University of Heidelberg

Department of Economics

Discussion Paper Series | No. 573

Learning under Ambiguity – A Note on the Belief Dynamics of Epstein and Schneider (2007)

Daniel Heyen

October 2014

Learning under Ambiguity – A Note on the Belief Dynamics of Epstein and Schneider (2007)

Daniel Heyen*
Department of Economics
Heidelberg University

October 9, 2014

Abstract Epstein and Schneider (2007) develop a framework of learning under ambiguity, generalizing maxmin preferences of Gilboa and Schmeidler (1989) to intertemporal settings. The specific belief dynamics in Epstein and Schneider (2007) rely on the rejection of initial priors that have become implausible over the learning process. I demonstrate that this feature of ex-post rejection of theories gives rise to choices that are in sharp contradiction with ambiguity aversion. Concrete, the intertemporal maxmin decision-maker equipped with such belief dynamics prefers, under prevalent conditions, a bet in an ambiguous urn over the same bet in a risky urn. I offer two modifications of their framework, each of which is capable of avoiding this anomaly.

Keywords learning under ambiguity, multiple prior, maxmin, ambiguity aversion

JEL D81, D83

*heyen@eco.uni-heidelberg.de

1 Introduction

Since Ellsberg (1961), ambiguity averse preferences as opposed to subjective expected utility preferences (SEU, Savage 1954) have been a vital economic research topic. One of the most influential axiomatization of ambiguity aversion is the multiple prior model with maxmin expected utility (MEU) of Gilboa and Schmeidler (1989). The framework of Gilboa and Schmeidler (1989), however, is atemporal and thus not capable of reflecting intertemporal ambiguity aversion. In a recent and influential paper, Epstein and Schneider (2007) (henceforth ES) develop, based on Epstein and Schneider (2003), a tractable framework of intertemporal maxmin preferences and thus opens ambiguity aversion to a dynamic learning environment. This framework has already been used extensively in various contexts, including financial markets (Condie and Ganguli 2011; Garlappi et al. 2007; Ju and Miao 2012; Leippold et al. 2008) and real options (Nishimura and Ozaki 2007; Riedel 2009).

The focus of this paper is a problematic characteristic of the belief dynamics in Epstein and Schneider (2007). The updating process in ES involves the rejection of initial beliefs that have become less plausible given the observed signal history. I demonstrate that the rejection of beliefs renders possible a switch in preferences. To illustrate this, I construct a simple example in which an ambiguity averse decision-maker switches to ambiguity loving behavior after observing one draw from a payoff-relevant urn. The reason for this instability of preferences is found in the rejection of initial beliefs in the updating procedure. It can happen that the uniform distribution, the standard initial prior of the SEU decision-maker, is rejected in the reevaluation procedure of ES. As a consequence, the set of updated beliefs does not include the posterior of the SEU decision-maker, who serves as the usual reference point to define ambiguity preferences. This gives rise to the switch to ambiguity loving behavior.[1] Furthermore, I show that this anomaly of intertemporal beliefs is not just an artifact, but rather a pervasive and general feature of the ES belief dynamics. Finally, I suggest two modifications of the ES setting to ensure stable ambiguity averse preferences over time.

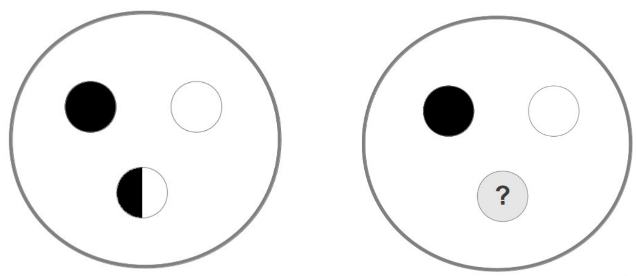
I proceed as follows. In section 2, I recapitulate the basic cornerstones of Epstein and Schneider (2007). The simple example demonstrating the ES anomaly in a standard setting is found in section 3. Section 4 generalizes the example to symmetric urns with an arbitrary number of balls and shows that the ES anomaly may occur under very general conditions. The two parts of the theorem are illustrated graphically in section 5. In section 6, I offer two alternatives to overcome the problems while keeping the general structure of ES. I conclude in section 7.

[1] ES notes that Ellsberg type behavior in the short run will converge (for a fixed composition urn) to ambiguity neutral behavior in the long run. The possibility of a change in preferences to ambiguity loving choices, however, was not mentioned.

2 Intertemporal maxmin preferences – The setting of Epstein and Schneider (2007)

The main purpose of this section is to recall the basic components of Epstein and Schneider (2007). All readers familiar with the ES framework may thus skip this section.

To reemphasize the motivation of their framework, I consider two urns that are simplifications of the scenarios in ES. Both urns only features risk or ambiguity, but not both kinds of uncertainty at the same time. These urns are used in section 3 to demonstrate the existence of the ES anomaly and are then generalized in section 4 to arbitrary symmetric settings.



Urn R: Risk   Urn A: Ambiguity

Figure 1

2.1 Unknown parameter

To motivate the introduction of the parameter space in ES, consider the two urns in Figure 1. Both urns contain exactly three balls (this assumption will be abandoned in section 4). In both urns there is, apart from a white and a black ball, a third ball that is either black or white. The composition of either urn is thus unknown to the decision-maker but will not change over the course of the experiment.[2] In the language of ES, the ratio of black balls in the urn is the unknown *parameter* θ with possible values in the set $\Theta = \{1/3, 2/3\}$. The key difference between both urns is that for the risky Urn R, the decision-maker knows that the color of the third ball has been determined via a fair mechanism, e.g. a fair coin. This is the same kind of uncertainty as in scenario 1 in ES. For the ambiguous Urn A however, the decision-maker has no information on the mechanism that determined the color of the third ball. In contrast to scenario 2 in ES, Urn A features pure ambiguity.

[2]Thus, we do not consider scenario 3 in Epstein and Schneider (2007) with unknown likelihoods

2.2 State space and recursive utility

In every period and for each urn, one ball is drawn and then put back (sampling with replacement). The period state space is $S_t = S = \{B, W\}$, identical for all times. We denote by $s_t \in S$ the color observed by the agent at time t. An agent's information at time t is the history $s^t = (s_1, \ldots, s_t)$. The natural full state space is S^∞.

The agent ranks consumption plans $c = (c_t)$ according to recursively defined utility,

$$U_t(c; s^t) = \min_{p \in \mathcal{P}_t(s^t)} \mathbb{E}^p \left[u(c_t) + \beta U_{t+1}(c; s^t, s_{t+1}) \right] \ , \tag{1}$$

where u and β have the usual properties. A central component in ES is $\mathcal{P}_t(s^t)$. This set of probability measures models beliefs about the next ball observed, s_{t+1}, given the history s^t. Such beliefs reflect ambiguity if $\mathcal{P}_t(s^t)$ is a non-singleton, which is the appropriate description if the draw will be made from Urn A. Beliefs about the next draw from Urn R can be, as usual, described with $\mathcal{P}_t(s^t)$ being a singleton. ES refer to $(\mathcal{P}_t)$ as the *process of one-step-ahead beliefs*. Specifying beliefs in this way ensures dynamic consistency of the decision framework, as is shown in Epstein and Schneider (2003). To further clarify this set of beliefs, let us turn to the learning structure in the ES setting.

2.3 Learning

By observing data, here a sequence (s_t) of white and black balls, the decision-maker tries to learn the color of the third ball and thus the unknown parameter $\theta \in \Theta$. Ambiguity in initial beliefs about parameters can be represented by a *set* $\mathcal{M}_0$ of probability measures on Θ. The size of $\mathcal{M}_0$ reflects the decision-maker's (lack of) confidence in the prior information on which initial beliefs are based.

In both urns, Urn R and Urn A, the likelihood of observing a black or a white ball is fully determined by the parameter θ, the ratio of black balls in the urn. Obviously, $l(s = B|\theta) = \theta$. Throughout this paper, we restrict ourselves to those settings of unique likelihood functions.[3] Beliefs about parameters and the likelihood function jointly determine the process of one-step-ahead beliefs

$$\mathcal{P}_t(s^t) = \left\{ p_t(\cdot) = \int_\Theta l(\cdot|\theta) d\mu_t(\theta) \ : \mu_t \in \mathcal{M}_t(s^t) \right\} \ , \tag{2}$$

where $\mathcal{M}_t(s^t)$ is the set of posterior beliefs after observing the data s^t. This set is basically the priorwise bayesian update of initial beliefs $\mu_0 \in \mathcal{M}_0$. ES, however,

[3] ES also allow for ambiguity in likelihoods, that is a set of likelihoods $\mathcal{L}$. At any point in time, any element of $\mathcal{L}$ might be relevant for generating the next observation. Multiple likelihoods refer to those components of a decision problem the decision-maker has decided that she will not try to (or is not able to) learn about. The findings of this paper are independent of $\mathcal{L}$, which is why I restrict this analysis to the simplest case with a single likelihood function.

incorporate the possibility that the decision-maker rejects at every t some implausible initial beliefs. The set of posteriors only include updates of initial beliefs that are not rejected in this process. The specific procedure of rejecting beliefs in ES, however, gives rise to the anomaly that is the topic of this paper. In the following subsection, we explain the rejection of beliefs in ES in detail.

2.4 Updating and reevaluation

To assess the plausibility of a "theory" $\mu_0 \in \mathcal{M}_0$ after having observed the history of signals s^t, ES use the data density evaluated at s^t. I denote this plausibility of a theory μ_0 given the data $s^t = (s_1, \ldots, s_t)$ by

$$\text{Plaus}(\mu_0; s^t) = \int_\Theta \prod_{j=1}^{t} l(s_j|\theta) d\mu_0(\theta) \ . \tag{3}$$

With the usual Bayesian updating, recursively defined by

$$d\mu_t(\cdot; s^t, \mu_0) = \frac{l(s_t|\cdot)}{\int_\Theta l(s_t|\theta') d\mu_{t-1}(\theta'; s^{t-1}, \mu_0)} d\mu_{t-1}(\cdot; s^{t-1}, \mu_0) \ , \tag{4}$$

ES define the set of posteriors

$$\mathcal{M}_t^\alpha(s^t) = \left\{ \mu_t(\cdot; s^t, \mu_0) \ : \mu_0 \in \mathfrak{M}_0^\alpha(s^t) \right\} \tag{5}$$

as the set of prior-by-prior updates of $\mathfrak{M}_0^\alpha(s^t)$. Here,

$$\mathfrak{M}_0^\alpha(s^t) = \left\{ \mu_0 \in \mathcal{M}_0 \mid \text{Plaus}(\mu_0; s^t) \geq \alpha \max_{\tilde{\mu}_0 \in \mathcal{M}_0} \text{Plaus}(\tilde{\mu}_0; s^t) \right\} \tag{6}$$

is the set of theories (i.e. initial priors) that are not rejected after having observed the signal history s^t. Rejected are those inital priors that fail a maximum likelihood test against the most plausible prior, and the parameter α governs how strict this maximum likelihood test is. ES consider $0 < \alpha \leq 1$ as possible values for the rejection parameter. Ruling out $\alpha = 0$ implies that they require the decision-maker to actually use this rejection device.

The likelihood-ratio test is more stringent and the set of posteriors smaller, the greater is α. In the extreme case $\alpha = 1$, only parameters that achieve the maximum likelihood are permitted. If the maximum likelihood estimator is unique, ambiguity about parameters is resolved as soon as the first signal is observed. More generally, we have that $\alpha > \alpha'$ implies $\mathcal{M}_t^\alpha \subset \mathcal{M}_t^{\alpha'}$. It is important that the test is done after every history. In particular, a theory that was disregarded at time t might look more plausible at a later time and posteriors based on it may again be taken into account.

2.5 Subjective expected utility vs. maxmin preferences

Some words on the relation between subjective expected utility (SEU) and intertemporal maxmin (MEU) preferences are in order to round up this section. For Urn R, the initial prior $(1/2, 1/2)$ that puts equal weight on both parameters is obviously the best description of the decision-maker's state of knowledge, irrespective of whether the decision-rule is SEU or MEU.

For Urn A this is different. An SEU decision-maker – who is the standard Bayesian decision-maker – is by definition characterized by a single belief. As objective knowledge about Urn A is not available, the question how the initial prior is determined is in general not easy to answer (Maskin 1979). Due to the perfect symmetry of this setting, however, it is clear that a Bayesian decision-maker would, according to the *principle of insufficient reason* (see for instance Gilboa 2009), hold the initial prior that assigns equal probability to $\theta = 1/3$ and $\theta = 2/3$. As a consequence, an SEU decision-maker sees no difference between Urn R and Urn A.

3 A simple example demonstrating the switch in preferences

In this section I design a simple example to illustrate the key problem in the ES setting. It involves an MEU decision-maker who initially features the usual ambiguity averse preferences. After observing one draw from the urns, however, she switches her preferences and partially exhibits ambiguity loving behavior.

3.1 Preliminaries

The example is constructed within the standard infinite horizon setting with S^∞. I will, however, only compare the betting behavior on the color of the next ball before and after a single signal realization s. That is, I will focus on the betting preferences at $t = 0$ and $t = 1$. For simplicity, I consider exclusively the two bets 1_B0 and 1_W0, where the bet 1_B0 involves a payment of 1 \$ if the color of the ball drawn is black and 0 \$ if the ball is white. The bet 1_W0 is defined similar.

The example rests on the simple three-ball-urns introduced in section 2, cf. Figure 1. Consequently, the parameter space consists of the two possible ratios of black balls in the urn, $\Theta = \{1/3, 2/3\}$. Any prior and posterior over the parameter space has the form $(\nu, 1-\nu)$ where the extreme points $(1, 0)$ and $(0, 1)$ correspond to full weight on the parameter $\theta = 1/3$ and $\theta = 2/3$, respectively.

For Urn R, both the SEU and the MEU decision-maker hold the uniform distribution as the initial prior. This is also the prior the SEU decision-maker associates with the ambiguous Urn A. The MEU decision-maker, in contrast, holds a set of initial priors regarding Urn A. For simplicity, let the set in this example be the full set of priors $\mathcal{M}_0 = \{(\nu, 1-\nu) \mid 0 \leq \nu \leq 1\}$. Let the rejection parameter α be $4/5$, which means that the MEU decision-maker only updates the initial priors with a plausibility of at least 0.8 of the maximal plausibility.

3.2 *Ambiguity aversion before observing the signal*

Let us first compare the betting preferences of the SEU and the MEU decision-maker before the signal has been observed. Irrespective of the color to bet on, the SEU is indifferent between the bet on Urn R and Urn A as for her both urns are essentially the same. The MEU decision-maker, however, has – irrespective of the color to bet on – a clear preference for betting on Urn R. The reason is basically expression (1). For bet 1_B0, where B is the favorable color, the worst scenario is that with the lowest number of black balls in the urn. Thus the MEU decision-maker rests her decision on the prior $(1, 0)$ that puts full weight on $\theta = 1/3$. The associated expected payoffs are $1/3\,\$$. A similar argument for bet 1_W0 shows that the worst prior is $(0, 1)$, again with the expected payoff $1/3\,\$$. This has to be compared to the expected payoffs for Urn R. Clearly, the prior $(1/2, 1/2)$ is for both bets associated with expected payoffs of $1/2\,\$$. Thus, the MEU decision-maker strictly prefers either bet in Urn R over Urn A. This is the well-known ambiguity averse behavior.

3.3 *Switch to ambiguity loving behavior after learning*

We now compare the betting behavior after a signal s has been observed. To make behavior comparable, we restrict to the case that the balls drawn from Urn R and Urn A have the same color. Without loss of generality, say $s = B$. This transforms the SEU decision-maker's belief from the initial prior $(1/2, 1/2)$, irrespective of the urn, into the posterior $(1/3, 2/3)$, reflecting the increased subjective probability for the scenario that the unknown ball is black.

The MEU decision-maker shares this view for Urn R, but naturally has a different take on Urn A. Here, the set of initial priors $\mathcal{M}_0$ is, according to (5), updated to the set $\mathcal{M}_1^\alpha(s)$. To recapitulate the ES procedure explained in section 2.4, the first task is to find the most plausible theory $\mu_0 \in \mathcal{M}_0$. This is clearly $(0, 1)$. The plausibility of this theory (cf. (3)) is $2/3$. With $\alpha = 4/5$, the MEU decision maker rejects all theories with a plausibility less than $4/5 \cdot 2/3$ and thus keeps the set $\mathfrak{M}_0(s) = \{(\nu, 1-\nu) \mid 0 \leq \nu \leq 2/5\}$. Finally, this set is updated to $\mathcal{M}_1^\alpha(s) = \{(\nu, 1-\nu) \mid 0 \leq \nu \leq 1/4\}$.

Let us again compare betting preferences of the SEU and the MEU decision-maker. The SEU remains, for either bet, indifferent between Urn R and Urn A. Turning to the MEU decision-maker, consider first the bet 1_W0. Urn R promises expected payoffs of $4/9\,\$$. For Urn A the worst belief is still $(0, 1)$, associated with expected payoffs of $1/3\,\$ < 4/9\,\$$. Thus, the MEU decision-maker still prefers the bet 1_W0 in Urn R over Urn A.

Most intriguing, this is different for the bet 1_B0. Urn R promises an expected payoff of 5/9 \$. For Urn A, the worst posterior in $\mathcal{M}_1^\alpha$ is $(1/4, 3/4)$. This translates into expected payoffs of 7/12 \$, which is larger than 5/9 \$. The maxmin decision-maker thus prefers the bet 1_B0 in Urn A over the same bet in Urn R. This very surprising and a clear contradiction to ambiguity averse preferences. The reason for that switch in behavior stems from the fact that in general ES reject too many theories. Here, with a rather high $\alpha = 4/5$, the critical ambiguity neutral SEU prior $(1/2, 1/2)$ is rejected. This prior is critical because with it also all 'pessimistic priors' that would give rise to ambiguity averse choices are rejected. As a consequence, the set of posteriors $\mathcal{M}_1$ only contains optimistic beliefs that give rise to ambiguity loving choices.

One could argue that this problematic betting behavior can be avoided by adequately choosing α. In the example, any $\alpha < 3/4$ would not give rise to the switch in preferences, at least not at $t = 1$. However, α was introduced by ES to be a characteristic of the decision-maker and it is thus unnatural to adjust α to the specific setting.

The next section will demonstrate that the ES anomaly is pervasive. For every $\alpha > 0$ there is a similar setting to that considered in the example for which such a disconnect in the behavior of the MEU decision-maker occurs. It is thus not possible to find a rejection parameter $0 < \alpha \leq 1$ that is not prone to the ES anomaly.

4 The general result

In this section, I first generalize the setting from urns with three balls to arbitrary symmetric settings. This provides the framework to demonstrate that the problematic characteristic of the ES updating procedure is not restricted to specific urns. Indeed, we can show that each pair of generalized urns has this property for some rejection parameter $0 < \alpha \leq 1$. Even more interesting and less obvious, for each $0 < \alpha \leq 1$ we can construct a pair of urns and a signal history s^t such that the MEU decision-maker features the switch to *ambiguity loving* behavior.

4.1 Generalized urns

The first step is to define generalizations of Urn R and Urn U. I restrict to symmetric settings in the sense that a priori both colors are interchangeable. The generalized Urn R and Urn A, I denote them by $U_R(n, k)$ and $U_A(n, k)$, respectively, have exactly $2n + k$ balls. It is known that n balls are black, n balls are white and each of the remaining k balls can either be black or white. As will become clear in what follows, the urns in the example in the previous section correspond to the case $n = k = 1$. To generalize the three balls urn example, we assume that the number of black balls within the k unknown balls in urn $U_R(n, k)$ is uniformly distributed.[4] The urn $U_A(n,k)$ is basically the same, but without information about the distribution of the k unknown balls. In either case, the parameter set is $\Theta = \{n/(2n+k), \ldots, (n+k)/(2n+k)\}$ with $\theta \in \Theta$ being the true fraction of black balls in the urn, unknown to the decision-maker. The period state space is again S with the full state space S^∞. The likelihood functions are fully specified by $l(s = B|\theta) = \theta$.

In terms of initial prior, it is clear that for $U_R(n,k)$ both SEU and MEU decision-maker have the initial prior $(1/(k+1), \ldots, 1/(k+1))$ on Θ. This is also the initial prior the SEU decision-maker holds for Urn A. In contrast, the MEU decision-maker operates with a set $\mathcal{M}_0$ of initial priors, which is by definition a subset of $\Delta(\{0, \ldots, k\})$. The natural assumption I make is that the uniform distribution is an element of the set of initial priors, $(1/(k+1), \ldots, 1/(k+1)) \in \mathcal{M}_0$.

It is convenient to focus on simple and tractable sets for $\mathcal{M}_0$. As a generalization of intervals around 1/2 in the case $k = 1$, I consider sets of the form

$$\Delta_\epsilon^{(k)} = \left\{ (\nu_0, \ldots, \nu_k) \mid \sum_i \nu_i = 1 \,,\, 0 \leq \epsilon \leq \nu_i \leq 1 - k\epsilon \leq 1 \, \forall i \;,\; \epsilon < \tfrac{1}{k+1} \right\} . \quad (7)$$

By construction, all these sets contain the uniform distribution. The full set of priors Δ is the special case $\Delta_\epsilon^{(k)}$ with $\epsilon = 0$, and for $\epsilon \to 1/(k+1)$ the set collapses to a singleton with the uniform distribution as the only element.

4.2 Theorem

I now have the toolkit to formulate the main finding. The intertemporal maxmin decision-maker in the ES sense is characterized by the parameter α that describes to what extent theories are rejected ex-post. I have shown in section 3 that also the uniform distribution, the initial prior of the SEU decision-maker, can be rejected in this process. As a consequence, all pessimistic priors are rejected as well, giving rise to ambiguity loving choices. This anomaly occurs under very general conditions. Concrete,

[4] An alternative generalization would be that of k independent coin flips. My choice, however, is more intuitive and technically simpler

Theorem 4.1. *The* ES *anomaly can be characterized by two statements.*

(i) For any pair $(n,k) \in \mathbb{N}^2$, any $\mathcal{M}_0 = \Delta_\epsilon^{(k)}$ and any bet there is a rejection parameter α and a signal history s^t such that after observing s^t the MEU *decision-maker exhibits ambiguity loving behavior, i.e. she prefers the bet in $U_A(n,k)$ over the same bet in $U_R(n,k)$.*

(ii) For any rejection parameter $\alpha \in (0,1]$ and any bet there is $(n,k) \in \mathbb{N}^2$, a set of initial prior $\mathcal{M}_0 = \Delta_\epsilon^{(k)}$ and a signal history s^t such that after observing s^t the MEU *decision maker prefers the bet in $U_A(n,k)$ over the same bet in $U_R(n,k)$.*

Proof. For both statements of the theorem I have to construct a suited signal history s^t. Without loss of generality, I consider the bet 1_B0. Accordingly, the signal history can be constructed as $s^t = (B, \ldots, B)$. The task then is to construct an appropriate t. For $\mathcal{M}_0 = \Delta_\epsilon^{(k)}$, the maximal plausible prior after observing only black balls is $(\epsilon, \ldots, \epsilon, 1 - k\epsilon)$. The ES anomaly occurs if the uniform distribution $(1/(k+1), \ldots, 1/(k+1))$ is rejected. This happens if its plausibility is lower than α times the maximal plausibility. This condition, after multiplying with $(2n+k)^t$, reads

$$\sum_{i=0}^{k} (n+i)^t \; < \; (k+1)\alpha \left(\epsilon \sum_{i=0}^{k-1} (n+i)^t + (1-k\epsilon)\,(n+k)^t \right) . \tag{8}$$

For part (i), I have to construct a suited rejection parameter $0 < \alpha \leq 1$. In the appendix I demonstrate that this is equivalent with $\alpha > \underline{\alpha}$, where $\underline{\alpha}$ is a number independent of t. I show $\underline{\alpha} < 1$ and thus the existence of a problematic rejection parameter α.

For part (ii), $\alpha \in (0,1]$ is given and the task is to construct an urn (n,k), a set of initial prior $\Delta_\epsilon^{(k)}$ and a signal history s^t that fulfills (8). With sensible choices, this reduces to finding a number of unknown balls k. It is obvious that the smaller α, the larger must be k. In the appendix I derive a condition for k of the form $k > \underline{k}(\alpha)$ that is sufficient for (8). □

5 Graphical illustration

This section is dedicated to the graphical illustration of the general findings. The first subsection 5.1 illustrates the first part of the theorem by demonstrating the existence of a problematic rejection parameter in any given setting, here $n = k = 1$. The second part of the theorem is made more concrete in subsection 5.2: A given rejection parameter that is innocent in some setting ($k = 1$) gets problematic in a higher dimensional setting ($k = 2$).

5.1 *First part of the theorem*

The first part of the theorem can be illustrated with $n = k = 1$. In Figure 2 we see the time series of non-rejected initial priors $\mathfrak{M}_0^\alpha(s^t)$ (top panel), the history of standard SEU beliefs (black line in the bottom panel) and the set of ES beliefs $\mathcal{M}_t^\alpha(s^t)$ (shaded area in the bottom panel) for different rejection parameter α. At $t = 0$, by definition $\mathfrak{M}_0 = \mathcal{M}_0$ which is chosen as the full set $\Delta_0^{(1)}$. The initial SEU prior is $(1/2, 1/2)$.

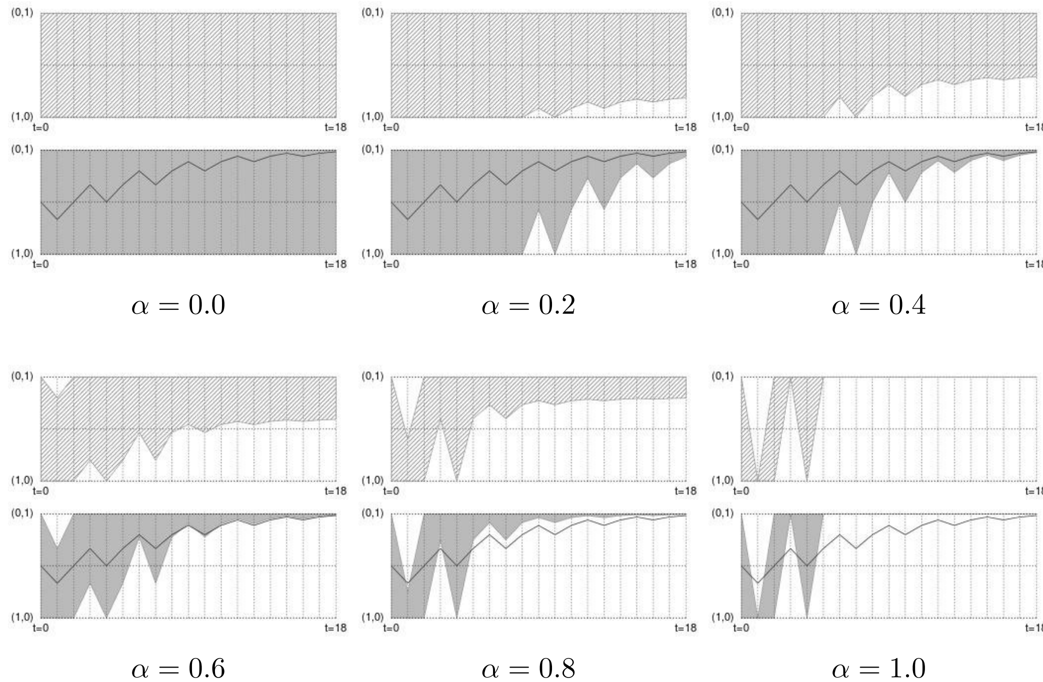


Figure 2: Illustration of the first part of the theorem. Each subfigure corresponds to a different rejection parameter α and presents the history of non-rejected initial priors $\mathfrak{M}_0^\alpha(s^t)$ (top panel), the history of standard SEU beliefs (black line in the bottom panel) and the set of ES beliefs $\mathcal{M}_t^\alpha(s^t)$ (shaded area in the bottom panel). For $\alpha > 0.5$, the ES anomaly occurs at some point in time.

The signal underlying all subfigures in figure 2 is the iteration of the sequence W, B, B (six times) and thus evidence for the theory $\theta = 2/3$, which corresponds to the belief $(0, 1)$. In the subfigures we see how this evidence is processed under different rejection parameter α. With $\alpha = 0$, which is a limiting case not permitted by ES, no theory in $\mathcal{M}_0$ is ever rejected. Thus $\mathfrak{M}_0(s^t)$ remains the full set $\mathcal{M}_0$ over the whole time (top panel). The update of the full set, however, is the full set again. This is why also $\mathcal{M}_t$ remains constant over time. The update of the SEU decision-maker, of course, is independent of α and converges to $(0, 1)$.

For increasing α more and more theories are rejected. Note that after observing the first signal, $s = W$, the theory $(1, 0)$ is the most plausible one. This is why theories in the neighborhood of $(0, 1)$ are rejected for adequately high α at time $t = 1$ (cf. top panel) but later, as $(0, 1)$ has become the most plausible theory, are element of the set $\mathfrak{M}_0$ again. After observing the first two signal realizations, $s_1 = W$ and $s_2 = B$, all theories are equally likely. This is reflected in the fact that $\mathfrak{M}_0^\alpha(s^2)$ is the full set, even for the most strict rejection parameter $\alpha = 1$. The same effect occurs at $t = 4$.

Central in this paper is the question under which conditions the SEU posterior is *not* in the set $\mathcal{M}_t$ of ES posteriors, potentially causing those problems delineated in section 3 and generally formulated in section 4.2. By definition of $\mathcal{M}_t$, this ES anomaly occurs if and only if the initial SEU prior $(1/2, 1/2)$ is not in $\mathfrak{M}_0$. We see this effect in the subfigures with $\alpha = 0.6$, $\alpha = 0.8$ and $\alpha = 1.0$. In the latter case, with the strictest rejection parameter possible, the set of non-rejected theories is either the full set (when the signal history is not conclusive) or the singleton with one of the extreme theories $(1, 0)$ or $(0, 1)$. As a consequence, also the posterior set $\mathcal{M}_t$ is either the full set or an extreme singleton. This shows the problem of the ES updating in a nutshell: The specific form of ES updating favors extreme beliefs, rather than 'smooth' beliefs around the SEU belief history. With $\alpha = 1.0$, the MEU decision-maker is positive that the unknown ball in the urn is white ($\mathcal{M}_1 = \{(1, 0)\}$) after observing the first signal realizations; after the second signal realization, she is clueless ($\mathcal{M}_2 = \Delta$); then positive that the color of the unknown ball is black ($\mathcal{M}_3 = \{(0, 1)\}$); then clueless again before remaining perfectly convinced that the unknown ball is black. This extreme form of reevaluation is due to $\alpha = 1$, but even smaller values give rise to a similar behavior.

Figure 2 thus illustrates the first part of the theorem: For a given urn, it states the existence of a signal history s^t and a rejection parameter α such that the setting described by those parameters features the ES anomaly. Figure 2 suggests that the ES anomaly does occur for $\alpha \geq 0.6$, but not for $\alpha \leq 0.4$. The proof of Theorem 4.1 in the appendix indeed shows that $\underline{\alpha} = 0.5$ is the relevant threshold for α.

This, in turn, implies that the setting $n = k = 1$ is innocent for $\alpha \leq 0.5$. Indeed, the ES anomaly does not occur even with extreme signal histories. This is demonstrated in Figure 3. Here, the rejection parameter $\alpha = 0.5$ is just small enough to ensure that the SEU posterior remains in the set of ES posteriors. Observing only black balls drawn, (0.1) is always the upper bound of the non-rejected initial priors (Figure 3, upper panel) and consequently also the upper bound of the set of posteriors (lower panel). The lower bound of the non-rejected priors $\mathfrak{M}_0$ converges to $(1/2, 1/2)$. Thus, the setting $\alpha = 0.5$ exhibits no ES anomaly.

5.2 Second part of the theorem

The second part of the theorem, however, shows that all rejection parameter $\alpha > 0$ are potentially problematic. An arbitrary $\alpha \leq 1/2$ might be innocent for the urn

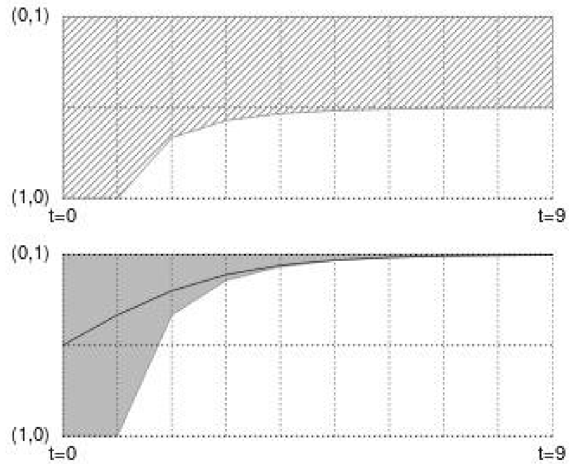


Figure 3: Illustration that $n = k = 1$ is innocent for $\alpha = 0.5$.

characterized by $k = 1$; there exists, however, a generalized urn (characterized by n and k) giving rise to the ES anomaly.

This part of the theorem is illustrated by Figure 4 with $\alpha = 1/2$, the signal history $s^t = (B, B, B, B, B)$ and $n = k = 2$.[5] As usual, the beliefs under $k = 2$ can conveniently be captured in a simplex. Each subfigure in Figure 4 corresponds to one point in time and shows the non-rejected theories $\mathfrak{M}_0$ with the uniform distribution $(1/3, 1/3, 1/3)$ marked with a black dot (top panel), SEU posteriors (black dots in the bottom panel) and set of ES posteriors (shaded area in the bottom panel).

We can see in the top panels that the uniform distribution $(1/3, 1/3, 1/3)$ is not in the set of admissible theories $\mathfrak{M}_0$ for all $t \geq 4$. As a result, the set of posteriors $\mathcal{M}_t$ does not contain the SEU update. This reflects the general theorem: A rejection parameter α, here $\alpha = 0.5$, might be innocent for certain settings, e.g. $n = k = 1$ (cf. Figure 3). It is always possible, however, to construct a generalized urn, here $n = k = 2$, such that the ES anomaly occurs after observing $s = B$ for a finite number of times.

The theorem demonstrates the problematic feature of the ES setting: It is not possible to avoid the ES anomaly when the positive rejection parameter $\alpha > 0$ ought to be *independent* of the decision-problem. In the next section, we offer two modifications of the ES framework in order to avoid the witch in ambiguity preferences.

[5]The number of unknown balls, k, determines the dimension of beliefs and is the relevant number. The number n only determines likelihoods and is thus of minor importance.

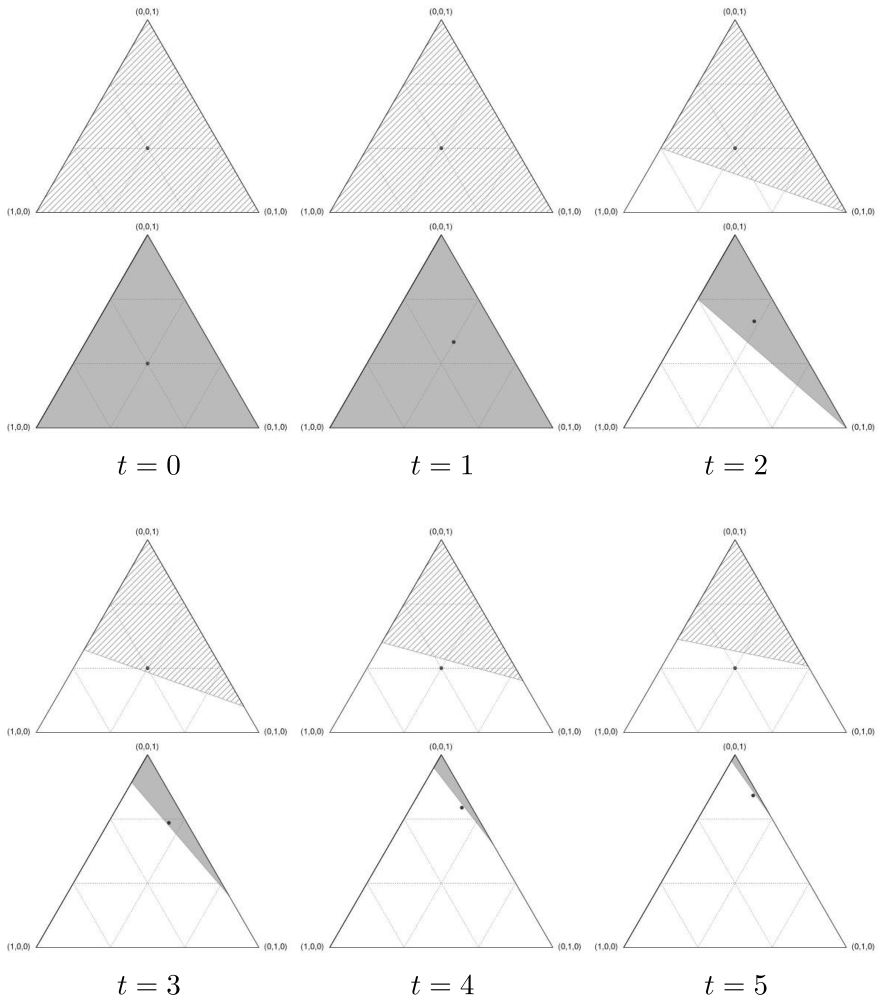


Figure 4: Illustration of the second part of the theorem with $\alpha = 0.5$ and $k = 2$. Each subfigure shows for a certain point in time the set of non-rejected initial priors $\mathfrak{M}_0^\alpha(s^t)$ (top panel), the standard SEU belief (black line in the bottom panel) and the set of ES beliefs $\mathcal{M}_t^\alpha(s^t)$ (shaded area in the bottom panel). The ES anomaly occurs at $t = 4$ and $t = 5$ when the uniform distribution is rejected.

6 Alternatives

In this final substantive section I offer two modifications of the ES framework, each of which is designed to overcome the ES anomaly. The first modification in 6.1 is basically a refinement of the ES approach and thus also involves the rejection of initial priors; the ES anomaly is avoided by defining a set of essential beliefs that are immune to rejection. The second modification, which I consider the preferable fix, has the charme of simplicity. I argue that the rejection of theories is not desirable anyway. Abstaining from the rejection of theories clearly avoids the ES anomaly. As

I will demonstrate in 6.2, simple restrictions on the set of initial priors $\mathcal{M}_0$ suffice to ensure well-behaved learning dynamics.

6.1 Refinement of the rejection of theories

The first modification of the ES framework declares certain theories as unrejectable and can thus avoid the ES anomaly. The decision-maker may feel that a certain set of theories $\mathcal{M}_0^{\text{ess}}$ is essential and thus should be immune to ex-post rejection, no matter how implausible the essential priors are in light of the signal history.

The modification of (6) is, for $\mathcal{M}_0^{\text{ess}} \neq \emptyset$,[6] given by

$$\mathfrak{M}_0^{\alpha}(s^t) = \left\{ \mu_0 \in \mathcal{M}_0 \mid \text{Plaus}(\mu_0; s^t) \geq \alpha \min_{\tilde{\mu}_0 \in \mathcal{M}_0^{\text{ess}}} \text{Plaus}(\tilde{\mu}_0; s^t) \right\} . \tag{9}$$

That is, a theory μ_0 is only rejected if it fails a maximum likelihood test against *all* theories in $\mathcal{M}_0^{\text{ess}}$. This ensures that even for $\alpha = 1$ all theories in $\mathcal{M}_0^{\text{ess}}$ are updated.

In our setting, a natural candidate for such an essential theory is the uniform distribution held by the SEU decision-maker as the prior distribution. As it might be desirable to ensure that the Bayesian update is an inner point of $\mathcal{M}_t$, we choose $\mathcal{M}_0^{\text{ess}} = \Delta^{(k)}_{\epsilon^{\text{ess}}}$ (cf. 4.1) with some $0 < \epsilon^{\text{ess}} < 1/(k+1)$.

In Figure 5 we illustrate the effect of this alternative definition on the dynamics of multiple beliefs. All parameters settings are as in Figure 2. The only difference is that the rejection of initial priors is based on (9) instead of (6).

Figure 5 shows that the ES anomaly is avoided by this alternative definition. Due to (9), the initial prior $(1/2, 1/2)$ is an element of the set of non-rejected priors for all t and all rejection parameter α (upper panels). Thus, the SEU posterior (black line in the bottom panel) is always element of set of modified ES posteriors. As the reference point for theory rejection is not the theory with the maximal plausible anymore, the rejection of theories is less strict even in cases that were not prone to the ES anomaly (compare $\alpha = 0.2$, $\alpha = 0.4$ and $\alpha = 0.6$ in Figure 5 to those in Figure 2). The convergence behavior of the modified ES framework is the same as in the initial framework: For $\alpha > 0$, the set of posteriors converges to the true distribution.

The appeal of this modified definition is that it is only a slight modification of the ES framework. It avoids the ES anomaly while preserving their general structure of reevaluation of theories. In the next subsection, I argue for an alternative modification that does not rest on the reevaluation of theories ($\alpha = 0$).

[6] If $\mathcal{M}_0^{\text{ess}} = \emptyset$ and thus no theory is regarded essential, the ES framework is unchanged. In that case the rejection of theories is defined by (6).

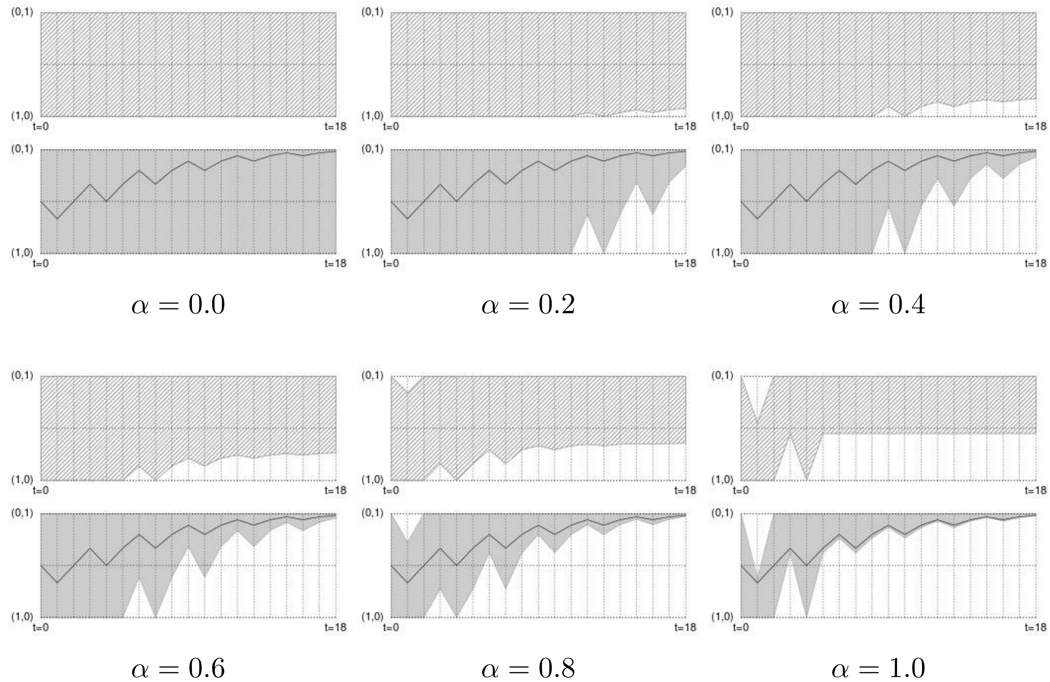


Figure 5: Refinement of the ES theory rejection with essential priors $\mathcal{M}_0^{\text{ess}}$. Each subfigure corresponds to a different rejection parameter α. In each subfigure you find the history of non-rejected initial priors $\mathfrak{M}_0^{\alpha}(s^t)$ (top panel), the history of standard SEU beliefs (black line in the bottom panel) and the set of modified ES beliefs $\mathcal{M}_t^{\alpha}(s^t)$ (shaded area in the bottom panel). The critical SEU prior, the uniform distribution, is never rejected.

6.2 A simple alternative to the rejection of theories

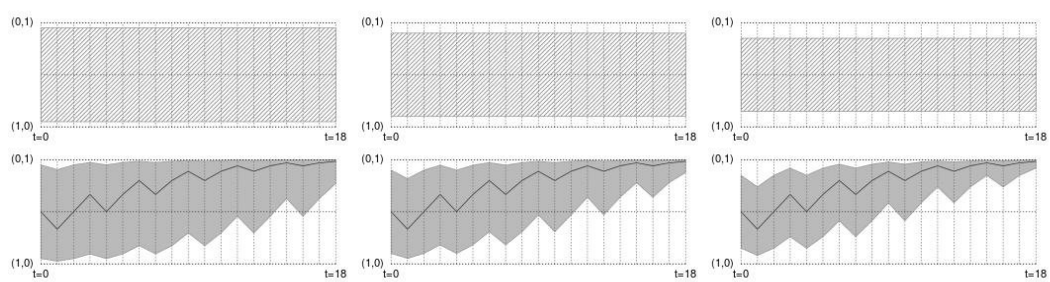
If a decision-maker holds the full set of priors $\Delta_0^{(k)}$ as the initial prior set $\mathcal{M}_0$, unrestricted updating (full bayesian updating, $\alpha = 0$) is not capable of reflecting learning as $\mathcal{M}_t = \mathcal{M}_0$ for all t, irrespective of the signal history (cf., for example, the first subfigure in Figure 2). As shown by ES, this undesirable feature can be avoided by using a theory rejection parameter $\alpha > 0$ (cf. Figure 2). This, however, in turn gives rise to a problem I coined the ES anomaly. Theorem 4.1 showed that this anomaly is a pervasive characteristic of the ES framework. The previous subsection introduced a moderate modification of this framework to circumvent this anomaly. In this subsection, I follow a different line of thought.

Reevaluation, that is the rejection of theories after observing a signal history, is not part of the standard Bayesian updating procedure. Clearly, an SEU decision-maker does *not* reevaluate its initial prior. She does not replace her initial prior by a more plausible one to update the latter. Rather, the initial prior (the uniform distribution in this paper) is the prior to be updated for all t and all conceivable signal histories. The information obtained over the learning process is reflected in the posterior and not used to reevaluate the initial prior.

I argue to keep this characteristic in the multiple prior setting. That is, I argue for defining the set of posteriors $\mathcal{M}_t$, under all conditions, as the update of the *full* initial prior set $\mathcal{M}_0$. In other words, I argue for extending the ES framework to $\alpha = 0$ and rule out all $\alpha > 0$. Reevaluation of initial priors is not necessary as the information provided by the signal history is reflected in the set of posteriors. This will become apparent below.

The price for this simple fix of the ES anomaly is that I have to avoid the trivial learning dynamics mentioned above. The solution is just the restriction $\mathcal{M}_0 \neq \Delta_0^{(k)}$. The price of this restriction, however, is low. Please recall that the axiomatization of maxmin preferences of Gilboa and Schmeidler (1989) regards the set of beliefs as an *endogenous* component of the ambiguity averse preferences. In other words, the set of beliefs, instead of reflecting objective uncertainty, reflects how strong the ambiguity averse preferences of the decision-maker are. With the inflexible and extreme maxmin rule the set of beliefs is basically the only way to express different degrees of ambiguity aversion. In that sense, the full initial prior set $\Delta_0^{(k)}$ would correspond to the most extreme uncertainty aversion possible. It seems not problematic to rule ot this extreme form of ambiguity aversion.

Due to the non-rejection of theories, one might suspect this modification $\alpha = 0$ to produce very large sets of posteriors. Figure 6, however, demonstrates that this is actually not the case. Even large prior sets narrow down substantially as the signals get more and more informative.



$\epsilon = 1/20$ $\epsilon = 2/20$ $\epsilon = 3/20$

Figure 6: Simple alternative without theory rejection. Initial priors $\mathfrak{M}_0^\alpha(s^t)$ (top panel), the history of standard SEU beliefs (black line in the bottom panel) and the set of ES beliefs $\mathcal{M}_t^\alpha(s^t)$ (shaded area in the bottom panel).

Figure 6 shows the set of non-rejected priors $\mathfrak{M}_0$, the SEU posteriors and set of ES posteriors for different initial prior sets $\mathcal{M}_0 \neq \Delta_0^{(k)}$. The set $\mathfrak{M}_0$ is, by definition of $\alpha = 0$, constant over time. The attractive features of this modification of the ES framework for applications are that (i) the procedure is simple, (ii) that the set of posteriors $\mathcal{M}_t$ follows a similar trajectory like the SEU posterior and yet (iii) the contraction of the set reflects the increased information about the true parameter θ. The belief dynamics in Figure 6 make the impression of being more "smooth" than those in Figure 5, let alone Figure 2. As this impression is further underpinned by the theoretical argument that the rejection of initial priors may be per se problematic, I argue to use this modification of the ES setting in applications.

7 Concluding discussion

The intertemporal maxmin framework of Epstein and Schneider (2007) involves the rejection of initial priors that have become implausible in light of the observed signal history. In this comment, I have demonstrated that these specific belief dynamics potentially give rise the so called ES anomaly, namely a problematic switch in ambiguity preferences. Those who apply the framework of Epstein and Schneider (2007) to model intertemporal ambiguity aversion should be aware of the potential switch in preferences and aim to avoid it.

I have offered two modifications of the ES framework. The first solution to the ES anomaly is to declare a set of priors essential and thus immune to rejection. To avoid the switch in ambiguity preferences, the uniform distribution would be such an essential prior. The second alternative, which seems the simpler and more appealing solution, abstains from the rejection of initial priors in any case and thus avoids the preference switch right from the start. As I have demonstrated, this modification leads to well-behaved belief dynamics if the set of initial priors is not the full set.

It is important to note that I have essentially focused in this comment on a reduced version of Epstein and Schneider (2007). Apart from multiple initial priors captured by the set $\mathcal{M}_0$, ES also allow for *multiple likelihoods* $\mathcal{L}$. The rejection of "theories" in ES does apply to initial priors in $\mathcal{M}_0$ *and* likelihoods in $\mathcal{L}$. The ES anomaly already occurs in the reduced setting with $\mathcal{L}$ being a singleton; a similar anomaly, however, may occur with multiple likelihoods when a standard likelihood is rejected over the course of the learning process. Future research ought to isolate the conditions for such a switch in detail. Both modifications of the ES framework presented in this paper, however, seem to be promising candidates to also fix the potential anomalies that may occur due to the reevaluation of multiple likelihoods.

Acknowledgements

I acknowledge the support of the German Science Foundation, grant no. GO 1604/2-1. I want to thank Boris Wiesenfarth for the discussions in the early stages of this paper. I am grateful to Timo Goeschl and Tobias Pfrommer for their helpful comments and suggestions.

References

Condie, S., Ganguli, J.V., 2011. Ambiguity and rational expectations equilibria. The Review of Economic Studies 78, 821–845.

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

Epstein, L.G., Schneider, M., 2003. Recursive multiple-priors. Journal of Economic Theory 113, 1–31.

Epstein, L.G., Schneider, M., 2007. Learning under ambiguity. The Review of Economic Studies 74, 1275–1303.

Garlappi, L., Uppal, R., Wang, T., 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. Review of Financial Studies 20, 41–81.

Gilboa, I., 2009. Theory of decision under uncertainty. volume 1. Cambridge university press Cambridge.

Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with non-unique prior. Journal of Mathematical Economics 18, 141–153.

Ju, N., Miao, J., 2012. Ambiguity, learning, and asset returns. Econometrica 80, 559–591.

Leippold, M., Trojani, F., Vanini, P., 2008. Learning and asset prices under ambiguous information. Review of Financial Studies 21, 2565–2597.

Maskin, E., 1979. Decision-making under ignorance with implications for social choice. Theory and Decision 11, 319–337.

Nishimura, K.G., Ozaki, H., 2007. Irreversible investment and knightian uncertainty. Journal of Economic Theory 136, 668–694.

Riedel, F., 2009. Optimal stopping with multiple priors. Econometrica 77, 857–908.

Savage, L.J., 1954. The foundations of statistics. Wiley, New York.

A Proof of the theorem

Proof. For both statements of the theorem I have to construct a suited signal history s^t. Without loss of generality, consider the bet $1_B 0$. Accordingly, I construct the signal history to generate the ES anomaly in the form of constantly observing $s = B$, $s^t = (B, \ldots, B)$. The task then is to construct an appropriate t. For $\mathcal{M}_0 = \Delta_\epsilon^{(k)}$, the maximal plausible prior after observing only black signals is $(\epsilon, \ldots, \epsilon, 1 - k\epsilon)$. The uniform distribution $(1/(k+1), \ldots, 1/(k+1))$ is rejected if its plausibility is lower than α times the maxmimal plausibility. This condition, after multiplying with $(2n+k)^t$, reads

$$\sum_{i=0}^{k} (n+i)^t \;<\; (k+1)\alpha \left(\epsilon \sum_{i=0}^{k-1} (n+i)^t + (1-k\epsilon)\,(n+k)^t \right) \;. \tag{10}$$

For part (i), I have to construct, for given n and k, a suited rejection parameter $0 < \alpha \leq 1$. The condition for the existence of such a rejection parameter is

$$\alpha > \underline{\alpha} := \frac{\sum_{i=0}^{k} (n+i)^t}{(k+1)\left(\epsilon \sum_{i=0}^{k-1} (n+i)^t + (1-k\epsilon)\,(n+k)^t\right)} \;. \tag{11}$$

The existence of a rejection parameter α giving rise to the ES anomaly is ensured if $\underline{\alpha} < 1$. Simple algebra leads to the equivalent condition

$$(1-(k+1)\epsilon) \sum_{i=0}^{k-1} (n+i)^t < (1-(k+1)\epsilon)\, k\, (n+k)^t \;. \tag{12}$$

By definition of $\Delta_\epsilon^{(k)}$, $(k+1)\epsilon < 1$. Furthermore, $\sum_{i=0}^{k-1} (n+i)^t < k(n+k)^t$. This proves $\underline{\alpha} < 1$. In particular, there are no further restrictions on t. This implies that for every urn there is a rejection parameter α such that the ambiguity loving behavior occurs already after observing *one* signal, $t = 1$.

For part (ii), $\alpha \in (0, 1]$ is given and I have to construct an urn (n, k), a set of initial prior $\Delta_\epsilon^{(k)}$ and signal history s^t that fulfills (8). As will become clear, it is helpful to choose $\epsilon = \frac{1}{k(k+1)\alpha}$, for which $\epsilon < \frac{1}{k+1}$ if $k > \frac{1}{\alpha}$. With that, condition (8) reads

$$\sum_{i=0}^{k} (n+i)^t < \tfrac{1}{k} \sum_{i=0}^{k-1} (n+i)^t + ((k+1)\alpha - 1)\,(n+k)^t \;. \tag{13}$$

Sufficient for this, by neglecting the first positive expression on the right hand side, is

$$\sum_{i=0}^{k} \left(\frac{n+i}{n+k} \right)^t < (k+1)\alpha - 1 \;. \tag{14}$$

As $((n+i)/(n+k))^t$ tends to 0 for $t \to \infty$ for all $0 \leq i < k$, there is a t such that $((n+i)/(n+k))^t < 1/k$ for all $0 \leq i < k$. Thus, a sufficient condition for the existence of a parameter k giving rise to the ESanomaly is the condition $1 + k \cdot \frac{1}{k} < (k+1)\alpha - 1$. This is equivalent with

$$k > \underline{k}(\alpha) = \frac{3}{\alpha} - 1 \; . \tag{15}$$

In particular, under this condition also $k > 1/\alpha$ and thus the ϵ-value I defined above is actually feasible.

I have shown the existence of a problematic urn for all rejection parameter α. It is intuitive that smaller α make higher k necessary. It is interesting that there are no restrictions on n, the number of known black and white balls, respectively. □