UNAWARENESS, PRIORS AND POSTERIORS

SALVATORE MODICA *

Abstract. This note contains first thoughts on awareness of unawareness in a simple dynamic context where a decision situation is repeated over time. The main consequence of increasing awareness is that the model the decision maker uses, and the prior which it contains, becomes richer over time. The decision maker is prepared to this change, and we show that if a projection-consistency axiom is satisfied unawareness does not affect the value of her estimate of a payoff-relevant conditional probability (although it may weaken confidence in such estimate). Probability-zero events however pose a challenge to this axiom, and if that fails, even estimate values will be different if the decision maker takes unawareness into account. In examining evolution of knowledge about relevant variable through time, we distinguish between transition from uncertainty to certainty and from unawareness to certainty directly, and argue that new knowledge may cause posteriors to jump more if it is also new awareness. Some preliminary considerations on convergence of estimates are included.

Keywords: Awareness of Unawareness, Model Uncertainty.
Mathematical Subject Classification: 91B06, 91B44, 62C10.
Journal of Economic Literature Classification: D83, C11.

1. Introduction

Decisions often depend on unobserved variables, and a distinction has been made between those whose relevance the subject is aware of and those which are 'totally out of her mind'. The concept of awareness was formally introduced in Modica–Rustichini [15] where it was defined as the union of certainty ("I know") and 'conscious uncertainty' ("I do not know but I know I do not know"). The literature has progressed significantly from [15] to Heifetz–Meier–Schipper [9], and we briefly report on that at the end of this section. Two issues that are still largely unexplored are *time* and *awareness of unawareness*, the underlying question being, Does the distinction between uncertainty and unawareness have a bearing on decisions? In particular, If the decision maker is 'aware of unawareness', what should she do? This note presents first thoughts on these matters.

In a simple setting where a decision situation occurs repeatedly *we focus on the probability judgments* on which decisions are based. We ask how these judgments might be affected by the fact that awareness of relevant aspects/variables of interest typically increases through time, this evolution being experienced by, hence known to, the decision maker. In our model the latter is interested at t in the value of a binary variable —one may think of

* November 2007. Address: Facoltà di Economia; Viale delle Scienze; 90128 Palermo; Italy. e–mail: modica@unipa.it.

an asset being high or low— at $t+1$. We show that under an 'unawareness projection' condition, the point estimate of the relevant conditional probability is unaffected by awareness of unawareness. The result is presented and critically discussed in section 3. Obviously, it is reasonable to think that the dm's *confidence* in this estimate is lower the less complete she thinks her current model is; this issue is not covered by this paper.

With a dynamic model in hand we also give a first look at the learning process, where increasing are both the set of variables which the subject is aware of and the set of variables which she actually observes. Indeed, if we follow the state of the subject's knowledge about a single variable in the course of time we may see it passing from unawareness to awareness and then —perhaps at the same time— to certainty. Accordingly, we will look at the two transitions: unawareness to awareness in general; and the particular transition from unawareness directly to certainty (the latter is quite common, for often one becomes aware of something by 'involuntarily' observing its effects). Concerning the former, the moment the individual becomes aware of the relevance of facts which had previously escaped him, a different view of the situation —past and future— materializes. We will trace the effects of this on subjective probabilities and see how *priors* change on these occasions. For priors we offer an explicit unawareness–based rationalization of phenomena —unawareness and changing priors— which are informally widely discussed in the statistics literature. [1] The direct transition form unawareness to certainty must be contrasted with that from (conscious) uncertainty to certainty: the first time a variable becomes observed, can it make a difference on *posteriors* whether the subject was previously aware of it or not? We argue that new knowledge which is also new awareness may have a stronger impact because the latter often gives a better understanding of the past.

Related Unawareness Literature. An extensive bibliography on unawareness has been collected by Burkhard Schipper [21]; we only mention a few papers more or less directly related to the topic of the present note. Probabilistic beliefs are introduced in Heifetz–Meier–Schipper [9] and Sadzik [20], the former also dealing with Bayesian games. The concept of awareness of unawareness is tackled, in the context of syntactic models of higher-order logic, by Ågotnes-Alechina [1], Board-Chung [2], Halpern-Regô [7] and Sillari [22]. Time is introduced in analysis of solution concepts for dynamic games in Feinberg [6], Li [14], and Regô-Halpern [18].

[1] For instance, Hill [10] writes: "[...] one makes an initial specification of model and prior, but certainly one knows that this does not fully represent his background knowledge, much less the truth, and consequently when the data strongly suggest a broadening, [...] one must feel completely open to respecify both the original model and the prior distribution. This may seem to be too adhoc, but I believe it is the only sensible way to do statistics. [...] The only problem is the difficulty in realistically specifying a "prior" distribution for parameters that one hadn't explicitly recognized before seeing the data." And Poirier [17], who calls these 'post–data priors', writes: "I believe such priors exist before the data, but that the difficulties of formulating priors cover them up. The data encounter that provokes the window [i.e. model] switch awakens the researcher [...] to rediscover this latent prior."

Related Economics and Decision-Theoretic Literature. The broad Economics issue to which the present discussion belongs is that of *model uncertainty*. It is a deep area of growing importance — the topic was at the center of Hansen's 2007 Richard Ely Lecture, see [8], and at the heart of Weitzman's explanation of the equity premium puzzle, see [23]—, in the context of which this note only makes a couple of marginal points, as we explain in this paragraph. In a more specific Decision-Theoretic context, Kilbanoff-Marinacci-Mukerji [12] axiomatize recursive preferences in a dynamic context (extending their atemporal results in [11]) by way of a functional involving probabilities on probabilities, or 'second order probabilities' as they call them, which are also the central object of the present paper.

The difference between model uncertainty with and without unawareness, and thus between the question we ask and the ones which are discussed in the literature to which we are referring, lies in the order of uncertainty. In that literature uncertainty concerns the models the agent can describe, in particular the possible probabilities underlying them. In the present setting the agent has no uncertainty about the model she can describe, in the sense that she is able to assess uniquely probabilities governing that model; on the other hand she is uncertain about what other models, which she is unable to conceive, might in fact be the true ones. So our question might be phrased thus: Beyond the uncertainty the agent may have on the models she can describe, does the higher layer of uncertainty, about models which she cannot describe, have a further impact on her behaviour?

In the sequel of the paper, after describing the setup in the next section, we discuss priors and estimation in section 3, and changes of the subject's model following occurrence of new awareness in section 4. A concluding comment is added in section 5.

2. The setup

Process and Agent. $\sigma_t \in \Sigma$ is the state of the world at $t \in \mathbb{T} = \{1, 2, \dots\}$, which evolves as a time–homogeneous Markov process; and $y_t \in Y = \{0, 1\}$ is a high–or–low dividend of an asset, whose value at $t+1$ depends on σ_t and y_t and is conditionally independent of σ_{t+1}. That is, with P denoting the probability on $(Y \times \Sigma)^{\mathbb{T}}$ governing the process, we take

$$\mathsf{P}\left(y_{t+1}, \sigma_{t+1} \mid (y_s, \sigma_s)_{s \leqslant t}\right) = \mathsf{P}\left(y_{t+1} \mid y_t, \sigma_t\right) \cdot \mathsf{P}\left(\sigma_{t+1} \mid \sigma_t\right).$$

At each t the decision maker is interested in the probability that $y_{t+1} = 1$ conditional on her information —because, say, she bets at t on the value of y_{t+1}.

The set $\boldsymbol{\Sigma}$. There is a countable set $\{\phi_1, \phi_2, \dots\}$ of 'facts' by which each state can be fully described; that is, we take

$$\Sigma = \{0, 1\}^{\{\phi_1, \phi_2, \dots\}}$$

so that at each t, σ_t is a sequence of zeros and ones corresponding to which facts are true and false. For more compact notation we set

$$\Phi = \{y, \phi_1, \phi_2, \dots\} \quad \text{and} \quad \Omega = \{0, 1\}^{\Phi}$$

so that the value of the asset is included in the description of the state, and P is on $\Omega^{\mathbb{T}}$. We may write $\omega = (y, \sigma)$, with $\sigma \in \{0,1\}^{\{\phi_1,\phi_2,\dots\}}$.

Agent Information. For each t, this is a σ–field $\mathcal{F}_t$ in $\Omega^{\mathbb{T}}$, to describe which we define, for $L \subseteq \Phi$:

- ω^L as the restriction of $\omega \in \Omega$ to L;
- $\tilde{\omega}^L_t \colon \Omega^{\mathbb{T}} \to \{0,1\}^L$ by $(\omega_1, \omega_2 \dots) \mapsto \omega^L_t$; and
- for $\phi \in \Phi$, $\tilde{\phi}_t = \tilde{\omega}^{\{\phi\}}_t$, so that $\tilde{\phi}_t((\omega_1, \omega_2 \dots)) = \omega_t(\phi)$.

The function $\tilde{\omega}^L_t$ just says whether the facts in L are true or false at t along a path; and $\tilde{\phi}_t$ does the same for the single ϕ.[2]

At each t the agent observes the value of the asset y_t, plus *some* other coordinates of ω_t, i.e. some facts; we let $L_t \subseteq \Phi$ be the set of coordinates observed at t. She does not forget past information (e.g. if she observes at s that ϕ is true, then at $t > s$ she remembers that ϕ was true at s); and $L_s \subseteq L_t$ if $s < t$ (if she observes the value of ϕ at s then she will be able to observe it at any $t > s$). Then, for some $L_t \subseteq \Phi$ and $t_\phi \leqslant t$, $\phi \in L_t$ it is

$$\mathcal{F}_t = \sigma(\{\tilde{\phi}_s : \phi \in L_t,\, t_\phi \leqslant s \leqslant t\}),$$

where $\sigma(\cdot)$ denotes here the sigma field generated by $(\cdot)$; with t_ϕ denoting the first time fact ϕ is observed, $\mathcal{F}_t$ describes which coordinates the agent observes at t and since when (for example we are assuming that $y \in L_t$ for each t, i.e. $t_y = 0$). Writing $\phi_s \in \mathcal{F}_t$ (with $s \leqslant t$) will have the obvious meaning.

A superset of L_t, denoted by A_t and also increasing with t, will represent the set of variables whose relevance the subject is aware of —those which she can list ‘by their names’. Those in $A_t \setminus L_t$ are not observed though present to the subject’s mind.

Estimation Goal. In the notation just introduced, the subject is interested in $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \mathcal{F}_t]$, which she has to estimate. The process law P is unknown to her.

A benchmark case is that of full awareness, $A_t = \Phi$, with lack of knowledge, $L_t \subset \Phi$ increasing through time, which is a hidden Markov model generalized to observing a more and more precise signal about unobserved variables, in the infinite–dimensional binary regression setting. In this situation the Bayesian subject would start with a prior μ on $\Omega^{\mathbb{T}}$ (which makes sense because $A_0 = \Phi$), update it to $\mu(\cdot \mid \mathcal{F}_t)$ and estimate the probability of $y_{t+1} = 1$ given the data as $\mu[\tilde{y}_{t+1} = 1 \mid \mathcal{F}_t]$.

Examples. Here are some simple examples to which we come back at the end of section 4. The first three are simple situations where the $\{\sigma_t\}$ process is iid, and we assume that the agent begins by observing just the value of y and is aware of nothing which may influence its evolution besides past values of y itself.

[2]Here the subscript obviously refers to time; to avoid confusion with the index of ϕ in the set $\{\phi_1, \phi_2, \dots\}$, in the sequel subscripts will always indicate time, and different facts will be denoted by different letters, like ϕ, ψ etc.

Example 1. The value of y at $t+1$ depends deterministically on the value of a certain $\phi \in \Phi \setminus \{y\}$ at t:

$$\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \tilde{\omega}_t] = \begin{cases} 1 & \text{if } \tilde{\phi}_t = 1 \\ 0 & \text{if } \tilde{\phi}_t = 0 \end{cases},$$

and $\mathsf{P}[\tilde{\phi}_t = 0] = \mathsf{P}[\tilde{\phi}_t = 1] = 1/2$ for all $t \geqslant 0$ (where t is time).

Example 2. Here ϕ_t triggers change in y_{t+1}:

$$\mathsf{P}[\tilde{y}_{t+1} = y_t \mid \tilde{\omega}_t] = \begin{cases} 1 & \text{if } \tilde{\phi}_t = 0 \\ 0 & \text{if } \tilde{\phi}_t = 1 \end{cases},$$

with $\mathsf{P}[\tilde{\phi}_t = 0] = \mathsf{P}[\tilde{\phi}_t = 1] = 1/2$ for $t \geqslant 0$ as before. Here the value of y_{t+1} depends also on y_t.

Example 3. In this variant ϕ 'acts' only if on a pair with another $\psi \in \Phi$:

$$\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \tilde{\omega}_t] = \begin{cases} 1 & \text{if } \tilde{\phi}_t = \tilde{\psi}_t \\ 0 & \text{if } \tilde{\phi}_t \neq \tilde{\psi}_t \end{cases},$$

and $\mathsf{P}[\tilde{\phi}_t = a, \tilde{\psi}_t = b] = 1/4$ for any $a, b = 0, 1$ and $t \geqslant 0$. In this case $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \tilde{y}_t, \tilde{\phi}_t] = 1/2$ for any value of $(\tilde{y}_t, \tilde{\phi}_t)$.

In each of these cases what the agent observes are repetitions of Bernoulli trials with probability 1/2 (which she can confirm if observations are repeated for long enough). How should her estimates be influenced if she suspects that any of the three above underlying processes may govern evolution of y?

Example 4 (Markov dynamics). We sketch here a 'minimal' non–trivial dynamics. The subject is aware of two facts ϕ and ψ, and only the first is observed; ϕ is thought to be negatively autocorrelated, ψ positively; and both are thought to have positive influence on y. More uncertainty concerns their joint effect, in that the subject thinks that: either both $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \tilde{\psi}_t = 0, \tilde{\phi}_t]$ and $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \tilde{\psi}_t = 1, \tilde{\phi}_t]$ depend significantly on the value of $\tilde{\phi}_t$; or the former does but the latter does not, and it is high; in other words perhaps $\psi = 1$ has an overwhelming influence, perhaps not. Here the more the observed relative frequency of the transition from y_t to $y_{t+1} = y_t$ depends on the value of y_t, being higher when $y_t = 1$, the more the hypothesis of large impact of $\psi = 1$ is favored.

Related Statistical Literature. Two strands of literature treat special cases of this model: that on nonparametric binary regression and that on hidden Markov models. In both cases the subject is aware of everything from the beginning, $A_t = \Phi$ all t.

The first, cfr. Diaconis–Freedman [4, 5], deals with the case where $L_t = \Phi$ at all t, y_t is independent of y_s, $s < t$ and the σ_t are independent, uniformly distributed:

$$\mathsf{P}(\omega_{t+1} \mid \{\omega_s\}_{s \leqslant t}) = \mathsf{P}(y_{t+1} \mid \sigma_t) \cdot \lambda(\sigma_{t+1}), \tag{1}$$

with λ uniform. Writing $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \sigma_t] = f(\sigma_t)$, they study —and under quite general conditions demonstrate— convergence to f of the posteriors, starting with a prior on f of the form $\sum_{k=0}^{\infty} w_k \pi_k / \sum_{k=0}^{\infty} w_k$, where π_k is supported on the class of functions f which depend only on the first k coordinates.

Hidden Markov models (HMM), cf. e.g. the recent book by Cappé-Moulines-Ryden [3], have $L_t = L \subsetneq \Phi$ constant over time, and letting $H = \Phi \setminus L$ have P of the form

$$\mathsf{P}(\omega_{t+1} \mid \{\omega_s\}_{s \leqslant t}) = \mathsf{P}(\omega_{t+1}^L \mid \omega_t^H) \cdot \mathsf{P}(\omega_{t+1}^H \mid \omega_t^H).$$

The purpose there is to estimate the current value of unobserved variables (we are interested in estimating the effect of unobserved variables on future observed ones).

It is to be stressed that when $L \neq \Phi$ the observed process is not Markov: $\mathsf{P}(\omega_{t+1}^L \mid \{\omega_s^L\}_{s \leqslant t}) \neq \mathsf{P}(\omega_{t+1}^L \mid \omega_t^L)$. The simplest example to see this is the following: there is an unobserved variable taking three possible values a, b, c which evolves deterministically: $a \to b \to c \to a$ (so it is Markov) and the observed y is with $y(a) = y(b) = 1$ and $y(c) = 0$; then y is also deterministic but not Markov (given $y_t = 1$, the value of y_{t+1} depends on y_{t-1}).

3. Unawareness and Priors

Awareness States. At time zero our individual can list only a set $A_0 \subseteq \Phi$ of facts which may be related to the evolution of y —the initial awareness set—, and she thinks that this A_0 may not be exhaustive (recall $A_0 \ni y$).

Being aware of the relevance of the facts in A_0 means having some idea about their joint evolution, and, in the context of this Bayesian model, *we formalize this as being able to assess a prior $\alpha(0)$ on the paths on $\{0,1\}^{A_0}$*. Then, from an ex–ante point of view, being aware of k new facts, that is forming an idea about their interaction with those already in the picture, can be formalized as specifying a prior $\alpha(k)$ on the paths on $\{0,1\}^{A_0 \cup \{\ell_1,\ldots,\ell_k\}}$, where $\{\ell_1, \ldots, \ell_k\}$ is just a set of meaningless 'labels' —it is the specification of $\alpha(k)$ which gives them meaning. At time zero uncertainty concerns k and the meaning these labels can acquire, that is, about k and $\alpha(k)$; and for each $\alpha(k)$, about the possible realizations of a still richer prior $\alpha(k+k')$ on the spaces of paths on $\{0,1\}^{A_0 \cup \{\ell_1,\ldots,\ell_{k+k'}\}}$, for some $k' > 0$. The $\alpha(k)$'s are interpreted as *awareness states*. Starting from $\alpha(0)$, realization of an $\alpha(k)$ occurs when k new facts are conceived which enter the picture as specified in $\alpha(k)$; the interpretation is analogous for subsequent steps. Before writing all this more precisely we now describe the construction in a concrete situation.

Illustration. In general, estimation is complicated by the fact that as we have seen the observation process is not Markov; and this is why we have to speak of laws on the space of paths. In this illustration we eliminate this problem to ease exposition by assuming that the $\{\sigma_t\}$ process is independent:

$$\mathsf{P}(\omega_{t+1} \mid \{\omega_s\}_{s \leqslant t}) = \mathsf{P}(y_{t+1} \mid \omega_t) \cdot \mathsf{P}(\sigma_{t+1}). \tag{2}$$

It is a slightly generalized version of the Diaconis–Freedman setting, in that y_{t+1} may depend on y_t besides σ_t. Then

$$\begin{aligned}\mathsf{P}(y_{t+1} \mid \mathcal{F}_t) &= \int_{\{0,1\}^{\Phi\setminus L_t}} \mathsf{P}(y_{t+1} \mid \mathcal{F}_t, \tilde{\omega}_t^{\Phi\setminus L_t})\, \mathsf{P}(d\,\omega_t^{\Phi\setminus L_t} \mid \mathcal{F}_t) \\ &= \int_{\{0,1\}^{\Phi\setminus L_t}} \mathsf{P}(y_{t+1} \mid \tilde{\omega}_t)\, \mathsf{P}(d\,\omega_t^{\Phi\setminus L_t} \mid \tilde{\omega}_t^{L_t}) = \mathsf{P}(y_{t+1} \mid \tilde{\omega}_t^{L_t}).\end{aligned}$$

Thus in this case estimating $\mathsf{P}[\tilde{y}_{t+1} \mid \mathcal{F}_t]$ amounts to estimating the probability of success $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \omega_t^{L_t}]$ for each $\omega^{L_t} \in \{0,1\}^{L_t}$ in a Bernoulli scheme, as in Diaconis–Freedman [4, 5].

Suppose now that $A_0 = \{y\}$. In this context an $\alpha(0)$ specifies what the subject thinks of the starting value of y and of the dependence of y_{t+1} on y_t, hence it is determined by a prior on the initial value of y and an estimated two–by–two transition matrix linking y_t and y_{t+1}.[3] Taking $A_0 = \{y\}$ means that she can think of nothing which may influence y (besides itself); still, she can imagine that there are other forces at play; for example, she may imagine that there is *one* fact ℓ, $k = 1$, on which y depends ($\alpha(1)$ here links y_{t+1} to y_t and ℓ_t); and she may think she might judge this fact to be 'crucial', or on the contrary not very influential; that is, at time zero she will have a prior on the possible forms $\alpha(1)$ will take. Similarly, she can think of the possibility of becoming aware of *two* relevant facts, which as of time zero may have any kind of joint influence on y; and this means having a prior on the possible future priors $\alpha(2)$, where each $\alpha(2)$ is an estimate of the dependence of y_{t+1} on y_t, ℓ_{1t} and ℓ_{2t}. And so on. Next: the discovery of e.g. a 'crucial' factor related to y may influence the (subjective) probability of discovering further crucial elements —she may think that after a crucial discovery it will be less likely that she will make another important one, or perhaps she may think the opposite, that further important findings are positively correlated with the first. In other words, each realization of $\alpha(1)$ may influence the possible evolution of the subject's ideas on the process; for example for different $\alpha(1)$'s she will have possibly different priors on subsequent realizations of $\alpha(2)$ (and $\alpha(3), \alpha(4)$, etc.). The same applies to any k and $k + k'$.

Priors on priors. To formalize let $\mathbb{N}_0 = \emptyset$, $\mathbb{N}_n = \{1, 2, \ldots, n\}$, $n \in \mathbb{N}$, and $\mathbb{N}_\infty = \mathbb{N}$; define a set of possible 'formal completions' of A_0 to a candidate Φ–set by letting, for $0 \leqslant k \leqslant \infty$, $\Phi(k) = A_0 \cup \{\ell_i \colon i \in \mathbb{N}_k\}$; and let $\Omega(k) = \{0,1\}^{\Phi(k)}$ and $\mathcal{A}(k) = \mathsf{M}\left(\Omega(k)^{\mathbb{T}}\right)$ —the set of probabilities on $\Omega(k)^{\mathbb{T}}$. Elements $\mathcal{A}(k)$ are the awareness states.

Then the subject's view at time zero of the possible evolution of her awareness, starting with $A_0 \subseteq \Phi$ and $\alpha(0) \in \mathcal{A}(0)$, is described by: *a prior*

[3] In this case there are iid repetitions: every time y_t is, say, equal to 0 the transition to y_{t+1} is governed by the same unknown probability $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \tilde{y}_t = 0]$, which has to be estimated. Here the subject will assess a non-degenerate prior on the value of this probability, which she then updates as data come in. The estimate in the text should be thought as the average of this subjective distribution (this is the usual Bayes estimate, arising from a quadratic loss function).

$\pi(\alpha(0)) \in \mathsf{M}\big(\cup_{k>0}\mathcal{A}(k)\big)$, *and for each* $\alpha(k) \in \mathcal{A}(k)$ *a prior* $\pi(\alpha(k)) \in \mathsf{M}\big(\cup_{k'>0}\mathcal{A}(k+k')\big)$.

Remarks. (1) $\pi(\alpha(k))$ also contemplates the possibility of being aware of *no* new facts, event represented by all priors in $\cup_{k'>0}\mathcal{A}(k+k')$ which give no role to the coordinates in $\Phi(k+k') \setminus \Phi(k)$.

(2) Define the expected belief in $\mathcal{A}(k+k')$ determined by a given $\alpha(k) \in \mathcal{A}(k)$, denoted $\overline{\alpha}(k+k';\alpha(k))$, by

$$\overline{\alpha}(k+k';\alpha(k))(X) = \pi(\mathcal{A}(k+k'))^{-1}\int_{\mathcal{A}(k+k')} \alpha(X)\,\pi(\alpha(k))(d\alpha),\ X \subseteq \Omega(k+k')^{\mathbb{T}}.$$

The 'looking forward' restriction: for $X \subseteq \Omega(k+k'+k'')^{\mathbb{T}}$

$$\overline{\alpha}(k+k'+k'';\alpha(k))(X) = \int_{\mathcal{A}(k+k')} \overline{\alpha}(k+k'+k'';\alpha)(X)\,\pi(\alpha(k))(d\alpha)$$

would then imply that $\pi(\alpha(k))$ need only be specified on $\mathcal{A}(k+1)$.

(3) The realization of awareness state $\alpha(k)$ coincides with the subject becoming aware of k new facts; if this occurs at time t, at that time the labels $\ell_1,\dots\ell_k$ become 'real' ϕ's in Φ, and the awareness set A_0 is enlarged to a corresponding $A_t \subseteq \Phi$.

A Consistency Axiom. We will state an axiom concerning the subject's 'internal' ex-ante consistency, analogous to the projective system of priors of Heifetz-Meier-Schipper [9]. The idea is that if the subject thought that the next awareness state would occur with a different average prior than current on the variables she is already aware of, she would start using that today. This conflict can be avoided by imposing the following 'looking backwards' restriction. Let the marginal of $\overline{\alpha}(k+k';\alpha(k))$ on $\Omega(k)^{\mathbb{T}}$ be defined in the natural way by identifying a path $x \in \Omega(k)^{\mathbb{T}}$ with the set of paths $\{z \in \Omega(k+k')^{\mathbb{T}} \colon \forall t\ \forall i \leqslant k\ z_t(\ell_i) = x_t(\ell_i)\}$ in $\Omega(k+k')^{\mathbb{T}}$; denoting this marginal by $\overline{\alpha}(k+k';\alpha(k))|_{\Omega(k)^{\mathbb{T}}}$, then the axiom requires that it be the same as $\alpha(k)$:

Projection Consistency:

$$\overline{\alpha}(k+k';\alpha(k))|_{\Omega(k)^{\mathbb{T}}} = \alpha(k)\,. \tag{3}$$

Remarks. (1) Notice the case of infinite A_0, or more generally of an $\alpha(\infty) \in \mathcal{A}(\infty)$ realized at some later time. In this case since $\mathcal{A}(\infty+k') = \mathcal{A}(\infty)$ we get a map taking $\alpha(\infty) \in \mathcal{A}(\infty)$ to $\pi(\alpha(\infty)) \in \mathsf{M}(\mathcal{A}(\infty))$, i.e. a Markov process on $\mathcal{A}(\infty)$ (where the index is not a time index: it may take any time to get from a state to another). Also, given $\alpha(\infty)$, the average $\overline{\alpha}(\infty;\alpha(\infty))$ defined as before by $\overline{\alpha}(\infty;\alpha(\infty))(X) = \int_{\mathcal{A}(\infty)} \alpha(X)\,\pi(\alpha(\infty))(d\alpha)$ for $X \in \Omega(\infty)^{\mathbb{T}}$ is again in $\mathcal{A}(\infty)$; then (3) says that $\alpha(\infty)$ is stationary, for it reads $\overline{\alpha}(\infty;\alpha(\infty)) = \alpha(\infty)$, that is

$$\alpha(\infty)(X) = \int_{\mathcal{A}(\infty)} \alpha(X)\,\pi(\alpha(\infty))(d\alpha), \quad X \subseteq \Omega(\infty)^{\mathbb{T}}.$$

(2) The notation being already overburdened the following point has been postponed: any $\alpha(k)$, once realized and as long as 'in use' (i.e. before the next awareness state materializes), is updated each period in the light of the

incoming data; in other words it becomes an $\alpha(k)[\cdot \mid \tilde{\omega}_t^{L_t}]$, call it $\alpha_t(k)$. The consistency restriction (3) refers to the latter; that is, the $\alpha(k)$ appearing in that equation should be replaced with $\alpha_t(k)$.

Given an awareness state $\alpha_t(k)$ at time t, the subject has average beliefs $\overline{\alpha}(k+k'; \alpha_t(k))$, $k' > 0$ on larger potentially relevant spaces; what use should she make of them in estimating future values of y? The following result shows that Projection Consistency implies that at least as far as the value of her point estimate of the probability that $y_{t+1} = 1$ is concerned, the answer is *None* —she can ignore them.

Proposition. *Suppose that the subjective probability of $y_{t+1} = 1$ at t given current awareness is arrived at by averaging over $k' > 0$ the values obtained via the expected beliefs $\overline{\alpha}(k+k'; \alpha_t(k))$. Under Consistency* (3)*, the resulting probability is the same as that obtained directly via $\alpha_t(k)$.*

Proof. Fix any k', and let $A = \Phi(k)$ and $U = \{\ell_{k+1}, \dots, \ell_{k+k'}\}$. The estimate based on $\beta \equiv \overline{\alpha}(k+k'; \alpha_t(k))$ given current awareness is

$$\int \beta\left(y_{t+1} \mid \{\omega_s^A\}_{s \leqslant t}, \{\omega_s^U\}_{s \leqslant t}\right) \beta\left(d\{\omega_s^U\}_{s \leqslant t} \mid \{\omega_s^A\}_{s \leqslant t}\right),$$

that is $\beta(y_{t+1} \mid \{\omega_s^A\}_{s \leqslant t})$; but axiom (3) implies that this is the same as the value obtained via $\alpha_t(k)$. □

The result states that he value of an estimate based on current awareness state should be not altered by mental constructions with no 'real' counterparts. If this is so, estimation proceeds as follows: use the initial prior $\alpha(0) \in \mathcal{A}(0)$ (on the paths on $\{0,1\}^{A_0}$), and update it to get at each t the estimate $\alpha(0)[\tilde{y}_{t+1} \mid \tilde{\omega}_t^{L_t}]$, until a new awareness state occurs, say at t_k, in the form of a prior $\alpha(k) \in \mathcal{A}(k)$, at which moment start again by using and updating this one to get the estimate $\alpha(k)[\tilde{y}_{t+1} \mid \mathcal{F}_t]$ for $t \geqslant t_k$ until another awareness state occurs; and so on. In other words, under (3) the π system is not needed as far as estimation is concerned.

We stress that the above proposition does not imply that the π is useless altogether, for assessing one's view of the possible evolution of the current awareness state determines the *degree of confidence* in current estimates. Indeed, it should be possible to use the present framework to argue that if the decision maker has a finite amount of money to bet over time, she will shift investments to later dates the more the larger the amount of relevant information she feels she is currently unaware of. Of course, even for estimation purposes the π system is necessary if condition (3) is violated, and this *can* happen as we discuss next.

Awareness of Unawareness and Events of Probability Zero. As pointed out by a referee to whom I remain in debt, there is an important qualification to the above axiom and proposition. The critical point concerns events which under the current model have zero probability. To visualize consider an example suggested by the referee, about the possibility that some individual correctly guesses a sequence of random numbers. To aid the discussion let y denote guessing right/wrong, and assume that our decision maker, being aware of nothing that may influence the process, attaches probability zero to any y-path containing only right guesses (perhaps after a short learning span). Let now ℓ be a label such that an α including ℓ (besides y), say $\alpha(\ell)$, gives positive probability to right guessing. Then axiom (3) prescribes that ex-ante, when the agent does not suspect of anything which might possibly give rise to a non-zero probability of right guessing, she should give zero π-probability to $\alpha(\ell)$ —which, incidentally, would imply that she should accept to bet any amount at any odds against right guessing. But this does not seem compelling; indeed, it does not sound *irrational* to be unwilling to enter those bets on the basis of an assignment of positive probability to unconceived worlds where events which have zero probability in the perceived world have positive probability. If the agent makes such probability judgments in violation of axiom (3), her subjective π system affects estimated probabilities, and concretely the bets she is willing to enter.

Whether the decision maker assigns them probability zero or not, unconceived worlds contradicting her initial view are sometimes the real ones, so that evidence accumulates which challenges the current view of reality. In the guessing example, the predictor might know the algorithm generating the sequence; then the ex-ante zero-probability event of continued right guessing occurs. What happens to the agent's π and/or awareness state? Most likely, in our view, she will not have to modify her π system, for she will soon come up with one or indeed more concrete stories which would reasonably give rise to the observed outcome with positive probability; that is, she will become aware of new things, and leave $\alpha_t(k)$ for a new awareness state. In other words, observing zero-probability events is in our view a common route to expanded awareness states.

Convergence to P. Going back to estimation —and recalling that the decision maker's purpose is that of getting, over time, at the true, P-based probability that $y = 1$ next period—, there are three problems specific to the present model: (i) $A_t \subsetneq \Phi$; (ii) $L_t \subsetneq A_t$; and (iii) the sets L_t and A_t are not constant. As long as $A_t \subsetneq \Phi$ we can only talk about convergence to the marginal of P on the paths on $\{0,1\}^{A_t}$; as long as $L_t \subsetneq A_t$, some variables in A_t will be 'hidden'; and lastly, any convergence result relies on observing some phenomenon (or at least its effects) for long enough, so the lack of constancy of our observation sets is another complication.

Starting from the last point: given that L_t and A_t increase, all we can do is fix, given t, $L = L_t$ and $A = A_t$ and look at the projection of the process on L and A. Of course as time passes that L will no longer be equal to current L_t, the latter having become presumably larger —but this problem is unavoidable given that L_t varies with time.

Since the process on $\{0,1\}^A$ is in general not Markov (by the same argument as on page 6), the HMM methods for extracting information on variables in $A \setminus L$ do not apply. One possibility is to aim at convergence in the space of paths on $\{0,1\}^L$, for which one can appeal to existing results on merging ('weak merging', given that at t we are only interested in y_{t+1}) à la Blackwell–Dubins, cf. Lehrer–Smorodinsky [13], which require conditions relating the initial prior to (the projection on L of) the true P: if those conditions are met, we get convergence for each fixed L. An alternative, at least for finite A, is to estimate the (marginal) *Markovian* dependence in the $\{\tilde{\omega}_t^{A\setminus L}\}$ process, by HMM methods; this would be more in the spirit of bounded rationality. Some results on convergence of maximum likelyhood estimators are in [19].

The bottom line for convergence of estimates of observed variables seems to be the following: if the agent is confident that L_t is approaching Φ, then she can be also confident that her estimates become good as time goes by. Otherwise, she can trust her estimates concerning marginals on L's observed for long periods; but such estimates can only be good on average (with respect to unobserved variables), and so their value will depend on the environment in which the decision maker operates. Take for concreteness example 1 on page 5, where there is a $\phi \in \Phi$ such that $y_{t+1} = 1$ iff $\phi_t = 1$ and ϕ is zero or one with probability 1/2 each t; suppose the subject is aware of nothing; then he essentially observes an iid fair coin, and after long enough she will find this out, i.e. will estimate that next value of y will be zero or one with probability 1/2; this is indeed correct on average; but it is not of much help if she is betting against somebody who observes ϕ!

4. New Awareness and Posteriors

We now discuss how priors are updated when new knowledge and/or new awareness actually occurs. The sigma–field $\mathcal{F}_t$ records the facts which get to the stage of being observed, and when a fact unobserved at $t-1$ is observed at t we see a corresponding change in $\mathcal{F}_t$. Can this change be different according to whether the subject was aware of the given fact at $t-1$ or not? According to the usual awareness definition it cannot: there is no distinction there between the transition from conscious uncertainty to certainty and that from unawareness to certainty.

But, suppose we are at t and $s < t$. If $\phi_s \notin \mathcal{F}_s$, i.e. the subject did not observe ϕ at s, then either she could not (although she would have wanted to), or she could but did not (and hence she was not aware of ϕ at s, for if she had been and could have observed ϕ, she would have). If it is for the first reason, observing ϕ_t can give no additional information on ϕ_s, $s < t$ —this we will call 'new knowledge'—; if it is for the second, then $\phi_t \in \mathcal{F}_t$ implies that at t she has realized the relevance of ϕ, and then not only will she record ϕ_t, but also ϕ_s, $s < t$ as far as she can —'new awareness', gives hindsight. The two concepts may be then distinguished as in the following

Definitions.

(i) *the subject has* new knowledge of ϕ at t *if* $t = \min\{s\colon \phi_s \in \mathcal{F}_s\}$*;*
(ii) *the subject has* new awareness of ϕ at t *if she has new knowledge of* ϕ *at* t *and there is* $\underline{s} < t$ *such that* $\underline{s} \leqslant s \leqslant t$ *implies* $\phi_s \in \mathcal{F}_t$*.*

Then, whether $\phi_s \notin \mathcal{F}_s$ for one reason or the other will have consequences the first day $t > s$ when $\phi_t \in \mathcal{F}_t$: new knowledge adds just one piece of data to $\mathcal{F}_t$; new awareness brings in a whole bunch of new data, and in this way it may have more impact on behaviour through probability updating.[4]

Priors and Posteriors in the Examples. We come back to the examples of page 4. In the first two we supposed that the agent observes just the value of y_t, and is aware of nothing which may influence its evolution. Then, in both versions he effectively observes repeated tosses of a fair coin: $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \{y_s\}_{s\leqslant t}] = 0.5$ for all $\{y_s\}_{s\leqslant t}$. And as long as she does not observe ϕ she can learn nothing about the underlying, hidden process. In the first for example, even if she were aware of ϕ she could never distinguish, without data on it, the true model from one where ϕ had no influence whatsoever on y, as it would be for example with an arbitrary distribution for ϕ_t and $\mathsf{P}[\tilde{y}_{t+1} = 1 \mid \tilde{\phi}_t] \equiv 1/2$.

The contrast between new knowledge and new awareness is particularly sharp in these situations: suppose that she is aware of ϕ from the outset, and that t is the first time she observes ϕ ('new knowledge'); all is as before up to time t, and from t onwards her estimate of $\mathsf{P}[\tilde{y}_{s+1} = 1 \mid \tilde{y}_s, \tilde{\phi}_s]$ is updated as data on the dependence of y_{t+1} on ϕ_t start accumulating. Suppose instead that she is initially unaware of ϕ until, at time t, not only does she becomes aware of it, but also observes its current *and* past values ('new awareness'); this has obviously a larger impact on her estimate *immediately at* t. Examples 1 and 2 are of course extreme cases, but more generally one may argue that in some concrete situations (stock market trading for example), some discontinuities in behaviour may be the consequence of many agents becoming aware of some relevant variables at the same time. In the third example 'new awareness' of ϕ alone makes no improvement at all, for the observation process remains an iid sequence.

In the context of example 4 with Markov dynamics we can look at the impact of new awareness *without* knowledge on priors and estimates. As we observed, in that example the more the observed relative frequency of the transition from y_t to $y_{t+1} = y_t$ depends on the value of y_t, being higher when $y_t = 1$, the more the hypothesis of big impact of $\psi = 1$ is favoured. In that case, following long strings of $y = 1$, chances are that $\psi = 1$ and so no weight is given to the value of ϕ for prediction of future y; while following time spells with high variability of y chances are higher that $\psi = 0$ and the value of ϕ will be more decisive. In the example the subject is aware of ϕ and ψ from time 0, but consider now what would happen if she were unaware of ψ at the beginning and became aware of it at a later time. Then until unawareness persists the subject can make no use of her observations on y to calibrate her estimated current influence of ϕ on y —which is quite different from the picture with awareness of both ϕ and ψ. So in this case the acquisition of new awareness, *even without knowledge*, has a sizable impact on estimates through the change in the subject's model which it generates.

[4]Terminology is imprecise in the definition in that 'new awareness' is really 'new awareness plus new knowledge', but we are considering the first time a fact becomes observed here: the situation of becoming aware of ϕ at t and not knowing its value implies that then $\phi_t \notin \mathcal{F}_t$ because she cannot observe it; and assuming that $\phi_s \in \mathcal{F}_s \Rightarrow \phi_t \in \mathcal{F}_t$ (recall $s < t$), it must also be $\phi_s \notin \mathcal{F}_s$ for $s < t$. So the first date u such that $\phi_u \in \mathcal{F}_u$ it will be 'new knowledge'.

5. Concluding Comment

The message of this note is that a decision maker in an imperfectly known world should be prepared to change her views, i.e. her model. This will affect her confidence in current probability estimates, but perhaps also their values.

The decision maker has been viewed in isolation, but the message would become sharper and, more importantly, richer in an interactive context. For when consequences depend also on others' behaviour, changes in one's view typically, and often painfully, occur by observing others' actions and payoffs. And in that context observation of actions and payoffs often suggest precise directions in which one's model should be revised.

References

[1] Ågotnes, T. and N. Alechina (2007): Full and relative awareness: A decidable logic for reasoning about knowledge of unawareness, mimeo, University of Nottingham
[2] Board, O. and K.S. Chung (2006): Object-based unawareness, in: Logic and the Foundations of Game and Decision Theory, Proceedings of the Seventh Conference, ed. G. Bonanno, W. van der Hoek, M. Woolridge
[3] Cappé, Olivier, Eric Moulines and Tobias Ryden (2007): *Inference in Hidden Markov Models*, Springer New York
[4] Diaconis, P. and D. Freedman (1993): Nonparametric Binary Regression: a Bayesian Approach, *The Annals of Statistics* **21**: 2108-2137
[5] Diaconis, P. and D. Freedman (1995): Nonparametric Binary Regression with Random Covariates, *Probability and Mathematical Statistics* **15**: 243–273
[6] Feinberg, Y. (2004): Subjective reasoning —games with unawareness, mimeo, Stanford University.
[7] Halpern, J. and L.C. Regô (2006): Extensive games with possibly unaware players, in: Proc. Fifth International Joint Conference on Autonomous Agents and Multiagent Systems: 744-751
[8] Hansen, Lars Pater (2007): Beliefs, Doubts and Learning: The Valuation of Macroeconomic Risk, *American Economic Review* **97**: 1–30.
[9] Heifetz, Aviad, Burkhard C. Schipper and Martin Meier (2007): Unawareness, Beliefs and Games, mimeo
[10] Hill, B. (1986): Some Subjective Bayesian Considerations in the Selection of Models, *Econometrics Reviews* **4**: 191–246
[11] Klibanoff, P., M. Marinacci, and S. Mukerji (2005): A Smooth Model of Decision Making under Ambiguity, *Econometrica* **73**: 18491892.
[12] Klibanoff, Peter, Massimo Marinacci and Sujoy Mukerji (2007): Recursive Smooth Ambiguity Preferences, mimeo
[13] Lehrer, E. and R. Smorodinsky (1996): Merging and Learning, in *Statistics, Probability and Game Theory: Papers in Honor of David Blackwell*, edited by T. Ferguson, L. Shapley and J. MacQueen, *IMS Lecture Notes—Monograph Series* **30**: 147–168
[14] Li, J. (2006): Dynamic games of complete information with unawareness, mimeo, University of Pennsylvania.
[15] Modica, S. and A. Rustichini (1994): Awareness and Partitional Information Structures, *Theory and Decision* **37**: 107-124
[16] Modica, S. and A. Rustichini (1999): Unawareness and Partitional Information Structures, *Games and Economic Behavior* **27**: 265–298
[17] Poirier, D. (1988): Frequentist and Subjectivist Perspectives on the Problem of Model Building in Economics, *Journal of Economic Perspectives* **2**: 121–144
[18] Regô, L. C. and J. Halpern (2007): Generalized solution concepts in games with possibly unaware players, mimeo, Federal University of Pernambuco.
[19] Rydén, T. (1997): On Recursive Estimation for Hidden Markov Models, *Stochastic Processes and their Applications* **66**: 79–96

[20] Sadzik, T. (2006): Knowledge, awareness and probabilistic beliefs, mimeo, Graduate School of Business, Stanford University

[21] Schipper, Burkhard (2007): The Unawareness Bibliography, available at `www.econ.ucdavis.edu/faculty/schipper`

[22] Sillari, G. (2006): Models of awareness, in: Logic and the Foundations of Game and Decision Theory, Proceedings of the Seventh Conference, ed. G. Bonanno, W. van der Hoek, M. Woolridge.

[23] Weitzman, Martin (2007): Prior Sensitive Expectations and Asset Return Puzzles, *American Economic Review*, Forthcoming