from the set Y, so $A\colon = Y$ and $M(y, \omega) = y$ for all ω. The game induced by M in D amounts to a classical Ellsberg decision problem: the agent can choose any lottery over b, r, and mid. The numbers are chosen such that the agent of the ambiguity neutral type ν is happy to choose the lottery $(\frac{1}{2}, \frac{1}{2})$ over b and r, whereas the ambiguity averse type β has $(0, 0)$ as his only weakly consistent optimal choice. So scf which coincides with this plan $\mathfrak{a}$ is semi-consistently implementable in the mechanism.

Now suppose that scf was semi-consistently implementable via any truth-telling mechanism $(\{\nu, \beta\}, \hat{M})$. Note first that $\hat{M} : \{\nu, \beta\} \times \Omega \to \Delta Y$ must be such that $\hat{M}(\nu, \omega) = (\frac{1}{2}, \frac{1}{2})$ and $\hat{M}(\beta, \omega) = (0, 0)$ for scf to be induced by the truth-telling strategy profile in the game induced by $(\{\nu, \beta\}, \hat{M})$ in D. But the agent of type β then prefers to announce ν, and scf is not implementable.

The logic that underlies the standard revelation principle goes as follows: Suppose some social choice function scf is implemented by some mechanism. Then we may construct another mechanism in which the agents directly report their types to the designer who then executes the equilibrium strategies for the agents in the old mechanism. Telling the truth then is optimal given that all other agents tell the truth, since the situation in which all agents tell the truth is equivalent to the situation in which all agents follow the equilibrium strategies. This equivalence however depends on the fact that agents choose lotteries if and only if each of the pure choices they randomize over is an optimal choice. The crux is that this equivalence need not hold for a semi-consistent agent.

In the example the semi-consistent agent of type β is ambiguity averse: he would like to choose a 50/50 mixture over b and r. But, being only semi-consistent, he is not able to choose this lottery. He knows that he would not follow through with it, given that mid is also available. So he has to resign himself to the lottery mid. However, any direct revelation mechanism implementing a social choice function in which the agent of type ν obtains a 50/50 lottery on b and r has to map the announcement of ν to that lottery. Such a mechanism then eliminates type β's commitment problem. If the agent of type β announces that he is of type ν he receives his most preferred lottery on $\{r, b, mid\}$ and the social choice function cannot be implemented.

While a mechanism $((A_i)_{i\in N}, M)$ may induce an extensive form game in which agents sequentially take actions, Lemma **??** shows that any strongly consistently implementable choice function is also strongly consistently implementable in a mechanism in which agents simultaneously choose their actions. It is moreover without loss of generality to identify these action spaces with the agents information partitions.

Can ambiguous mechanisms make a difference? To address this questions I impose some further conditions to align the present model as closely as possible with the standard model. Ambiguity then only appears as a feature of the function $M : A \times \Omega \to \Delta Y$.

In the next result I show that the introduction of ambiguous information alone does not change the set of of implementable social choice functions. For this result I adopt the standard assumptions that all agents are expected utility maximizers with respect to $\mathcal{Q}\wedge\mathcal{R}$-measurable acts and that the social choice function to be implemented is $\mathcal{Q}$-measurable. So ambiguity only enters as a feature of the mechanism via the feature of the function $M : A \times \Omega \to \Delta Y$. The only difference between the present setup and that of Bose and Renou [3] is that the agents in Theorem 6 are consistent or semi-consistent. The agents in Bose and Renou [3] are inconsistent.

Theorem 6 *Fix a mechanism design problem $D = ((\mathcal{Q}_i, u_i, \succsim_i)_{i\in N}, H)$ such that the restriction of any $(u_i, \succsim_i)$ to the set of $\mathcal{Q}$-measurable acts has an expected utility representation. Either assume that $\pi(\theta) > 0$ holds for all $\theta \in \mathcal{Q}$ or that there exists a very bad social choice y^b. Consider the set of $\mathcal{Q}$-measurable social choice functions. Such a social choice function scf is (consistently or semi-consistently) implementable in D if and only if it is implementable in some $D' = ((\mathcal{Q}_i, u_i, \pi_i)_{i\in N}, H)$ where (u_i, π_i) coincides with $(u_i, \succsim_i)$ on the set of all $\mathcal{Q} \wedge \mathcal{R}$-measurable acts.*

Proof **to be written** □

We furthermore obtain a corollary to Theorem **??** the result on observational equivalence.

Corollary 2 *Fix a mechanism design problem $D = ((\mathcal{Q}_i, u_i, \succsim_i)_{i\in N}, H)$ if the social choice function scf is consistently implementable in D then it is implementable in some $D' = ((\mathcal{Q}_i, u_i, \pi_i)_{i\in N}, H)$.*

The next result shows that ambiguous communication can make a difference if we do not only consider $\mathcal{Q}$-measurable social choice functions but also such social choice functions that let the outcomes depend on some (possibly ambiguous) randomization device. Such a social choice function $scf : \Omega \to \Delta Y$ is then $\mathcal{Q}'$-measurable where $\mathcal{Q}'$ is strictly finer than $\mathcal{Q}$.

Theorem 7 *Fix a mechanism design problem $D = ((\mathcal{Q}_i, u_i, \succsim_i)_{i\in N}, H)$ and a social choice function $scf : \Omega \to \Delta Y$. Say all agents share a prior π^* on $\mathcal{Q}$. a) If $scf : \Omega \to \Delta Y$ is weakly consistently implementable in D, then scf need not be implementable in any $D' = ((\mathcal{Q}_i, u_i, \pi_i)_{i\in N}, H)$ where each π_i has π^* as its $\mathcal{Q}$-marginal. b) If $scf : \Omega \to \Delta X$ is strongly consistently implementable* *in D, then scf is implementable in some* $D' = ((\mathcal{Q}_i, u_i, \pi_i)_{i\in N}, H)$ *where each* π_i *has* π^* *as its* $\mathcal{Q}$*-marginal.*

Proof Define a mechanism design problem $D = \big((\mathcal{Q}_i, u_i, C^{**}))_{i\in\{1,2\}}, h\big)$ with $\mathcal{Q}_1 := \{\lambda, \rho\}$ and $\mathcal{Q}_2 := \{\lambda^*, no, \rho^*\}$ where $\lambda^* \subset \lambda$ and $\rho^* \subset \rho$. The set of social choices is $Y = \{l, m, r\}$, the partition of payoff relevant events is $\mathcal{R} = \{\lambda, \rho\}$. Payoffs $u_i(H(y, \omega))$ are given by the following table

	l	m	r
λ	$1, 21$	$10, 10$	$0, 0$
ρ	$0, 0$	$10, 10$	$1, 21$

payoffs $u_i(H(y, \omega))$

With respect to $\mathcal{Q}$-measurable acts $F : \Omega \to \Delta X$ the two agents are expected utility maximizers; all priors $\pi \in C^{**}$ agree on any $\theta \in \mathcal{Q}$ and we have $\pi(\lambda) = \pi(\rho) = \frac{1}{2}$ and $\pi(\lambda^*) = \pi(\rho^*) = .1$.

Fix a $\mathcal{Q}$-measurable social choice function $scf : \Omega \to \{l, m, r\}$ with $scf(\lambda^*) = l$ and $scf(\rho^*) = r$. I next show that scf can only be weakly consistently implementable in D if $scf(\omega) \neq m$ for all $\omega \in \Omega$. Suppose not. By Lemma **??** scf would have to be implementable with a direct revelation mechanism. First consider the case that $scf(no) = m$. If agent 2 tells the truth his ex ante utility is $.1 \times 21 + (.4 + .4) \times 10 + .1 \times 21$.

$$\pi(\lambda^*)u_2(H(scf(\lambda^*), \lambda))) + \pi(no)u_2((H(scf(no), \cdot))) + \pi(\rho^*)u_2(H(scf\rho^*, \rho)) = \\ .1 \times 21 + .8 \times 10 + .1 \times 21$$

If agent 2 deviates to the strategy according to which he truthfully reports λ^* and ρ^*, but declares λ^* when he receives the signal no he increases his utility to

$$\pi(\lambda^*)u_2(H(scf(\lambda^*), \lambda))) + \pi(\lambda \cap no)(H(scf(\lambda^*), \lambda))) + \\ \pi(\rho \cap no)u_2(H(scf(\lambda^* \cap \rho), \rho) + \pi(\rho^*)u_2(H(scf(\rho^*), \rho)) \geq \\ .1 \times 21 + .4 \times 21 + .4 \times 0 + .1 \times 21.$$

The inequality follows from the fact that $u_2(scf(\lambda^* \cap \rho), \rho)$ cannot be below 0, as the mechanism must choose l, r, or m - even when it is clear that at least one agent lies.

If $scf(no \cap \lambda) = m$ while $scf(no \cap \rho) \neq m$ then agent 1 obtains at most

$$\pi(\lambda^*)u_1(H(scf(\lambda^*), \lambda))) + \pi(no \cap \lambda)u_1((H(scf(no \cap \lambda), \lambda))) +$$
$$\pi(no \cap \rho)u_1((H(scf(no \cap \rho), \rho))) + \pi(\rho^*)u_1(H(scf\rho^*, \rho)) =$$
$$.1 \times 1 + .4 \times 10 + .4 \times u_1((H(scf(no \cap \rho), \rho))) + .1 \times 1 \leq$$
$$.1 \times 1 + .4 \times 10 + .4 \times 1 + .1 \times 1$$

where the inequality follows from $u_1((H(scf(no \cap \rho), \rho)))$ being at most 1 as $scf(no \cap \rho) \neq m$. If agent always claims he observed λ he increases his utility to

$$\pi(\lambda^*)u_1(H(scf(\lambda^*), \lambda))) + \pi(no \cap \lambda)u_1((H(scf(no \cap \lambda), \lambda))) +$$
$$\pi(no \cap \rho)u_1((H(scf(no \cap \lambda), \rho))) + \pi(\rho^*)u_1(H(scf\rho^* \cap \lambda, \rho)) =$$
$$.1 \times 1 + .4 \times 10 + .4 \times 10 + .1 \times u_1(H(scf\rho^* \cap \lambda, \rho) \geq$$
$$.1 \times 1 + .4 \times 10 + .4 \times 1 + .1 \times 0$$

where the inequality follows from 0 being a lower bound on $u_1(H(scf\rho^* \cap \lambda, \rho)$. Mutatis mutandis agent 1 is better off always claiming his signal was ρ if $scf(no \cap \rho) = m$ and $scf(no \cap \lambda) \neq m$. Since we only consider $\mathcal{Q}$-measurable social choice functions scf here, this is an exhaustive list of cases.

Now consider the mechanism $M = \big((\mathcal{Q}_i)_{i \in N}, M\big)$ in which the designer may partition any event in $\mathcal{Q}$ more finely. Say he designer has access to another partition $\mathcal{D} = \{L, M, R\}$. Assume that the agents' set of beliefs on $\{L, M, R\}$ is as in Example 1. Let the set of beliefs be given by $\alpha \in [.1, .2]$ and the following matrix

	L	M	R
λ^*	$.1$	0	0
λ	$.4 - \alpha$	α	0
ρ	0	$.3 - \alpha$	$.1 + \alpha$
ρ^*	0	0	$.1$

the set of priors C^{**}

The $\mathcal{Q} \wedge \mathcal{D}$-measurable social choice function $scf^* : \Omega \to Y$ with $scf^*(\lambda^*) = scf^*(L) = l$, $scf^*(\rho^*) = scf^*(R) = r$, and $scf^*(M) = m$ is implementable.

The mechanism M' does not assume any extraneous knowledge by the designer. Loosely following Bose and Renou [3] we can think of the following process. First the agents truthfully report their types. The designer then knows the partition $\mathcal{Q} = \{\lambda^*, no \cap \lambda, no \cap \rho, \rho^*\}$. The designer then uses an ambiguous randomization device to partition the events $no \cap \lambda$ and $no \cap \rho$. The designer could for example draw a ball from an urn filled with red and black balls in unknown proportion. The draw of a black ball is mapped to L if the agents reported no and λ. The draw of a red ball is mapped to R if the agents reported no and ρ. The remaining draws are mapped to the event M if no was reported. Given that agent 1 truthfully reports his information agent 2's decision problem is very similar to the problem studied in Example 1. He now has an incentive to truthfully reveal no. Now only M is mapped to m, but the ambiguity in the mechanism was constructed such that in the event M agent 2 faces enough ambiguity to prefer the insurance option m to l and r. □

A Appendix

Proof of parts a) and c) of Theorem 5

Fix a mechanism $M : A \times \Omega \to \Delta Y$ and consider the induced game $\Gamma = \big((\mathcal{Q}_i, A_i, u_i, \succsim_i)_{i \in N}, G\big)$ with $G(a,\omega) = \sum_{y \in Y} \big(M(a,\omega)(y)\big) H(y,\omega)$ for all $(a,\omega) \in A \times \Omega$.

a) Say Γ has a consistent equilibrium q^* that induces scf. We therefore have $scf(\omega) = \sum_{\mathfrak{a} \in \mathcal{A}} q^*(\mathfrak{a}) M\big(\mathfrak{a}(\omega), \omega\big)$ for all $\omega \in \Omega$. Define the direct revelation mechanism $\hat{M} : B \times \Omega \to \Delta Y$ such that $\hat{M}(b,\omega) = \sum_{\mathfrak{a} \in \mathcal{A}} q^*(\mathfrak{a}) M(\mathfrak{a}(b), \omega)$ for all $(b,\omega) \in B \times \Omega$. The strategy profile $q \in \Delta\mathcal{B}$ in the direct revelation mechanism induces the social choice function that maps any ω to

$$\sum_{\mathfrak{b} \in \mathcal{B}} q(\mathfrak{b}) \hat{M}(\mathfrak{b}(\omega), \omega) = \sum_{\mathfrak{b} \in \mathcal{B}} q(\mathfrak{b}) \sum_{\mathfrak{a} \in \mathcal{A}} q^*(\mathfrak{a}) M(\mathfrak{a}(\mathfrak{b}(\omega)), \omega)$$

The truth-telling strategy profile $t \in \Delta\mathcal{B}$ in $\hat{M}$ induces the social choice functions scf in D: we have

$$\sum_{\mathfrak{b}\in\mathcal{B}} t(\mathfrak{b})\hat{M}(\mathfrak{b}(\omega),\omega) = \hat{M}(\theta(\omega),\omega) = \sum_{\mathfrak{a}\in\mathcal{A}} q^*(\mathfrak{a})M(\mathfrak{a}(\omega),\omega)$$

where the first equality follows from t assigning probability 1 to θ and the second follows from the definition of $\hat{M}$.

If an agent i deviates to q_i' in the new mechanism while all other agents follow the truth-telling strategy he induces the social choice function that maps each ω to

$$\begin{aligned}
&\sum_{\mathfrak{b}\in\mathcal{B}} q_i'(\mathfrak{b}_i)t_{-i}(\mathfrak{b}_{-i})\hat{M}\big(\mathfrak{b}(\omega),\omega\big) = \\
&\sum_{\mathfrak{b}\in\mathcal{B}} q_i'(\mathfrak{b}_i)t_{-i}(\mathfrak{b}_{-i})\sum_{\mathfrak{a}\in\mathcal{A}} q^*(\mathfrak{a})M\big(\mathfrak{a}((\mathfrak{b}(\omega)),\omega\big) = \\
&\sum_{\mathfrak{b}_i\in\mathcal{B}_i} q_i'(\mathfrak{b}_i)\sum_{\mathfrak{a}\in\mathcal{A}} q_i^*(\mathfrak{a}_i)q_{-i}^*(\mathfrak{a}_{-i})M\big((\mathfrak{a}_i(\mathfrak{b}_i(\omega)),\mathfrak{a}_{-i}(\omega)),\omega\big) = \\
&\sum_{\mathfrak{a}_i\in\mathcal{A}_i} q_i''(\mathfrak{a}_i)\sum_{\mathfrak{a}_{-i}\in\mathcal{A}_i} q_{-i}^*(\mathfrak{a}_i)M\big((\mathfrak{a}(\omega)),\omega\big) = \\
&\sum_{\mathfrak{a}\in\mathcal{A}} q_i''(\mathfrak{a}_i)q_{-i}^*(\mathfrak{a}_i)M\big(\mathfrak{a}(\omega),\omega\big)
\end{aligned}$$

where

$$q_i''(\mathfrak{a}_i) = \sum_{(\mathfrak{a}_i',\mathfrak{b}_i)\in S(\mathfrak{a}_i)} q_i'(\mathfrak{b}_i)q^*(\mathfrak{a}_i')$$

and where $S(\mathfrak{a}_i')\colon = \big\{(\mathfrak{a}_i',\mathfrak{b}_i) : \mathfrak{a}_i(\omega) = \mathfrak{a}_i'(\mathfrak{b}_i(\omega) \text{ for all } \omega\in\Omega\big\}$ is the set of all combinations of strategies $\mathfrak{a}_i'\in\mathcal{A}$ and $\mathfrak{b}_i\in\mathcal{B}_i$ for which the concatenation $\mathfrak{a}_i'\circ\mathfrak{b}_i$ that appears in the above expression equals $\mathfrak{a}_i$.

Since q is a consistent equilibrium we have:

$$\begin{aligned}
&u_i\big(\sum_{\mathfrak{a}\in\mathcal{A}} q^*(\mathfrak{a})\sum_{y\in Y} M((\mathfrak{a}(\cdot),\cdot))(y)H(y,\cdot)\big) \succsim_i \\
&u_i\big(\sum_{\mathfrak{a}\in\mathcal{A}} q_i'(\mathfrak{a}_i)q_{-i}^*(\mathfrak{a}_i)\sum_{y\in Y} M((\mathfrak{a}(\cdot),\cdot)(y))H(y,\cdot)\big)
\end{aligned}$$

and agent i prefers the truth-telling strategy t_i to all other strategies q_i' under the direct revelation mechanism given that the other agents follow the truthtelling strategy.

c) Say Γ has a consistent equilibrium $\mathfrak{a}^*$ that induces scf. We therefore have $scf(\omega) = M\big(\mathfrak{a}^*(\omega), \omega\big)$ for all $\omega \in \Omega$. Define the direct revelation mechanism $\hat{M} : B \times \Omega \to \Delta Y$ such that $\hat{M}(b, \omega) = M(\mathfrak{a}^*(b), \omega)$ for all $(b, \omega) \in B \times \Omega$. The strategy profile $q \in \Delta \mathcal{B}$ in the direct revelation mechanism induces the social choice function that maps any ω to

$$\sum_{\mathfrak{b} \in \mathcal{B}} q(\mathfrak{b}) \hat{M}(\mathfrak{b}(\omega), \omega) = \sum_{\mathfrak{b} \in \mathcal{B}} q(\mathfrak{b}) M(\mathfrak{a}^*(\mathfrak{b}(\omega)), \omega)$$

The truth-telling strategy profile $t \in \Delta \mathcal{B}$ in $\hat{M}$ induces the social choice functions scf in D: we have

$$\sum_{\mathfrak{b} \in \mathcal{B}} t(\mathfrak{b}) \hat{M}(\mathfrak{b}(\omega), \omega) = \hat{M}(\theta(\omega), \omega) = M(\mathfrak{a}^*(\omega), \omega)$$

where the first equality follows from t assigning probability 1 to θ and the second follows from the definition of $\hat{M}$.

If an agent i deviates to q_i in the new mechanism while all other agents follow the truth-telling strategy he induces the social choice function that maps each ω to

$$\begin{aligned} \sum_{\mathfrak{b} \in \mathcal{B}} q_i(\mathfrak{b}_i) t_{-i}(\mathfrak{b}_{-i}) \hat{M}\big(\mathfrak{b}(\omega), \omega\big) &= \\ \sum_{\mathfrak{b} \in \mathcal{B}} q_i(\mathfrak{b}_i) t_{-i}(\mathfrak{b}_{-i}) \; M\big(\mathfrak{a}^*(\mathfrak{b}(\omega)), \omega\big) &= \\ \sum_{\mathfrak{b}_i \in \mathcal{B}_i} q_i(\mathfrak{b}_i) M\big((\mathfrak{a}_i^*(\mathfrak{b}_i(\omega)), \mathfrak{a}_{-i}^*(\omega)), \omega\big) &= \\ \sum_{\mathfrak{a}_i \in \mathcal{A}_i} q_i(\mathfrak{b}_i \mid \mathfrak{a}_i = \mathfrak{a}_i^* \circ \mathfrak{b}_i) M\big((\mathfrak{a}_i(\omega), \mathfrak{a}_{-i}^*(\omega)), \omega\big) & \end{aligned}$$

If q_i assigns probability 1 to some strategy $\mathfrak{b}_i$ then the social choice function $M\big((\mathfrak{a}_i^*(\mathfrak{b}_i \omega), \mathfrak{a}_{-i}^*(\omega)), \omega\big)$ is induced. Since $\mathfrak{a}^*$ is a semi-consistent equilibrium we have:

$$u_i\Big(\sum_{y\in Y} M((\mathfrak{a}^*(\cdot),\cdot))(y)H(y,\cdot)\Big) \succsim_i$$
$$u_i\Big(\sum_{y\in Y} M((\mathfrak{a}_i(\cdot),\mathfrak{a}^*_{-i}(\cdot)),\cdot)(y))H(y,\cdot)\Big)$$

for any $\mathfrak{a}_i \in \mathcal{A}_i$, in particular ass $\mathfrak{a}_i$ that arise as the composition $\mathfrak{a}^*_i \circ \mathfrak{b}_i$. Consequently the truthtelling is among the best pure strategies given that everyone else tells the truth.

To see that

that only assign positive probability to optimal pure strategies $\mathfrak{a}_i$. We in particular obtain that

still not yet there

References

[1] Aumann, R.: "Subjectivity and correlation in randomized strategies", *Journal of Mathematical Economics*, 1 (1974), 6796.

[2] Azrieli, Y. and R. Teper: "Uncertainty Aversion and Equilibrium Existence in Games with Incomplete Information", *Games and Economic Behavior*, 73 (2011), 310-317.

[3] Bose, S. and L. Renou: "Mechanism Design with Ambigous Communication Devices", *Econometrica*, 82 (2014), 18531872.

[4] Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio, "Uncertainty Averse Preferences", *Journal of Economic Theory*, 146, (2011) pp. 1275-1330.

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

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

[7] Ellis, A.: "Condorcet Meets Ellsberg", forthcoming *Theoretical Economics*

[8] Hanany, E. and P. Klibanoff, "Updating Preferences with Multiple Priors", *Theoretical Economics*, 2, (2007), pp. 261-298.

[9] Hanany, E. and P. Klibanoff, "Updating Ambiguity Averse Preferences", *The B.E. Journal of Theoretical Economics*, 9, (Advances), article 37.

[10] Kellner, C. and M. Le Quemet: "Endogenous ambiguity in cheap talk" *mimeo* Bonn University 2015.

[11] Klibanoff, P., "Stochastically Independent Randomization and Uncertainty Aversion", *Economic Theory*, 18, (2001), pp. 605-620.

[12] Klibanoff, P., M. Marinacci, and S. Mukerji, "A Smooth Model of Decision Making under Ambiguity", *Econometrica*, 73, (2005), pp. 1849-1892.

[13] Kuzmics, C. "Abraham Wald's complete class theorem and Knightian uncertainty" *mimeo* (2015) University of Bielefeld. Christoph Kuzmics

[14] Machina M., "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty", *Journal of Economic Literature*, 27 (1989), 1622- 1668.

[15] McClennen, E.: *Rationality and dynamic choice: foundational explorations* (1990), Cambridge University Press, Cambridge.

[16] Raiffa, H. "Risk, Ambiguity, and the Savage Axioms: Comment", *The Quar- terly Journal of Economics*, 75, (1961) 690-694.