Learning and Discovery

Boyan Jovanovic*

October 12, 2009

Abstract

Bayesian inference is a process of eliminating parameter values that do not explain the data and shifting posteriors towards values that do explain them. There is no room in it for 'discovery'. I study belief processes that allow for discovery. I then ask when one can approximate discovery-induced beliefs by a Bayesian belief mechanism. This allows us to define optimal decisions and rational valuations in environments with discovery.

PRELIMINARY

1 Introduction

Learning and discovery appear to be two qualitatively different ways in which beliefs change. By 'learning' we shall mean the usual form of statistical updating of beliefs in light of evidence. By a 'discovery' we shall mean a change in beliefs such that the posterior puts positive mass on an event on which the prior had a zero mass. Bayes learning is learning from experience alone. Such 'experience learning' leaves no room for 'insight' which brings about a new hypothesis never before conceived of. More precisely, it provides no way for evidence to convert a zero-probability event into one to which one would assign a positive probability.[1]

This paper shows that in a class of models previously discussed by Easley and Kiefer (1988) in which discovery can be modeled in a standard way. In this model, if

*I thank L. Blume, T. Cogley, S. Galanis, A. Kochov, M. Kredler, R. Lucas, Y. Nyarko, E. Ozbay, R. Radner, J. Schlossberger, J. Stoye, H. Tretvoll and A. Tsyvinski for comments and the Kauffman Foundation and the NSF for support.

[1]Let A be an event to which the agent assigns a probability zero, i.e., $P(A) = 0$. Then for any other event B,

$$P(A \cap B) = P(A \mid B)\, P(B) = 0$$

so that $P(B) > 0 \Longrightarrow P(A \mid B) = 0$, whereas if $P(B) = 0$, $P(A \mid B)$ is undefined.

a Bayesian receives a large signal pointing to an event to which he assigned a negligible but positive prior belief, the resulting shift in beliefs will be indistinguishable from one that results from a discovery.In other words, discovering a new state has roughly the same implications for observables as does a strong signal pointing to a state that received a positive but very low prior weight. This equivalence then allows us to place an value to information in such situations. And such situations are firm that invests in basic research.

The paper proceeds in three steps, described in Sections 2, 3 and 4. Section 2 develops a model of belief evolution when discoveries can occur over time. It adds an awareness-growth mechanism to the Easley Kiefer (1988) formulation of the Bayes decision model. Assuming a hypothetical prior covering the entire state space, the prior is then restricted to have its support on a strict subset of the state space. That subset is the agent's initial awareness. Awareness is said to grow when this conditioning set grows. An exogenous mechanism for discovery is introduced. Together with the initial prior, this mechanism implies that beliefs can evolve in a non-Bayesian way, with a discovery mass appearing on sets that were previously of measure zero. .

Section 3 asks when one can approximate discovery-induced beliefs arbitrarily closely by a Bayesian belief mechanism. This can be done if we are allowed to introduce a new set of signals. Having approximated the discovery process in this way, we obtain a model of decisions and valuations in situations in which something like discovery takes place, but which we can model in the standard way.

Section 4 does the reverse; it asks when models of learning can be approximated by models of discovery. In particular, it shows how two standard models of technological progress – Telser (1982) or Muth (1986) on the one hand, and Prescott and Mehra (1994) on the other – can be re-cast in a model in which the agent discovers new technologies, thereby improving his decisions.

Section 5 formulates the equivalence question more generally, but there is no general result yet. Section 6 discusses how one may wish to restrict the discovery process in light of previous research on similar topics.

Related work divides into several strands. Decision theory stresses that unawareness of an event is different from having beliefs that assign a probability zero to it. Dekel, Lipman and Rustichini (1998), Li (2008), Galanis (2008) and Ozbay (2008, Sec. 4) discuss this point and they further discuss many other related papers. Applied work also routinely deals with environments in which there discoveries such as inventions; it typically assumes that agents form beliefs over payoffs, costs or productivity. Models of technological improvement arising from research such as Telser (1982), Muth (1986) and Kortum (1997) proceed in this way, as do research-driven-growth models of Romer (1990) and Aghion and Howitt (1992). Empirical work on patents on a firm's stock price value by Pakes (1985) assumed that the firm invents something; the content of that invention does not matter, only the amount by which it lowers production costs and raises the firm's value. In other studies summarized by Griliches (2000), basic research is held as having a higher probability than applied research of producing an outcome in the right tail of the payoff distribution.

The (exogenously-timed) introduction of a new parameter value into the prior leads to a re-evaluation of the evidence and a possibly a dramatic shift in beliefs. Other models that feature dramatic shifts in beliefs are the Bayesian model of Cogley and Sargent (2005), and non-Bayesian models of Young (02, Sec. 8.3) and Cho and Kasa (2009) who study how agents would periodically switch models whenever they fail a statistical test and Kocherlakota (2007) has a related discussion. Nyarko (1991) and Evans and Honkapohja (2001) extensively discuss learning of (sometimes) misspecified equilibrium models. Venezia (1985) and Auerswald *et al.* (2000) discuss other learning algorithms, and a paper on paradigm shift is Bramoullé and Saint Paul (2007).

All this relates to the issue of the "directedness" of discovery. Sometimes an agent's awareness grows as a result of an action taken by an opponent in the game, an action that may simply be a self-interested announcement by another agent as in Ozbay (2008), and in this sense directed towards that agent's interests. Sims (1971) and Radner (2002) offer non-Bayesian treatments of model revision, with each revised model having more parameters than the previous model, so that model revision is directed towards the next layer of complexity. Under full awareness Rothschild (1974), Jovanovic and Rob (1990) and Jovanovic and Nyarko (1996) used information theory to study directed search in variations of the multi-armed bandit model.

2 Model

We shall introduce discovery into the model of Easley and Kiefer (1988). Let us first review that model with some minimal modifications. Time is discrete. In each period t the decision-maker chooses an action $x_t \in X$. At the end of the period t he observes $y_t \in Y$ and receives a reward $U(x_t, y_t)$. The discount factor is δ. The random variable y_t has density

$$y \sim f(y \mid x, \theta) \tag{1}$$

where $\theta \in \Theta$ is a parameter. Define the one-step-ahead Bayes map

$$\mu' = B(x, y, \mu) = \frac{f(y \mid x, \theta)\, \mu(\theta)}{\int_\Theta f(y \mid x, \theta)\, d\mu(\theta)}. \tag{2}$$

This first-order difference equation for μ leads to the Bellman equation

$$V(\mu) = \max_{x \in X} \int_Y \left[U(x, y) + \delta V(B(x, y, \mu))\right] f(y \mid x, \theta)\, dy d\mu(\theta). \tag{3}$$

Equations (1)-(3) represent the Easley-Kiefer model. Let us now modify it by adding limited awareness into the model. By 'Awareness' we shall mean simply the subset of Θ on which beliefs have their support. 'Unawareness' of a subset of Θ is equivalent to having beliefs that assign measure zero to that subset.[2]

Awareness-A beliefs.—With awareness $A \subset \Theta$, the agent's actual belief is the restriction of μ to his awareness set, A:

$$m(\theta) = \mu(\theta \mid A) = \begin{cases} \frac{\mu(\theta)}{\mu(A)} & \text{for} \quad \theta \in A, \\ 0 & \text{for} \quad \theta \notin A. \end{cases} \tag{4}$$

A growth in awareness then occurs when A gets larger.[3] In order for m to continue to be defined for any $A \subset \Theta$, μ must be defined over all of Θ. This means that we must start with a hypothetical prior over Θ from which we can iterate using (2) conditional on A.[4]

3 Discovery approximated by standard learning

We now give two examples in which discovery differs from learning, but in which the two can produce the same outcomes for the variables that an outsider could observe.

3.1 Example 1: Investments with Binomial outcomes

A gamble pays a dollar each time a coin comes up heads, and nothing if it comes up tails. Let θ be the probability of heads, and let θ be zero, so that the coin always comes up tails. Let $y_t \in \{0, 1\}$ be the gross payoff to the gamble, so that $y_t = 0$ is observed by the agents for all t. We shall ask how this evidence, tails for ever, will affect the evolution of beliefs of two risk-neutral agents, Agent 1 and Agent 2. Agent 1 experiences a discovery whereas Agent 2 does not.

Let $\Theta = \{\theta_1, \theta_2\}$ where $\theta_1 = 1/2$ and $\theta_2 = 0$. This is the universe of the states that θ can assume. We choose the prior μ_0 so that it assigns equal probability to the two events.

Agent 1.—Let Agent 1's initial awareness be $A_0 = \{\theta_1\}$. Substituting A_0 for A in (4) we see that Agent 1 starts out believing dogmatically that the coin is fair, i.e., that $\theta = \theta_1 = 0.5$. At date T, the agent discovers that θ could be zero. He never conceived of such a possibility before date T but, having discovered it, he considers it as likely *a priori* as the possibility that $\theta = 1/2$. Absent any evidence, that is, he

[2] I use "awareness" for want of a better term, recognizing that decision theorists have stressed that unawareness of an event and the assignment of a zero probability to that event, are qualitatively different – e.g., Dekel *et al.* (1998), Galanis (2008), Li (2008) and Ozbay (2008).

[3] To simplify, let $f(y \mid x, \theta) > 0$ for all $(x, y, \theta) \in X \times Y \times \Theta$. Then no observation is logically impossible for any $\theta \in \Theta$, and the support of beliefs over Θ cannot shrink.

[4] This hypothetical prior is similar in nature to the 'belief function' in Ozbay (2008). More generally, it is similar to how in games one specifies beliefs over types that take a particular off-equilibrium action.

assigns equal probabilities over the two events. Then his prior is 0.5 on $\theta = 1/2$ and 0.5 on $\theta = 0$. With these new priors and the evidence that the coin came up tails T periods in a row, his expected gross payoff, $E(y_t)$, falls suddenly from 0.5 for $t < T$ to $\frac{1}{1+2^t}$ for $t \geq T$:

$$E(y_t) = \begin{cases} \frac{1}{2} & \text{for } t < T \\ \frac{1}{1+2^t} & \text{for } t \geq T, \end{cases} .$$

Let us plot this as a function of t for three alternative discovery dates $T = 2, 12, 21$. These are the three heavy purple lines in Fig. 1.

Agent 2.—Now consider agent 2 who, from the outset, is aware of the possibility that $\theta = 0$, who has a prior of μ that $\theta = 0$ and $1 - \mu$ that $\theta = 1/2$. His posterior belief over $\theta = 0$ is

$$\pi(\theta = 0 \mid t \text{ tails and no heads}) = \frac{\mu}{\mu + (1-\mu)\left(\frac{1}{2}\right)^t} = \frac{\mu 2^t}{\mu 2^t + 1 - \mu}$$

Agent 2's expected gross payoff is

$$E^*(y_t) = \frac{1}{2}\left(1 - \frac{\mu 2^t}{\mu 2^t + 1 - \mu}\right) = \frac{1}{2}\frac{1-\mu}{\mu 2^t + 1 - \mu}$$

In Figure 1 we plot $E^*(y_t)$ for six values of $\mu = \{10^{-1}, 10^{-2}, ..., 10^{-6}\}$. We find that even as μ becomes quite small and the possibility of the coin being biased becomes quite remote, the model cannot generate the sudden drop in expected values. The drop takes place mostly over about 10 periods. This Figure and the resons for the shape of the curves are similar to Figures 3 and 9 of Cogley and Sargent (2005).

Actions.—Suppose that taking the gamble costs 25¢. This can be thought of as the cost of investing in the project and the up-front costs of hiring the factors or production. Then until date T, firm 1 would keep losing money and then after date T investment would cease. For firm 2, investment would cease when E^* drops below 0.25. The optimal stopping boundary is strictly below 0.25 (because of the incentive to experiment and the option of stopping if the news is unfavorable) and approaches 0.25 as $t \to \infty$.

Value of firm.—If we choose μ so that we fit actions, output and profits, clearly the Bayesian model with full awareness is unable to match the stock-price drop. If the business was public and if the public had the same beliefs, the market value of that business would experience a sudden crash in the first model, and would take at least 10 periods to do so in the second. On the other hand if one is interested in the behavior of the price of a stock of such a company and is willing to tweak the model and assume that some periods produce a larger number of signals, then a faster market crash is possible in the Bayesian model. For instance, if in Figure 1 we take the right-most curve corresponding to $\mu = 10^{-6}$, we could assume that at $t = 17$, the agent receives not one signal but ten. This would produce a rapid crash of the stock price, but would interfere on observations at the frequency at which trials actually took place. Therefore the signals would need to be unrelated to the firm's output.

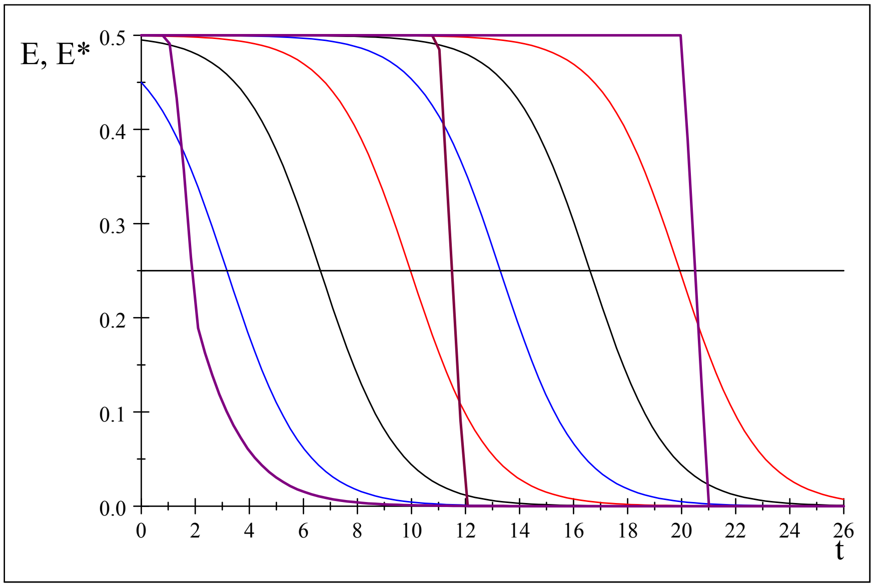

Figure 1: THE BINOMIAL CASE

3.2 Example 2: Normal outcomes

Since the normal distribution is a limit of additions of small binomial trials, we shall in this second example reach conclusions similar to those reached in the binomial case. Consider a sequence of normally-distributed gambles that pay

$$y_t = \theta + \varepsilon_t \tag{5}$$

for $t = 0, 1, ...,$ where $\varepsilon \sim N(0, \sigma^2)$. Again, let $\Theta = \{\theta_1, \theta_2\}$ where $\theta_1 = 1/2$ and $\theta_2 = 0$, and again μ_0 assigns equal probability to the two events. To correspond to the above case, we suppose that the true $\theta = 0$ and ask how evidence generated this time by (5) will affect the evolution of beliefs of Agent 1 and Agent 2.

Agent 1.—He again believes dogmatically that $\theta = 0.5$, and again at date T he discovers that θ could be zero, and assigns equal prior probabilities over $\theta = 0$ and $\theta = 0.5$. Let $\bar{y}_t = \frac{1}{t}\sum_{s=0}^{t-1} y_s$. In contrast to the Binomial case, conditional on the true θ being zero, the evidence, $\bar{y}_t \sim N\left(0, \frac{1}{t}\sigma^2\right)$ is a random variable. There must always be a positive drop at the date T, but the size of the drop of Agent 1's expectations of y_t at date T is then also a random variable, shifting from 0.5 to

$$E(y_T) = \frac{1}{2}\frac{1}{1 + \exp\left\{-\frac{T}{2\sigma^2}\left[\bar{y}_T^2 - \left(\bar{y}_T - \frac{1}{2}\right)^2\right]\right\}}$$

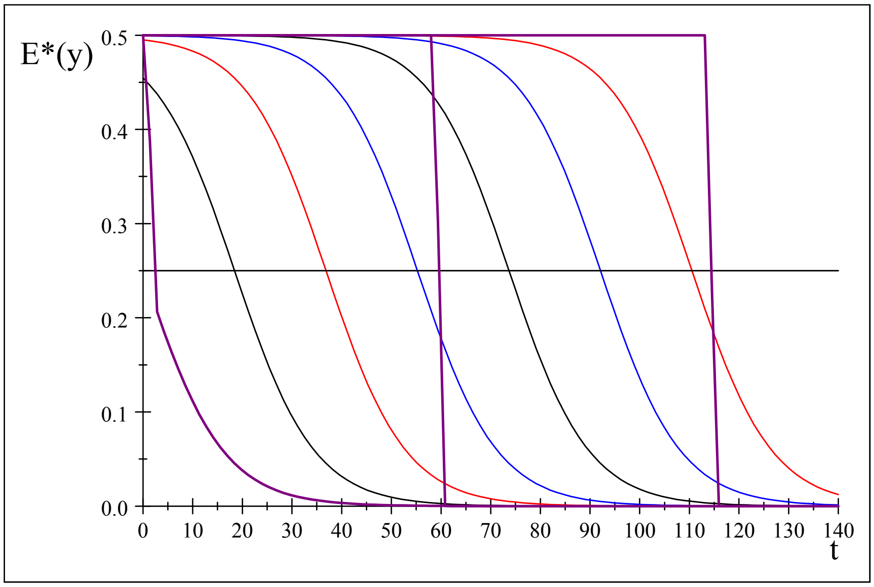

Figure 2: The Normal case

At the median, $\bar{y}_t = 0$, and so the median post-discovery expected payoff is

$$E_{\text{MED}}\left(y_t\right) = \frac{1}{2}\frac{1}{1+\exp\left(\frac{t}{8\sigma^2}\right)} \quad \text{for } t \geq T$$

Assuming that $\sigma = 1$, let us plot this as a function of t for three alternative discovery dates $T \in \{2, 60, 115\}$. These are the three heavy purple lines in Fig. 2.

$$\min\left(1, \max\left(0, T-t\right)\right)\frac{1}{2} + \left(1 - \min\left(1, \max\left(0, T-t\right)\right)\right)\frac{1}{2}\frac{1}{1+\exp\left(\frac{T}{8}\right)}.$$

Agent 2.—From the outset this agent assigns probability μ that $\theta = 0$, and $1-\mu$ that $\theta = 1/2$. His date-t posterior expectation is

$$E^*\left(y_t\right) = \frac{1}{2}\frac{1}{1+\frac{1-\mu}{\mu}\exp\left\{-\frac{T}{2\sigma^2}\left[\bar{y}_t^2 - \left(\bar{y}_t - \frac{1}{2}\right)^2\right]\right\}}.$$

At the median, $\bar{y}_t = 0$, and so the median post-discovery expected payoff is

$$E^*_{\text{MED}}\left(y_t\right) = \frac{1}{2}\frac{1}{1+\frac{\mu}{1-\mu}\exp\left(\frac{t}{8\sigma^2}\right)}.$$

In Figure 2 we plot $E^*_{\text{MED}}\left(y_t\right)$ for $\sigma = 1$, six values of $\frac{\mu}{1-\mu} = \left\{10^{-1}, 10^{-2}, ..., 10^{-6}\right\}$.When $\sigma = 1$, learning is very protracted and the full-awareness model cannot replicate a sudden drop in expectations. But when σ is small, information accumulates rapidly, because evidence accumulates at the rate and a combination of a small μ and small σ can indeed generate a precipitous drop in expectations. Of course we are not generally free to choose σ – it is determined by the properties of the y_t sequence. Again, one could add a signal at date t other than y_t, a signal not observed by the analyst, and thereby allow the full-awareness Bayesian model to fit the data.

3.3 Example 3: A learning curve

A particularly transparent effect of simulating a discovery at some date by endowing agents with a precise signal at that date is in the context of learning curves. As in Jovanovic and Nyarko (1996) suppose that x_t is an action, y_t a random variable and that output (and utility) is

$$U(x, y) = 1 - (y - x)^2,$$

where

$$y_t = \theta + \varepsilon_t$$

Let $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and suppose that the agent observes not only the realizations of output, but y_t itself. Let $\Theta = R$, and let $\mu_0 = N(\bar{\theta}, \sigma_\theta^2)$ be the normal prior over Θ.

Agent 1.—For any subset $A_0 \subset \Theta$, Agent 1 has the prior $\mu_0(\theta \mid A_0)$. Optimal decisions of an agent with awareness A are that

$$x_t = \int \theta d\mu(\theta \mid A)$$

If awareness were never to change so that A were for ever to remain fixed, decisions would converge to the "best" point in A, i.e.,

$$x_t \to\in \arg\min_{\theta \in A} \mid \theta - \theta^* \mid .$$

and expected output and utility would converge to

$$E(U_t) = 1 - \sigma_\varepsilon^2 - \min_{\theta \in A} (\theta - \theta^*)^2$$

If awareness would then grow to A', expected output would generally increase to $1 - \sigma_\varepsilon^2 - \min_{\theta \in A'} (\theta - \theta^*)^2$. Thus the learning curve would generally exhibit a jump at any discovery that was closer to the true θ. Clearly, such a learning curve would have concave portions interrupted by upward jumps, roughly as found in some empirical work summarized in Muth (1986).

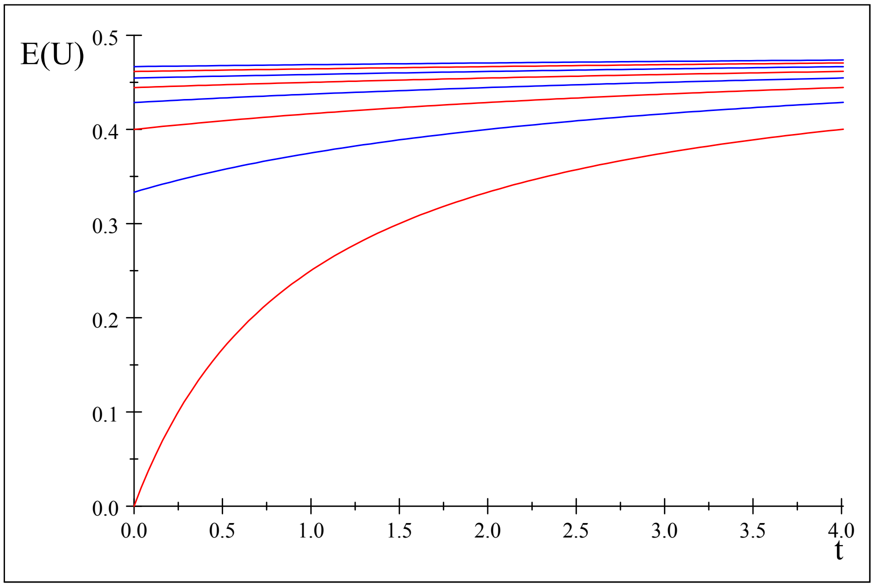

Figure 3: THE EFFECT ON EXPECTED OUTPUT OF FOUR PRECISION UNITS

Agent 2.—A standard learning structure with a normal prior $\mu_0 \sim N\left(\bar{\theta}, \sigma_\theta^2\right)$ on R produces a posterior $\mu_t = N\left(m_t, \sigma_{\theta,t}^2\right)$, where the posterior mean and variance are

$$m_t = \frac{\sigma_\theta^{-2} + \sigma_\varepsilon^{-2} \sum_{s=0}^{t-1} y_s}{\sigma_\theta^{-2} + \sigma_\varepsilon^{-2} t}, \qquad \text{and} \qquad \sigma_{\theta,t}^2 = \frac{1}{\sigma_\theta^{-2} + \sigma_\varepsilon^{-2} t}.$$

The optimal decision is $x_t = m_t$, and expected output at date t is

$$E\left(U_t\right) = 1 - \sigma_\varepsilon^2 - \sigma_{\theta,t}^2. \tag{6}$$

The learning curve (6) is plotted for $\sigma_\theta^2 = \sigma_\varepsilon^2 = 0.5$. Each curve in Fig. 3 is a horizontal displacement to the left by two time units or four information-precision units. The vertical distances between each pair of curves represent the jump in expected output that would result from the a sudden arrival of four precision units of information.

4 Learning approximated by discovery

In this section we start with a standard learning setup and ask what the corresponding discovery model would be. Both examples involve sampling technologies or wages from an urn as in Burdett (1978), Telser (1982), and Muth (1986).

4.1 Example 4: Search

First we present the search model and then show that the outcomes of search can be generated by a model in which the parameters of a production function are gradually discovered.

4.1.1 Search

Suppose a risk-neutral agent samples y from a distribution $G(y)$, for $y \in [0,1]$. Suppose that y is the output of a technology if it is used for production. New technologies are sampled (with no resource cost) at the rate of one per period. The technology in use is then the highest sampled to date so that one i.i.d. sample is taken per period. The distribution of the maximum of t draws

$$Y_t = \max_{0 \le i \le t-1} (y_i)$$

is

$$\Pr(Y_t \le Y) = G^t(Y)$$

If one can also produce while sampling, Y_t, then $G^t(Y)$ is also the distribution of output at date t. When there is no sampling cost, this is the model of Burdett (1978) and Muth (1985), and when there is a sampling cost it is the model of McCall (1965).

4.1.2 Discovery

Now we show that the same decision problem can be framed in the discovery model. Assume that y_t is output produced using the production function

$$y_t = 1 - (x_t^p - \theta)^2 + \varepsilon_t, \tag{7}$$

where $x^p \in R$ is a 'production' decision (later there will also be a 'discovery' decision, $x^d \in \{0,1\}$), $\theta \in \Theta \subseteq R$ is a parameter, and $\varepsilon_t \in R$ is an i.i.d. disturbance. The agent does not know θ, however. Instead, he is aware of a finite subset A of Θ, and his prior, denoted by $\mu_A(\theta)$, is uniform over A so that

$$\mu_A(\theta) = \begin{cases} \frac{1}{\#A} & \text{for } \theta \in A \\ 0 & \text{for } \theta \notin A \end{cases} \tag{8}$$

If the agent's awareness were to grow from A to A', say, where A' is a superset of A, then prior beliefs would again be given by (8) but with A' in place of A.

The discovery process.—Let discovery occur via i.i.d. sampling of θ from a distribution with C.D.F. $H(\theta)$ that has support Θ. This means that discovery is 'undirected' in the sense that the true θ is not involved. Then instead of writing (36) and (37) in distribution form, it is easier to write the difference equation for A'

$$A' = A \cup \theta \quad \text{with prob. } dH(\theta) \tag{9}$$

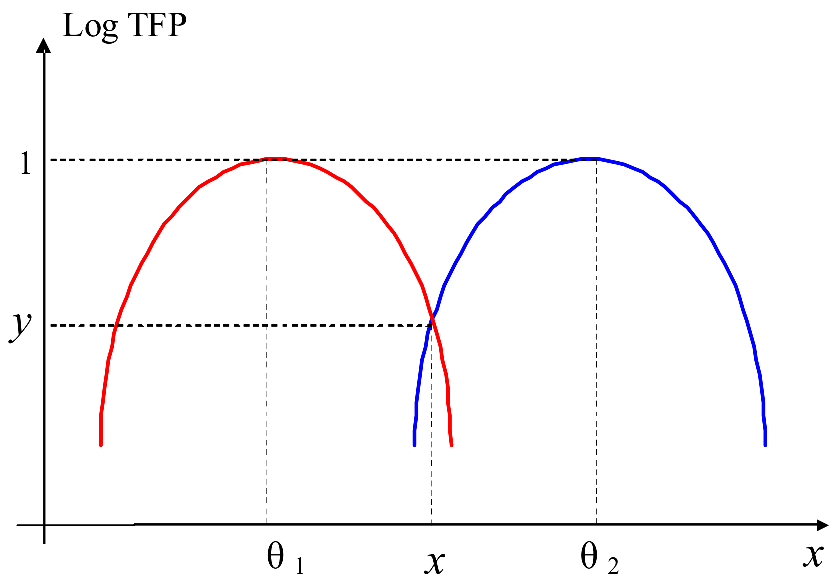

Figure 4: ANY (x, y) IMPLIES THAT $\theta \in \{\theta_1, \theta_2\}$

so that awareness remains unchanged with prob. $\int_A dH(\theta)$.

The role of ε.—If $\varepsilon \equiv 0$, the agent would quickly be able to learn the true θ; (7) has two solutions for θ:

$$\theta = x^p \pm \sqrt{1-y}.$$

These two solutions are illustrated as θ_1 and θ_2 in Figure 4. Therefore one needs just two independent observations of (x, y) to solve for θ, and this would contradict the slower discovery process that we shall posit below.

We now proceed in two steps. First, we show that conditional on deciding to discover a new θ, the distribution sampling, the distribution of y_t as given by (7) will be $G(y)$ different way of deriving $G^t(Y)$ as the distribution of the largest y sampled.

4.1.3 Deriving $G(y)$ via discovery

Additional signals on θ.—Before starting out, the agent has seem T signals s_i with density

$$\xi(s \mid \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(s-\theta)^2} \tag{10}$$

for[5] $i = -T, -T+1, ..., -1$.

[5]The idea is that the agent already has T observations of y at date $t = 0$

Optimal decisions.—Since (9) does not involve x^p, the production decision is static. Being risk neutral, the agent solves at each date

$$\min_{x\in R} E\left((x^p-\theta)^2 \mid h^{t,T},A\right)$$

where $h^{t,T}=\left(x^t,y^t,s^T\right)$ and where

$$x^t=(x_0,...,x_{t-1})\,,\; y^t=(y_0,...,y_{t-1})\,,\quad \text{and}\quad s^T=(s_{-T},s_{-T+1},...,s_{-1})\,.$$

The solution is to set x^p to equal the date-t expectation of θ:

$$x_t=E\left(\theta \mid h^{t,T},A\right). \tag{11}$$

Now let T get large. Then the information in the history of s_t dominates any information that will be received via (y_t). Ignoring the information in y would mean that the expectation in (11) would be taken with respect to the posterior distribution which, in light of the uniform prior reads

$$\mu\left(\theta \mid h^{t,T},A\right)=\frac{\prod_{i=-T}^{-1}\xi\left(s_i\mid\theta\right)}{\sum_{\theta\in A}\prod_{i=-T}^{-1}\xi\left(s_i\mid\theta\right)}, \tag{12}$$

for $\theta\in A$, where ξ is defined in (10).

Let θ^* denote the true value of θ. Then as $T\to\infty$ and the information becomes perfect, the posterior mean converges to

$$E\left(\theta \mid y^t,s^T\right)\overset{\text{a.s.}}{\to}\arg\max_{\theta\in A}\int \ln\xi\left(s\mid\theta\right)\xi\left(s\mid\theta^*\right)ds, \tag{13}$$

which also is the limit of the maximum-likelihood estimate of θ restricted to A. Although he has unlimited information that should clearly indicate θ^* when $\theta^*\in A$, when $\theta^*\notin A$, the agent obstinately treats that information as having arisen from an unlikely sequence (s_t,ε_t) governed by that value of $\theta\in A$ that best explains s^T.

Then

$$x_t=E\left(\theta \mid y^t,s^T\right)\overset{\text{a.s.}}{\to}\in\arg\min_{\theta\in A_t}|\theta-\theta^*| \tag{14}$$

We include the additional '$\in$' in (14) because there may be more than one argmin. Now let $\sigma^2\to 0$, in which case, since $\varepsilon\overset{\text{i.p.}}{\to}0$.[6] In that case we can easily show equivalence of the discovery model and the search model.

[6]This simplifies the relation between G and H. The probabilistic removal of ε will mean that we

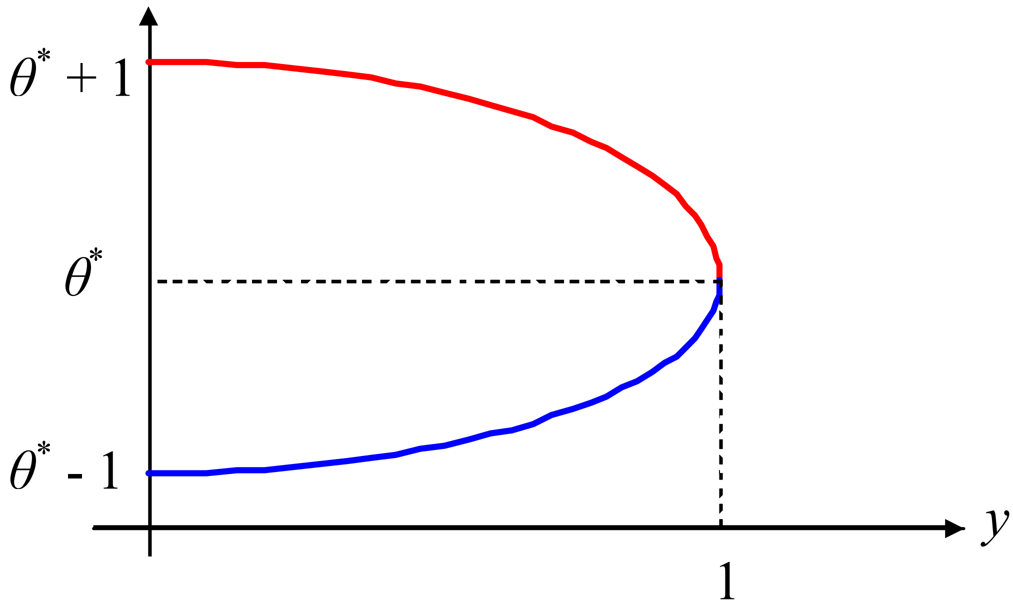

Figure 5: The Technology correspondence (15)

Equivalence.—Therefore for each $G(y)$ there exists a discovery process generating the same observations, in probability. If the agent starts with $A_0 = \varnothing$ and if he were to make one discovery per period, call it θ_t such that

$$y_t = 1 - (\theta_t - \theta^*)^2,$$

then the sequences of outputs will be the same at each date.

The technology correspondence.—For each output level $y < 1$, there are two technologies that give rise to it. Therefore we have the technology correspondence

$$\theta = \theta^* \pm \sqrt{1-y}, \tag{15}$$

which is depicted in Figure 5. If any measurable selection $\psi(y)$ from this correspondence is consistent with $G(y)$, in the sense the the distribution over Θ that it gives

can associate output deterministically with the pair (θ^*, A) as we shall see below. In the theory of search by which new methods of production, as summarized by the cost of production, are drawn from a distribution (Telser (1982), Muth (1986), Kortum (1997)), the technology in use is the most efficient hitherto sampled, and it is used until a still better technology is discovered. In this approach, a sampled technology is an 'inspection good' – having discovered a new technology, the potential user does not have to try it in order to be able to evaluate it and compare it to the other technologies that he knows. Since most technologies must be tried before their full potential is known and is realized, this approach may work if time periods are long enough so that any experimentation can be treated as occupying a negligible fraction of any period, and that the bulk of the time in each period, the technology used is the best hitherto tried.

rise to, call it $H^{\psi}(\theta)$, is consistent with G, then then all selections are consistent with G. That is, for any subset $B \subset \Theta$,

$$H^{\psi}(B) = G\left(\psi^{-1}(B)\right). \tag{16}$$

Let us focus on just two selections from the technology correspondence: The smallest and the largest.

The discovery process that generates $G(y)$.—It is enough that we find just one selection from (15). The selection that maximizes θ of course obtains when, for each y, we take the larger solution in (15) for θ, i.e., $\theta(y) = \theta^* + \sqrt{1-y}$. Then $\theta_t \leq \theta \Longleftrightarrow y_t \geq 1 - (\theta - \theta^*)^2$, and the CDF for θ needed to generate G is

$$H^+(\theta) = 1 - G\left(1 - (\theta - \theta^*)^2\right) \qquad \text{for } \theta \in [\theta^*, \theta^* + 1].$$

Alternatively, consider the smaller solution in (15) for θ_t, i.e., $\theta_t = \theta^* - \sqrt{1 - y_t}$. Then $\theta_t \leq \theta \Longleftrightarrow y_t \leq 1 - (\theta^* - \theta)^2$, and the CDF for θ needed to generate G is

$$H^-(\theta) = G\left(1 - (\theta^* - \theta)^2\right) \qquad \text{for } \theta \in [\theta^* - 1, \theta^*].$$

It does not matter which is chosen or, indeed, if a linear combination of the two is chosen as follows

$$H(\theta; \alpha) = \alpha I_{\{[\theta^*-1,\theta^*]\}} H^-(\theta) + (1-\alpha)\, I_{\{[\theta^*,\theta^*+1]\}} H^+(\theta) \tag{17}$$

for $\alpha \in [0, 1]$. If all we can observe are the y's, the parameter α is not identified, and neither, of course is the selection ψ.

Example of a uniform G.—To illustrate (17), assume $G(y) = y$ for $y \in [0, 1]$, and suppose that $\theta^* = 0$. Then

$$\begin{aligned} H^+(\theta) &= \theta^2 \qquad \text{for } \theta \in [0,1] \text{ and} \\ H^-(\theta) &= 1 - \theta^2 \qquad \text{for } \theta \in [-1, 0] \end{aligned}$$

and the densities are

$$\begin{array}{ll} h^+(\theta) = 2\theta & \text{for } \theta \in [\theta^*, \theta^* + 1] \text{ and} \\ h^-(\theta) = -2\theta & \text{for } \theta \in [\theta^* - 1, \theta^*] \end{array} \tag{18}$$

and we illustrate them in Figure 6, where we also have extended their definition to the interval $[-1, 1]$. Then any density of the form

$$h^{\alpha}(\theta) = \alpha h^-(\theta) + (1-\alpha)\, h^+(\theta) \tag{19}$$

$\alpha \in [0, 1]$ also generates the C.D.F. $G(y) = y$.

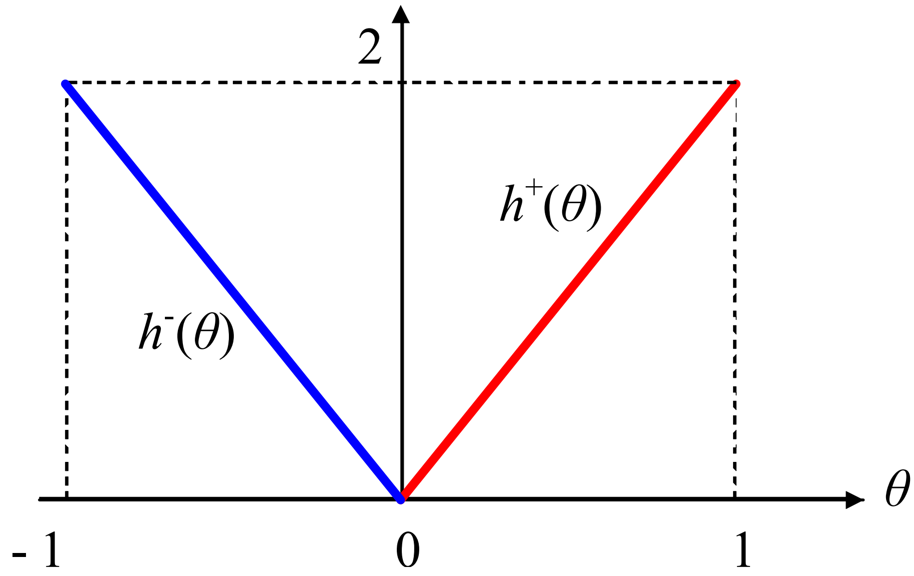

Figure 6: TWO DISCOVERY DENSITIES CONSISTENT WITH A UNIFORM $G(y)$

4.1.4 Sampling costs and the search/discovery decision

While the production decision is static, when there is a sampling cost c, the search decision is dynamic. But having derived $G(y)$ via the distribution H in (17) and, for a specific example, in (18), it is straightforward to add a search-investment decision $x^d \in \{0, 1\}$. The decision is $x = \left(x^p, x^d\right)$. As $T \to \infty$, so that

$$U(x, y) = y - cx^d.$$

As $\sigma \to 0$,

$$y \to 1 - (x^p - \theta^*)^2 = 1 - \min_{x \in A} (x - \theta^*)^2 \tag{20}$$

That is, as $\sigma \to 0$, the following three truths emerge

- $f(y \mid x, \theta)$ becomes degenerate, and we substitute y out in (38) via (20),
- $\mu(\theta \mid h, A)$ has converged with all its mass to $\arg\min_\theta (\theta - \theta^*)^2$
- V becomes stationary as h, being fully informative about θ^*, no longer matters.

Then (38) reads

$$\begin{aligned} V(A) &= \max_{(x^p,x^d)\in A\times\{0,1\}} \left\{ \left[\int_A \left(U(x,y) + \delta \int_{\mathcal{B}(\Theta)} V(A')\, d\phi(A' \mid A,h,x) \right) d\mu(\theta \mid h, A) \right] \right\} \\ &= 1 - \min_{x^p\in A} (x^p - \theta^*)^2 + \max_{x^d\in\{0,1\}} \left\{ -cx^d + \delta \max \int_\Theta V(A\cup\theta)\, dH(\theta) \right\} \\ &= 1 - [\rho(A,\theta^*)]^2 + \max\left\{ \delta V(A), -c + \delta \int_\Theta V(A\cup\theta)\, dH(\theta) \right\} \end{aligned} \tag{21}$$

where

$$\rho(A,\theta^*) = \min_{\theta\in A} |\theta - \theta^*| \tag{22}$$

is the Hausdorff distance between θ and A in the Euclidean norm. Then since the operator is a contraction, we find that V must be of the form $V(A) = \hat{V}(\rho[A,\theta^*])$. Then (21) reads

$$v(\rho) = 1 - \rho^2 + \max\left\{ \delta v(\rho), -c + \delta \int_0^\infty v(\min(\rho,\rho'))\, dP(\rho') \right\}. \tag{23}$$

where $P(\rho')$ is the distribution of ρ' implied by $H(\theta)$ Change variables from ρ to $\hat{y}(\rho) = 1 - \rho^2$ and note that any $H(\theta)$ satisfying (16) then gives rise to the transform $G(y) = H(\theta)$. Then define w by

$$w\left(1-\rho^2\right) = v(\rho)$$

Then (23) reads

$$v(y) = y + \max\left\{ \delta v(y), -c + \delta \int_0^\infty v(\max(y,y'))\, dG(y') \right\}$$

and the discovery decision then is

$$x^d = \begin{cases} 1 & \text{for } y < y^* \\ 0 & \text{for } y \geq y^* \end{cases} \tag{24}$$

where y^* solves the equation

$$c = \delta \int_0^\infty [v(y') - v(y)]\, dG(y'). \tag{25}$$

The decision rule characterized by (24) and (25) is the optimal stopping rule for an infinitely-lived agent who samples at a cost c from a distribution G that he knows. then his search decisions and reservation y would be the same as if discovery was i.i.d., with distribution H given in (17).

So far we assumed that the agent's guess about ϕ is correct, which then allows the model to generate search from a distribution G that the agent knows. If the agent does not know ϕ and learns about it along the way, then this would have a counterpart in search theory of the agent learning about G. Rothschild (1974) showed there would be a sequence of stopping set $\left(Y_t^S\right)$ such that the agent stops sampling when $y^t \in Y_t^S$ for the first time. In the discovery model the agent would have stopping sets $\left(\Theta_t^S\right)$ such that he would stop sampling when $\theta^t \in \Theta_t^S$ for the first time.

4.2 Example 5: Growth in the Prescott-Mehra model

As with Example 4, we first outline the usual assumptions about the stochastic process, and we then show how it may arise via a process of discovery. Only a sketch will be provided here because the development is quite similar to that in Example 4.

4.2.1 Exogenous growth

Let y be output and consumption (there is no saving). Let

$$z_t = \frac{y_{t+1}}{y_t}$$

be the growth factor of output. Suppose, as Mehra and Prescott (1985) do, that z is first-order Markov:

$$\Pr\left(z_{t+1} \leq z' \mid z_t = z\right) = G\left(z', z\right). \tag{26}$$

Assume, additionally, that $z_t \geq 1$.

4.2.2 Discovery

We now derive (26) via discovery. The following procedure works if z_t is non-decreasing. Let utility be iso-elastic

$$U(y) = \frac{y^{1-\gamma}}{1-\gamma}, \text{ and } y_t = \frac{1}{|x_t - \theta|^k} + \varepsilon_t. \tag{27}$$

where, again, $x_t \in R$ is a decision, $\theta \in \Theta \subseteq R$ an unknown parameter, and ε_t an i.i.d. disturbance. Then

$$U\left(y\right) = -\frac{|x_t - \theta|^{k(\gamma-1)}}{\gamma - 1} + \hat{\varepsilon}_t \tag{28}$$

so that

$$k = \frac{2}{\gamma - 1} \Longrightarrow U\left(y\right) = -\frac{\left(x - \theta\right)^2}{\gamma - 1} + \hat{\varepsilon}_t. \tag{29}$$

Optimal decisions.—Discovery occurs exogenously, independently of the action taken. When (29) holds, it is again optimal to set the decision at the posterior mean of θ. Discovery again occurs exogenously, and we shall assume that a very precise signal of θ is available, as in Example 1, in the sense that T gets large, leading to (11), and (13). Assuming (10) and (29) the optimal policy again given by (14). We shall henceforth send the ε_t to zero in probability and ignore them. As in Example 1, their role is to avoid logical difficulties for the agent. Then output is

$$y_t = \frac{1}{\min_{x \in A_t} |x - \theta^*|^k}$$

Growth.—The growth factor between t and $t+1$ is

$$z_t = \left(\frac{|x_t - \theta^*|}{|x_{t+1} - \theta^*|} \right)^k \tag{30}$$

Since $x_t - \theta^*$ is non-increasing, this construction can work only if $z_t \geq 1$ for all t. Assuming that this is true, we can devise a process for x_t such (30) holds. As in example 1, define ρ again as in (22) to obtain

$$\rho_t = |x_t - \theta^*|$$

as the fraction of the gap between ideal practice, θ^*, and best practice x_t, i.e., the knowledge left undiscovered. Then the stochastic process ρ_t must be non-increasing and satisfies

$$\Delta \ln \rho_t = -\frac{1}{k} \ln z_t \leq 0.$$

If, for example, $z_t \in (z_1, ..., z_N)$ can take on N states and that it is a first-order Markov chain, as in Mehra and Prescott (1985) who assumed that $N = 2$, then $\Delta \ln \rho_t$ is itself a first-order Markov chain taking on values

$$\Delta \ln \rho_t \in \left\{ -\frac{\ln z_1}{k}, ..., -\frac{\ln z_N}{k} \right\}.$$

The process θ_t is once again not uniquely determined, for the same reason that an entire family of densities for θ in (19) is a solution to the search problem.

5 General approach

Can the Bayesian model reproduce the time paths of payoffs (e.g., gambling profits), actions (e.g., decisions of whether to put up the ante and gamble again) and value (the present value of profits)? We seek a Bellman equation such as (3), but with μ replaced by m, and yet where discovery is possible.

Law of motion for m.—We now derive conditions under which m obeys a first-order representation of the form

$$m' = \hat{b}\left(x, y, A', m\right), \tag{31}$$

which generalizes (2) to allow for discovery.

If $A' = A$, $\hat{b}$ is the Bayes map, but when $A' \supset A$, determining $m'\left(A' - A\right)$ in general requires knowledge of the entire sequence $h^t = \left(x^t, y^t\right)$ which would then be used to update m according to (4) with A' in place of A. But (31) will hold if we can invert m_t to obtain all the information about θ that h^t contains. By an application of the inverse function theorem the latter will generically be true if the number of sufficient statistics is less than the number of elements of A.

Let $L_t\left(h^t, \theta\right) = \prod_{s=0}^{t-1} f\left(y_s \mid x_s, \theta\right)$ denote the likelihood function. Assume that under repeated sampling from (1), the information in $\left(x^t, y^t\right)$ has a k-vector of sufficient statistics, $T\left(h^t\right) \in \mathbf{T} \subseteq R^k$. Denote the law of motion for T by $T\left(h^{t+1}\right) = \tau\left(x, y, T\left(h^t\right)\right)$, or simply by[7]

$$T' = \tau\left(x, y, T\right). \tag{32}$$

Since T is sufficient for θ, Fisher-Neyman factorization states that

$$L\left(h^t, \theta\right) = L_1\left(h^t\right) L_2\left(\theta, T\left(h^t\right)\right) \tag{33}$$

for all h^t, and

$$m\left(\theta\right) = \mu\left(\theta \mid T, A\right) = \frac{L_2\left(\theta, T\right) \mu_0\left(\theta\right)}{\int_A L_2\left(\theta', T\right) d\mu_0\left(\theta'\right)} \quad \text{for } \theta \in A \tag{34}$$

Let $\Delta\left(\Theta\right)$ be the set of full-awareness beliefs reachable from μ_0 by observing some feasible $\left(h^t\right)$.[8].

Let $\Delta_k\left(\Theta\right)$ denote the subset of $\Delta\left(\Theta\right)$ entailing a support of at least k elements. These are the feasible beliefs of someone who is aware of at least k distinct θ's. Let $\mathcal{M}_k$ be the set of measures over this set.

Now define the map (34) as

$$m = \phi_A\left(T\right),$$

so that $\phi_A : T \to \Delta\left(\Theta\right)$. The next result concerns the invertibility of this map for A fixed. When can we recover T from m that has a support A?

Proposition 1 *For $m \in \hat{M} \subset \mathcal{M}_k$ as given in (34) assume that the Jacobian of the #A-equation system on the RHS of (34) satisfies*

$$rank\left[\frac{\partial \mu\left(\theta \mid T, A\right)}{\partial T}\right]_{\#A \,\times\, k} = k. \tag{35}$$

[7] Examples are provided below.

[8] Thus $\Delta\left(\Theta\right)$ is the union of the posteriors obtainable at any date from the supports of h^t as h^∞ ranges over its feasible set.

Then $\phi_A^{-1}(m)$ exists on $\hat{M}$ and the representation (31) exists on $\cdot_k(\Theta)$.

Proof. Then the inverse function theorem implies that for each $m \in \hat{M}$ there is a function $\phi^{-1} : \hat{M} \to R^k$ giving $T = \phi_A^{-1}(m)$. Now let

$$\begin{aligned} m'(\theta) &= \mu\left(\theta \mid \tau\left[x, y, \phi_{A'}^{-1}(m)\right], A'\right) \quad \text{for } \theta \in A' \\ &\equiv \hat{b}(x, y, A', m). \end{aligned}$$

which is of the form (31) as desired. ■

Examples are provided in the Appendix.

5.0.3 Discovery and the Generalized Bayes map

It remains for us to specify a process for A which we shall combine with that defined by (31).

The law of motion for A.—Discovery thus should depend on A and on m which reflect the agent's awareness and his beliefs over the models he is aware of, and on the possible conflict between these models and the evidence, y^t– such a conflict presumably stimulates new discovery. And A is the support of m. Therefore, the evolution of A depends on m alone, in addition to possibly depending on x_t. Then, for any $A' \in \mathcal{B}(\Theta)$, let

$$\Pr\left(A_{t+1} = A' \mid x_t = x,\ m_t = m\right) = \alpha(A', x, m). \tag{36}$$

Combining this with (31), we obtain the generalized Bayes operator

$$m' = b(x, y, m),$$

where

$$b(x, y, m) = \int_{\mathcal{B}(A')} \hat{b}(x, y, A', m)\, \alpha(A', x, m)\, dA'. \tag{37}$$

This 'generalized' Bayes map only exists when the likelihood function offers sufficient statistics. We stress it only because it will deliver us a Bellman equation analogous to (3).

Preferences.—The agent's utility is $\sum_{t=0}^{\infty} \delta^t U(x_t, y_t)$ but he does not maximize $E\left\{\sum_{t=0}^{\infty} \delta^t U(x_t, y_t)\right\}$ where E is the standard expectations operator. But this operator does not allow extension of mass to sets of zero measure. For reasons we discussed at the outset, the standard updating of probabilities would preclude the agent from realizing that discoveries are possible and, hence, would exclude any motive for the agent to take actions that would raise the chances of making such discoveries. It also would lead to a different (and probably lower) lifetime utility. The agent will evaluate the current reward using awareness A, but he also will recognize the evolution of A in the sense to be made precise below.

The Bellman equation

We write the Bellman equation corresponding to these preferences. It differs from the standard treatment in just one respect: Instead of using the Bayes map (2), we use the generalized Bayes map (37). Preferences are defined recursively:

$$V(m) = \max_{x \in X} \int_Y [U(x, y) + \delta V(b(x, y, m))] f(y \mid x, \theta)\, dy dm(\theta). \tag{38}$$

This approach has two arguably desirable features

(i) the agent's action be independent of states whose possible existence is not yet discovered.

(ii) he correctly judges the consequences for tomorrow's continuation value of the discoveries that may occur as a result of the actions taken today.

General comparison

The first three examples show that it will not do for the Bayes counterpart to have only y as the signal. We need a proxy for the discovery shock.

	Discovery	Bayes
utility	$U(x,y)$	
likelihood	$f(y \mid x, \theta)$; $\theta \in \Theta$	$g(y, z \mid x, \gamma)$; $\gamma \in \Gamma$
prior	$\mu_0(\theta \mid A)$	$\lambda_0(\gamma)$
updating	$\mu'_{\mid A'} = B(x, y, \mu \mid A')$	
awareness growth	$A' = \alpha\left(x, y, \mu_{\mid A}\right)$	$\lambda' = B(x, y, z, \lambda)$
	$m \equiv \mu_{\mid A}$	
	$b(x, y, m) = \int_{\mathcal{B}(\Theta)} \hat{b} \alpha dA'$	
policy function	$x = h(m)$	$x = H(\lambda)$
value	$V(m)$	$W(\lambda)$

where $x = H(\lambda)$ and W solve,

$$W(\lambda) = \max_{x \in X} \left\{ \int_{\Gamma \times Y \times Z} [U(x, y) + \delta W(B(x, y, z, \lambda))\, g(y, z \mid x, \gamma)]\, dydzd\lambda(\gamma) \right\} \tag{39}$$

Nothing is said about the dimension of Γ relative to that of Θ.

6 Restricting the discovery process

The new primitive concept added in this paper is the discovery process. We showed that once such a process is specified, we can then construct a Bayes model with no discovery that is observationally equivalent. But now we think about some reasonable ways of restricting the discovery process, based on related work which we divide into several topics.

Undirected search.—In this approach the action x_t would be the sampling rate. Conditional on a draw, a value θ' would be drawn from Θ or a subset of Θ. Perhaps new ideas $\theta' \in \Theta$ are generated via a stochastic process that is perhaps influenced by beliefs m (which have support A) and by decisions x^t such as R&D effort. Let $\psi(\theta', \theta)$ be a distribution over θ' conditional on θ which states that if the agent thinks of θ today, then θ' will occur to him tomorrow. Let x_t be the number of ideas θ_i with $i = \{1, 2, ..., x_t\}$ sampled and suppose that each idea is drawn from

$$\text{Prob}\{\theta_{i,t} \leq \theta' \in \Theta \mid m_t\} = \int_A \psi(\theta', \theta)\, dm. \tag{40}$$

Then any $\theta_{i,t} \notin A$ is a discovery. If the process in ψ is highly autocorrelated and if the dispersion of θ' is small, the agent can search only a small neighborhood of A, with most of the ideas sampled being duplicates of old ones. Examples of this approach are Telser (1982), Muth (1986) and Kortum (1997).

Directed search.—Bayesian treatments of the bandit problem and its various elaborations are in Rothschild (1974), Jovanovic and Rob (1990) and Jovanovic and Nyarko (1996). There is no discovery in these models in the sense that the parameter space is known from the outset.

Hypothesis testing and model revision.—Discovery could, however, be more directed, such as the algorithms for model revision proposed by Sims (1971) and more formally by Radner (2002). The agent from time to time enlarges the set of model parameters that he wishes to entertain. One may limit discoveries to those that improve the agent's understanding of the world according to some statistical criterion. He may assign zero probability to any new idea that seems improbable enough in light of the evidence. For example, the discovery it should pass a likelihood-ratio test. Thus one could constrain

$$\text{Admissible discoveries} = \left\{\theta' \in \Theta \mid \frac{L(h^t, \theta')}{\max_{\theta \in A} L(h^t, \theta)} \geq \lambda\right\}. \tag{41}$$

The parameter $\lambda \geq 0$ indicates the height of the hurdle that a discovery must clear. Indeed, if $\lambda > 1$, with probability 1 discovery must stop short of the true θ. For any λ, the process is in line with the view of Galanis (2008 Section 6, page 17) argues that awareness grows when observing something which you thought was impossible. We can think of "impossibility" as a zero denominator.

Let us describe Radner (2002) in more detail. Hew proposes a model-revision procedure. Let N and k be positive integers and let $M(k)$ be the set of N-state Markov chains of order k. Let $\Theta = \cup_{k=0}^{\infty} M(k)$ and $A_k = \cup_{j=0}^{k} M(j)$. Radner assumes that A_k is augmented to A_{k+1} at a pre-specified date t_k where (t_k) is a sequence increasing in k. Since the dimension of the parameter space increases geometrically with k, $t_{k+1} - t_k$ must also increase with k if the agent is to get to learn much of anything. Radner's term 'model revision' is equivalent to a 'discovery' in our sense. Let us relate this to eq. (4) in particular. Now $M(k)$ can be considered as a zero-measure subset of $M(k+1)$. and $\Delta(M(k))$ the set of measures over $M(k)$. Then letting $\mu_k \in \Delta(M(k))$, let the full-awareness prior be

$$\mu_0 = \sum_{k=0}^{\infty} \alpha_k \mu_k.$$

Substituting this expression for μ on the RHS of (4) leads to

$$\mu_{A_k}(\theta) = \frac{\sum_{j=0}^{k} \alpha_j \mu_j(\theta)}{\int_{A_k} \sum_{j=0}^{k} \alpha_j d\mu_j(\theta)} \quad \text{for } \theta \in A_k, \quad \text{and } \mu_{A_k}(\theta) = 0 \qquad \text{for } \theta \notin A_k.$$

Game theory and decision theory.—When revelations of elements of $\Theta - A$ are made strategically, one can specify the subjective probability $\hat{\phi}$ using information about the opponent's incentives to reveal various states, as Ozbay (2008) argues. Kochov (2009) provides axioms that one would find reasonable in the context of unawareness, and discusses optimal decisions that way.

7 Conclusion

The paper defined discovery as an extension of the support of beliefs. It then showed that one can approximate discovery-induced beliefs by a Bayesian belief mechanism featuring precise signals pointing to events having small but positive probability. This made it possible to define optimal decisions and rational valuations in environments with discovery.

All this was shown in the context of several examples, and still to be derived is a general result concerning the correspondence between models of discovery and differently parametrized models with no discovery.

References

[1] Aghion Philippe and Peter Howitt. A Model of Growth Through Creative Destruction. *Econometrica* 60, No. 2. (Mar., 1992), pp. 323-351.

[2] Auerswald, Philip, Stuart Kauffman, José Lobo and Karl Shell. "The production recipes approach to modeling technological innovation: An application to learning by doing." *Journal of Economic Dynamics and Control* 24, no. 3 (March 2000): 389-450.

[3] Bramoullé, Yann and Gilles Saint Paul. "Research Cycles." 2007.

[4] Burdett, Kenneth. "A Theory of Employee Job Search and Quit Rates." *American Economic Review* 68, no. 1 (March 1978): 212-20.

[5] Cogley, Timothy and Thomas J. Sargent. "The Conquest of US inflation: Learning and Robustness to Model Uncertainty." *Review of Economic Dynamics* 8 (2005) 528–563.

[6] Cho, In-Koo & Kenneth Kasa; Simon Fraser University "Learning About Identification." University of Illinois 2009.

[7] Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini, "Standard State-Space Models Preclude Unawareness." *Econometrica* 1998, 66 (1), 159-173.

[8] Easley, David and Nicholas Kiefer. "Controlling a Stochastic Process with Unknown Parameters." *Econometrica* 56 (1988): 1045-64.

[9] Epstein, Larry and Stanley Zin. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57, No. 4 (Jul., 1989): 937-969.

[10] Evans, George and Seppo Honkapohja. *Learning and Expectations in Macroeconomics*, Princeton University Press,Princeton, NJ, 2001.

[11] Galanis, Spyros. "Trade and the Value of Information under Unawareness."

[12] Griliches, Zvi. *R&D, Education, and Productivity: A Retrospective.* Harvard University Press, 2000.

[13] Jovanovic, Boyan and Yaw Nyarko. "Learning by Doing and the Choice of Technology," *Econometrica* 64, no. 6 (December 1996): 1299-1310.

[14] Jovanovic, Boyan and Rafael Rob. "Long Waves and Short Waves: Growth through Intensive and Extensive Search." *Econometrica* 58, no. 6 (December 1990): 1391-1409.

[15] Kocherlakota, Narayana. “Model Fit and Model Selection.” *Federal Reserve Bank of St. Louis. Review* July/August 2007, 89(4), pp. 349-60.

[16] Kochov, Asen. “A Model of Limited Foresight.” U. Rochester, June 2009

[17] Kortum, Samuel. “Research, Patenting, and Technological Change.” *Econometrica* 65, no. 6 (1997): 1389-1419.

[18] Li, Jing. “A Note on Unawareness and Zero Probability.” PIER WP 08-022. January 2008.

[19] Mehra, Rajnish and Edward C. Prescott. “The Equity Premium: A Puzzle.” *Journal of Monetary Economics* 15, no 2 (March 1985): Pages 145-161

[20] Muth, John F., “Search Theory and the Manufacturing Progress Function.” *Management Science* (August 1986), pp. 948-962

[21] Nyarko, Yaw. “Learning in Mis-specified Models and the Possibility of Cycles.” *Journal of Economic Theory* 55, no. 2 (December 1991): 416-427.

[22] Ozbay, Erkut. “Unawareness and Strategic Announcements in Games with Uncertainty.” UMD February 2008.

[23] Pakes, Ariel. “On Patents, R&D, and the Stock Market Rate of Return.” *Journal of Political Economy* 93, no.2 (April 1985): 390-409.

[24] Radner, Roy. “Bayesian Analysis and Model Revision for k’th Order Markov Chains with Unknown k.” Stern School, NYU, October 2002.

[25] Romer, Paul. “Endogenous Technological Change.” *Journal of Poliical Economy* 1990.

[26] Rothschild, Michael, 1974. “Searching for the Lowest Price When the Distribution of Prices is Unknown.” *Journal of Political Economy* 82, no. 4 (July 1974): 689-711.

[27] Sims, Christopher. “Distributed Lag Estimation When the Parameter Space is Explicitly Infinite- Dimensional.” *Annals of Mathematical Statistics* 42, No. 5 (October 1971): 1622-1636.

[28] Telser, Lester. “A Theory of Innovation and Its Effects.” *Bell Journal of Economics* 13, no. 1 (1982): 69-92.

[29] Venezia, Itzhak. “On the Statistical Origins of the Learning Curve.” *European Journal of Operations Research*, 19 (1985): 191-200.

[30] Young, Peyton. *Strategic Learning and its Limits*. Oxford University Press 2002.

7.0.4 Appendix: Examples of (32)-(35)

TO BE COMPLETED

Example 1: *Normal distribution.*—Let $y = x + \theta + \varepsilon$ and let $\varepsilon \sim N\left(0, \sigma_\varepsilon^2\right)$ so that

$$p\left(y \mid x, \theta\right) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\left(y - x - \theta\right)^2\right).$$

and suppose that σ_ε is known. Let $\bar{y} - \bar{x} \equiv z$ be the sample mean of t x-adjusted signals. Then let $T \equiv (z, t) \in R^2$ be the sufficient statistic for θ. Thus $k = 2$. The transition function (32) for T is

$$T' \equiv \tau\left(x, y, T\right) = \begin{pmatrix} \frac{1}{t+1}\left(tz + y - x\right) \\ t + 1 \end{pmatrix}.$$

Example 2: *Binomial distribution.*—Let x be constant (i.e., no decision) and let (1) be given by

$$y = \begin{cases} 1 & \text{with Prob } \theta \\ 0 & \text{with Prob. } 1 - \theta \end{cases}$$

with $\theta \in [0, 1]$. Then in t trials and k successes the likelihood is $\theta^k \left(1 - \theta\right)^{t-k} \equiv L_2\left(\theta, T\right)$, where $T = (k, t)$. The transition (32) reads

$$T' \equiv \tau\left(y, T\right) = \begin{pmatrix} k + I_{\{y=1\}} \\ t + 1 \end{pmatrix},$$

where I is the indicator function.