

September 17, 2008  
Larry Karp  
ARE 263

## 1 Anticipated Learning

"Active learning" describes the situation where the decisionmaker (DM) manipulates a state variable in order to gain information. For example, a monopoly might believe that the stock of loyal customers (a state variable) changes as a result of advertising, but not know the exact relation between advertising and the change in the stock, i.e. the monopoly might not know a parameter value in the equation of motion. The monopoly might choose its advertising level (a control variable) over time in order to learn about the unknown but fixed parameter value. The model in the Clarke and Mangel, where an animal decides which patch to search for food, contains another example of active learning.

"Passive learning" describes the situation where learning is exogenous to the DM's actions. For example, a DM will not manipulate the stock of GHG in order to learn more about the relation between these stocks and climate change. I will only discuss passive learning; but see section 1.3.3.

Modeling learning requires using a state variable to describe information. I will discuss two types of approaches, using either a discrete or a continuous distribution. In both cases I use examples rather than a general framework, in order to make the ideas as clear as possible and keep notation to a minimum.

Throughout this section, a central assumption is that the DM *anticipates* learning. If the DM happens to learn, but does not anticipate that he will learn in the future, the problem is very different (and much simpler and less interesting).

### 1.1 Discrete distribution

Suppose that the single period payoff is

$$U(c) - \Delta d(S)$$

Where  $S$  is the stock of GHG (a state variable) in the current period and  $c$  is emissions of GHG (the control variable) in a period. The benefit of emissions is  $U$  and the damages associated with the stock is  $\Delta d$ , where  $d$  is a known function and  $\Delta$  is unknown. Section 1.2.1 considers the case where  $\Delta$  is a random variable, and we learn about its distribution. Section 1.2.2 considers the case where  $\Delta$  is a fixed parameter and we obtain information about the value of this parameter.

The stock of GHGs evolves according to

$$S' = f(S, c) \tag{1}$$

where  $f$  is the growth equation and  $S'$  is the stock in the next period. Recall that I use the convention that a variable without a time subscript is the value of the variable in the current period, and the variable with a prime is the value of that variable in the next period.

## 1.2 Learning about $\Delta$

First, I will discuss two ways to think about learning. We can either assume that  $\Delta$  is a random variable and we learn about its distribution, or we can treat  $\Delta$  as a fixed but unknown number, which we learn about over time. Then I will consider two ways of solving the problem (using either model of learning): either dynamic programming or stochastic programming.

### 1.2.1 $\Delta$ is a random variable with unknown distribution

First, suppose that we treat  $\Delta$  as a random variable. We do not know the distribution of  $\Delta$  but (for the purpose of modeling) we think that it is reasonable to assume that it is a draw from one of two distributions. The problem is to choose the control rule for emissions, taking into account that in the future we will have better information than today concerning which of the two possible distributions is correct.

Denote these two distributions as  $x_1$  and  $x_2$ . Associated with each distribution there are two possible outcomes,  $G$  and  $B$ . Under the first distribution the probability is  $q$  that the realization of  $\Delta$  is  $G$ ; under the second distribution the probability is  $r$  that the realization of  $\Delta$  is  $G$ .  $G$  and  $B$  are numbers, with  $G < B$ . The two realizations correspond to low damage ( $G$ ) and the high damage ( $B$ ) outcomes.

Table 1 gives the outcomes and probabilities associated with these two distributions. These are *conditional probabilities*, the probability of an event conditional on  $x$ .

realization of $\Delta$	$x = x_1$	$x = x_2$
$G$	$q$	$r$
$B$	$1 - q$	$1 - r$

Table 1: The conditional distributions of random variable  $\Delta$

If  $q < r$  then GHG stocks present a greater danger if the "truth" is  $x = x_1$  rather than  $x = x_2$ .

For the purpose of the model,  $q$  and  $r$  are taken to be objective probabilities. The DM is uncertain which distribution is correct, and at time  $t$  has subjective probability  $p_t$  that  $x = x_1$ . The subjective probability at time  $t$  that  $x = x_2$  is therefore  $1 - p_t$ . Denote  $P(p_t, \Delta_t)$  as the posterior probability, the subjective probability that  $x = x_1$  when the prior is  $p_t$  and you observe  $\Delta = \Delta_t$ :

$$p_{t+1} = P(p_t, \Delta_t) = \Pr \{x = x_1 \mid \Delta_t, p_t\}. \quad (2)$$

Using Bayes' Rule we can write the values of  $P(p_t, \Delta_t)$  for the two possible realizations of  $\Delta$ :<sup>1</sup>

$$P(p_t, G) = \frac{p_t q}{p_t q + (1 - p_t) r} \quad (4)$$

$$P(p_t, B) = \frac{p_t (1 - q)}{p_t (1 - q) + (1 - p_t) (1 - r)}. \quad (5)$$

*In this model, the subjective probability,  $p$  is a state variable, the equation of motion of which is given by equation 2.* Note that the evolution of  $p$  is stochastic, since the evolution depends on the realization of a random variable,  $\Delta$ . Also note that the evolution does not depend on actions that the DM takes; learning is passive.

---

<sup>1</sup>To obtain equation 4 and 5, use the rule

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

to write

$$P(B \mid A) = \frac{P(A \mid B) P(B)}{P(A)}. \quad (3)$$

Associate the event  $A$  with  $\Delta_t = G$  and the event  $B$  with  $x = x_1$ . Equations 4 and 5 then follow directly from the formula 3.

Increasing the number of possible outcomes of  $\Delta$  does not greatly increase the size of the problem – it just means that we have more possible outcomes, i.e. more equations of the form of equations 4 and 5. In contrast, increasing the number of possible distributions increases the dimensionality of the state variable, and significantly increases the size of the problem (which has important implications on the feasibility of obtaining a numerical solution). If there are  $n$  possible distributions you need  $n - 1$  state variables to describe beliefs; each state variable is the subjective probability that a particular distribution describes the world. Since the probabilities sum to 1, you only need  $n - 1$  numbers to keep track of the  $n$  probabilities.

### 1.2.2 The "star information structure"

Kolstad (JEEM 1996) uses an alternative called the "star information structure". In this setting  $\Delta$  is a parameter (rather than a random variable) that takes a particular value, either  $G$  or  $B$ , but the DM does not know which value it takes. Let  $g$  be a signal that makes me think it is more likely that  $\Delta = G$ , and  $b$  be a signal that makes me think it is more likely that  $\Delta = B$ . If at time  $t$  my subjective probability that  $\Delta = G$  is  $p_t$ , then if I (the DM) observe the signal  $g$  I update my probability according to

$$p(p_t, g) = \lambda + (1 - \lambda) p_t.$$

If I observe the signal  $b$  I update my probability to

$$p(p_t, b) = 0\lambda + (1 - \lambda) p_t.$$

My updated subjective probability is a convex combination of the previous probability and the state of being certain ( $p = 1$  or  $p = 0$ ). If  $\lambda = 1$  then after one observation I know the true value of  $\Delta$ . If  $\lambda = 0$  my observation in the current period provides no information, and I never change my subjective beliefs. As  $\lambda$  ranges from 0 to 1, the signal becomes more informative.

Figure 1 represents the star information structure when there are three (rather than two) possible values of  $\Delta$ , call them  $G$ ,  $B$ ,  $U$  and let  $p_j$   $j = 1$  for  $G$ ,  $j = 2$  for  $B$ , and  $j = 3$  for  $U$  be the subjective prior probabilities of each of the outcomes. A point such as "A" in the simplex represents a particular prior distribution. If, for example,  $\lambda = \frac{1}{2}$  as in the Figure, then if my signal in the present period is  $g$  then "I move" half way from my prior probability, point A, to being certain that  $\Delta = G$ . This certainty is represented by the

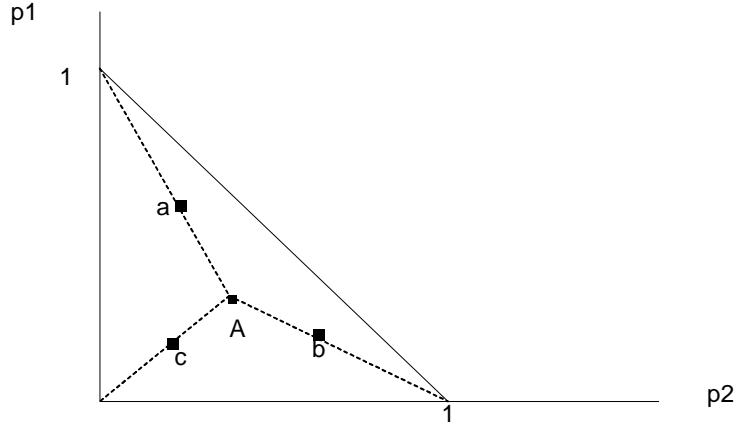


Figure 1: The "star" when there are three possible values and  $\lambda = \frac{1}{2}$

point  $(0,1)$  in the simplex. The point "a" in the figure is half way between the point  $A$  and  $(0,1)$ ; point "a" is my posterior if  $\lambda = \frac{1}{2}$ , my prior is "A" and my signal is  $g$ . Similarly, if my signal in the current period is  $b$  my posterior probability is point "b" and if the signal is  $u$  my posterior is point "c".

### 1.2.3 Comparing the two ways of modeling learning.

The star information structure has the advantage that it uses a single parameter,  $\lambda$ , to model the speed of learning. The disadvantage of that approach is that it does not permit a clear distinction between risk (the fact that there is objective randomness) and uncertainty (the fact that you do not know the probability distribution of a random variable). If  $\Delta$  is really a parameter that takes a single value (rather than a r.v.), then one observation should be enough to learn that value, unless there is some other randomness – but that additional randomness is not modeled. For example, as  $p$  approaches either 0 or 1, the amount of randomness goes to 0. It seems (to me) that the difficulty of learning arises because I get noisy signals, i.e. because the relation between damages and the stock of GHGs really is stochastic, not

merely uncertain. As my uncertainty about this relation diminishes, there remains the objective stochasticity. However, the star information structure apparently assumes that as my uncertainty about the relation diminishes, the objective stochasticity also diminishes – and this strikes me as unreasonable.

In contrast, the specification of underlying distributions recognizes that there is exogenous randomness; this is what makes it difficult to learn the true distribution. Even after the DM knows the true distribution, i.e. as  $p$  approaches 0 or 1, there is still randomness. The disadvantage of this approach is that it is less parsimonious – the modeler has to specify both  $q$  and  $r$ , instead of the single parameter  $\lambda$ . I find this approach more appealing than the star information structure, so I use it below.

### 1.3 Describing the optimization problem

There are two approaches to describing – and then solving – the optimization problem. You can write it as a dynamic programming (DP) or as a stochastic programming (SP) problem. I consider each of these in turn.

#### 1.3.1 The DP problem

I will consider an autonomous control problem: there is an infinite horizon and no explicit time dependence, apart from discounting at a constant discount factor  $\beta$ . In this problem there are two state variables,  $S$  and  $p$ . (I assumed that  $S$  evolves deterministically, but it is straightforward to add a random component to equation 1. This change requires that in solving the problem you take expectations with respect to the additional random component.)

Denote the value function as  $J(p, S)$  and write the DPE as

$$J(p, S) = \max_c E_{\Delta} \{U(c) - \Delta d(S) + \beta J(p', S')\}$$

subject to equations 1 and 2.

In taking expectations of  $\Delta$  I use the subjective distribution:

$$\begin{aligned} \Pr(\Delta = G) &= pq + (1 - p)r \\ \Pr(\Delta = B) &= 1 - [pq + (1 - p)r]. \end{aligned}$$

The realization of  $\Delta$  determines not only this period's damages, but also the subjective beliefs in the next period,  $p'$ .

Because  $\Delta$  appears linearly in the current period payoff, I can write

$$E\Delta = G(pq + (1 - p)r) + (1 - [pq + (1 - p)r])B. \quad (6)$$

Given  $p$ , there are only two possible values of  $p'$  (depending on whether the current realization of  $\Delta$  is  $G$  or  $B$ ). I can list the two possible outcomes of  $p'$ , say  $p'_G$  and  $p'_B$ , corresponding to the outcomes  $G$  and  $B$  in the current period, and write

$$EJ(p', S') = [pq + (1 - p)r]J(p'_G, S') + [1 - [pq + (1 - p)r]]J(p'_B, S'). \quad (7)$$

Using equations 6 and 7 I can remove the expectations operator from the DPE.

**Digression: anticipation** Note that if the DM either does not learn, or does not anticipate learning, the problem becomes much simpler. In that case, we have a single state variable,  $S$ , and we treat  $p$  as a parameter. Now the DPE is

$$J(S) = \max_c \{U(c) - (E\Delta)d(S) + \beta J(S')\} \quad (8)$$

subject to equations 1,

where equation 6 gives the formula for  $E\Delta$ . If the DM really never learns, equation 8 is the DPE in every period.

Perhaps the regulator learns about  $\Delta$  but simply does not anticipate learning. In that case, he still solves the DPE 8, but (to the DM's surprise) he has a different value of  $p$  in each period, as he acquires information. My point here is that it is not simply the fact of learning, but the anticipation of learning that determines the nature of the problem that the DM solves.

**Digression: a different state variable** Suppose that the problem is non-autonomous, so we have some reason to keep track of calendar time. For example, there may be a finite horizon, or maybe there is some exogenous (e.g. technical) change that we need to keep track of. In this situation, we have to add the argument  $t$  to the value function. Set the initial time equal to 0 so  $t$  gives both calendar time and the number of times I have observed the realization of  $\Delta$ .

In this case, an alternative to using  $p$  as the state variable is to use  $n =$  number of times I observed the realization  $\Delta = G$ . If I started out with

subjective probability  $p_0$  that  $x = x_1$  and  $n$  times I observed  $\Delta = G$ , and  $t - n$  times I observed  $\Delta = B$ , then I can use Bayes' Rule to calculate my current subjective probability  $p_t$ .

In other words, it does not matter whether I treat the arguments of the value function as  $S, p, t$  or  $S, n, t$ . In both cases I need to keep track of three arguments. (See the reading by Clarke and Mangel for an example of the latter approach.)

However, consider the following minor change in the problem. Suppose that there are 3 possible realizations<sup>2</sup> of  $\Delta, G, B, U$ ; as before, there are only two possible distributions –  $x$  takes one of only two values. If I treat  $p$  as my state variable, the addition of a third outcome to  $\Delta$  means that I need one more equation for  $p'$  (like equations 4 and 5), so I need to do a few more calculations when I compute the expectations. However, I do not need to increase the number of state variables. We will see that the numerical complexity depends on the number of state variables. In contrast, if I decide to keep track of outcomes, I now need to keep track of  $n$ , the number of times I observed  $G$ , and  $m$ , the number of times I observed  $B$  (so that I can calculate  $t - n - m$ , the number of times I observed  $U$ ). In this case, the minor change increases the dimension of the state variable – and that is a headache.

The point here is that often there are different ways to define the state variable. You want as parsimonious a representation as possible.

### 1.3.2 The SP problem

The SP problem requires keeping track of "histories". A history is a sequence of outcomes. Suppose the initial period is time  $t = 0$ . By definition, there are no histories at  $t = 0$ . At  $t = 1$  there are two possible histories  $\{G\}$  and  $\{B\}$ ; at  $t = 2$  there are 4 possible histories,  $\{G, G\}, \{G, B\}, \{B, G\}$ , and  $\{B, B\}$ . Define a particular history at time  $t$  as  $h_t$ .

You can calculate the subjective probabilities of each of these histories. For example, at time 0 the subjective probability of history  $\{G\}$  is simply  $p_0$ , the initial subjective probability. The subjective probability of  $\{G, G\}$  is  $p_0 P(G, p_0)$ , where I use equation 4 to calculate  $P(G, p_0)$ . Using calculations of this sort I can calculate the probability of any history. Define this probability as  $\mu(h_t, t)$ . For each time  $t$  and history  $h_t$  I can also calculate my subjective probability that  $\Delta = G$ ; denote this probability as  $\pi(h_t, t)$ .

---

<sup>2</sup>Any guesses for what  $U$  stands for?

For example, I can calculate my posterior probability that  $x = x_1$  after history  $h_t$ ; denote this posterior as  $p_t = P(h_t, p_0)$  to reflect the fact that the posterior depends on both my initial prior  $p_0$  and the history  $h_t$ . With this notation

$$\pi(h_t, t) = p_t q + (1 - p_t) r.$$

The SP is to maximize the expectation, over “future histories”, of the present value of the sequence of future payoffs. The SP problem is

$$\max_{\{c(h_t, t)\}} \sum_{t=0}^{\infty} \beta^t \left[ \sum_{h_t} \mu(h_t, t) \left( U(c) - \underline{[\pi(h_t, t) G + (1 - \pi(h_t, t) B) d(S_t)]} \right) \right]$$

subject to equation 1.

The underlined term gives expected damages in each period conditioned on the history. Note that the optimization chooses a level of emissions for each possible history, for each time.

In most cases we are really interested in behavior only during the next few periods (each of which might consist of a decade). For most problems, the decision rules are insensitive to the choice of final horizon,  $T$ , provided that  $T$  is sufficiently big; this is because with a discount factor  $\beta < 1$ , the very distant future does not matter much. Therefore, there is little cost to realism in replacing the infinite horizon with a finite horizon, so that you have a finite number of possible histories, and thus a finite number of choice variables  $c(h_t, t)$ . In this case, the SP problem can be solved as a non-linear programming problem, e.g. by using GAMS. Computational constraints may still limit the number of possible histories.

The advantage of SP over DP is that you can solve the former using a program like GAMS, whereas the DP generally requires some programming. As a practical matter, the DP approach can handle larger problems (more histories).

### 1.3.3 Digression: active and passive learning again

In the models above, the DM does not manipulate the state variable in order to learn; in that sense the learning process is exogenous, i.e. learning is passive. However, this model can be used to study other types of learning that are more relevant to environmental problems than active learning.

For example, in the Star Information Structure, the parameter  $\lambda$  measures the speed of learning. It might be possible to increase  $\lambda$  by paying a cost.

By solving the problem for different values of  $\lambda$  we can see how  $\lambda$  affects the value of the payoff, and in this way determine how much we should spend to increase  $\lambda$ .

In the model with explicit distributions, learning can occur more rapidly if we obtain better signals. For example, the outcome  $G$  might include "very good" ( $VG$ ) and "etzi-ketzi" ( $EK$ ). We can compare the value of the program when we observe only  $G$  and  $B$  with the value when we observe  $VG$ ,  $EK$ , and  $B$ . We can compare the gain in value, due to the more precise signals (and faster learning) with the cost of more precise signals.

The point of these examples is that we can use models of passive learning to think about certain kinds of activities that change the speed of learning.

## 1.4 Continuous distribution

The model with discrete distributions is attractive because it permits a very general representation of the different possibilities. The disadvantage of that model is that it is not parsimonious. As I noted above, if I have  $n$  possible distributions, I need  $n - 1$  state variables to represent my current information; each of these state variables is my subjective probability that one of the possible distributions is the correct distribution.

An alternative uses continuous distributions. This alternative is more flexible in that it allows infinitely many rather than several possibilities. However, it restricts each of these infinitely many possibilities to be a member of a certain family. In addition, this modeling approach requires that my prior distribution and likelihood function are conjugates – meaning that my posterior distribution has the same form as my prior.

### 1.4.1 Example 1: Poisson and gamma distributions

This example uses the climate change problem. I have chosen the distributions to illustrate the methods, not because I necessarily think that they provide a good representation of the climate change problem.

Suppose that my single period payoff is

$$U(c) - \Delta d(S),$$

where  $c$  is emissions,  $S$  is the stock of GHGs,  $U$  and  $d$  are known functions, and  $\Delta$  is a random variable. In each period I observe actual damages  $\Delta d(S)$

and I know the function  $d$ , so I can calculate the realization of  $\Delta$  in that period.

For the purpose of obtaining a tractable model, I will suppose that I know that  $\Delta$  has a Poisson distribution, with parameter  $\lambda$ :

$$\Pr(\Delta = s \mid \lambda) = \frac{e^{-\lambda} \lambda^s}{s!},$$

where  $s$  is an element of the set of positive integers. With this distribution,

$$E\Delta = \text{var}(\Delta) = \lambda.$$

I model uncertainty by assuming that I do not know the true value of  $\lambda$ . For example, if  $\lambda$  is a very small number, then the expected damage of  $S$  is small; but if  $\lambda$  is large, expected damages are large. I model my uncertainty about  $\lambda$  by choosing my prior distribution to be gamma with parameters  $r$  and  $t$ :

$$\Pr(\tilde{\lambda} = \lambda \mid r, t) = \frac{e^{-\lambda t} (\lambda t)^{r-1} t}{(r-1)!}.$$

(More precisely, this is a special form of the gamma distribution, called the Erlang distribution.) Using this density,

$$E\tilde{\lambda} = \frac{r}{t} \quad \text{and} \quad \text{var}(\tilde{\lambda}) = \frac{r}{t^2}.$$

The choice of initial values of  $r$  and  $t$  provides considerable flexibility in modeling my uncertainty about  $\lambda$ . Suppose that my current values of the parameters are  $r, t$  and in this period I observe a particular value of  $\Delta$ ; I noted above that I can calculate this realization because I observe actual damages and know the function  $d$  and the value of  $S$ . Due to the fact that the gamma and Poisson are conjugates, my posterior on  $\lambda$ , after observing  $\Delta$ , is also a gamma with parameters

$$\begin{aligned} r' &= r + \Delta \\ t' &= t + 1. \end{aligned} \tag{9}$$

Note that the evolution of the parameter  $r$  is stochastic, because it depends on the realization of  $\Delta$ , a random variable. However, the evolution of  $t$  is deterministic. This fact is useful, because it means that I can include non-stationarity (e.g. technical change) in the problem without increasing the dimension of the state variable.

For this problem, I need three state variables,  $S$ ,  $t$ , and  $r$ . The expectation of the single period payoff, given  $r, t$ , and  $S$  is

$$U(c) - d(S)E_\lambda (E_{\Delta|\lambda}) \Delta = U(e) - d(S)\frac{r}{t}.$$

The fact that  $\Delta$  enters the problem linearly leads to a particularly simple expression for the expected single period payoff. If  $\Delta$  had entered the payoff non-linearly, the expression for the expected single period payoff would have required more calculation, but that is really a fairly minor issue, since I will have to use numerical methods to solve this problem in any case.

The DPE is

$$J(S, r, t) = \max_c \left[ U(c) - d(S)\frac{r}{t} + \beta E_\Delta J(S', r + \Delta, t + 1) \right]$$

subject to equation 1.

The formula for the last expectation is

$$E_\Delta J(S', r + \Delta, t + 1)$$

$$\int_0^\infty \left( \sum_{s=0}^{s=\infty} \frac{e^{-\lambda} \lambda^s}{s!} \right) \left( \frac{e^{-\lambda t} (\lambda t)^{r-1} t}{(r-1)!} \right) J(S', r + s, t + 1) d\lambda.$$

The numerical solution to this problem is complicated by the fact that  $r$  and  $t$  increase without bound. To solve it numerically one would have to truncate the state space. For example, suppose that learning occurs only a finite number of times; in this case,  $t$  is finite. Rather than letting  $r$  be any positive integer, approximate the Poisson distribution by assuming that it takes one of a finite number of values. These numerical issues are important, but not the focus of this lecture.

Once again, it is worth pointing out that a fundamental feature of this problem is the *anticipation* of learning, not simply the fact of learning. If I do not anticipate learning, I can treat  $E\Delta$  as a parameter, rather than a function of the state variable. Without anticipated learning, the state variable is simply  $S$  and the DPE is

$$J(S) = \max_c [U(e) - d(S)E_\Delta + \beta J(S')]$$

subject to equation 1.

The fact that I anticipate learning greatly complicates the problem, because it means that I have to keep track of a larger state variable (one consisting of 3 rather than 1 variable).

### 1.4.2 Example 2: LQ functional forms and normal distributions

This example actually makes two points. First, I want to show how you can make some headway using the LQ functional form. Second, I want to show how you can calibrate a model if you additionally assume normal likelihood and priors.

**The LQ functional form** I will extend the LQ model considered in an earlier set of notes, by allowing the damage parameter to be uncertain. This model can also include asymmetric information about costs. (See Karp and Zhang JEEM 2005, the source of this section of the notes). I ignore the problem of asymmetric information here.

Let the benefit of emitting at rate  $x$  be

$$f + ax_t - \frac{b}{2}x_t^2.$$

The stock grows according to

$$S_{t+1} = \delta S_t + x_t$$

where  $0 < \delta < 1$  is a known parameter.

The damages associated with a stock of  $S_t$  are

$$D(S_t, \omega_t; G^*) = \frac{G^*}{2} (S_t - \bar{S})^2 \omega_t.$$

Here  $G^* > 0$  is a fixed but unknown number and  $\omega_t$  is the realization of a random variable with known distribution. In each period you can observe damages and you know the stock  $S_t$  and the parameter  $\bar{S}$ . This information allows you to calculate  $G^* \omega_t$ , but you are not able to infer the fixed value of  $G^*$ . For example, you cannot tell whether high damages were a result of a high realization of  $\omega$  or a high value of  $G^*$ . However, after observing a succession of damages, you begin to get a better idea of the level of  $G^*$ .

In order to be able to use standard dynamic programming methods, we need to be able to describe the subjective distribution of  $G^*$  using a finite number of parameters. Those parameters are elements of the state vector. I will suppose here that the subjective distribution of  $G^*$  at time  $t$  is defined by two moments, the mean and variance,  $\chi_t \equiv (G_t, \sigma_{G,t}^2)$ . However, you can think of  $\sigma_{G,t}^2$  either as a vector of higher moments or a scalar (the variance).

The regulator cannot predict his future subjective expectation, so his current subjective expectation is an unbiased estimator of its future value, i.e.  $E_t G_{t+\tau} = G_t$  for  $\tau \geq 0$ .

The form of my prior and my likelihood (assuming that they are conjugates) determine how I update the state  $\chi$ . That is, the prior and the likelihood function lead to a system of equations analogous to equation 9. At least one equation in this system must be stochastic, to reflect the fact that I do not know what information I will get, so I do not know what my beliefs will be in the future.

For the time being, put aside the question of the exact updating system of equations for  $\chi$ . Assume that I have such a system. Write  $\bar{\omega} = E_\omega \omega$ . Since, by assumption, I know the distribution of  $\omega$ , I treat  $\bar{\omega}$  as a fixed parameter. The DPE is

$$J(S, \chi) = \max_x \left[ f + ax - \frac{b}{2}x^2 - \left( E_{G^*|\chi} \frac{G^*}{2} \right) (S_t - \bar{S})^2 \bar{\omega} + \beta E_{\chi'|\chi} J(S', \chi') \right].$$

A nice feature of this problem is that the value function is linear-quadratic in  $S$ . That is,  $J$  has the form

$$J(S, \chi) = \xi(\chi) + \gamma(\chi)S + \frac{\rho(\chi)S^2}{2}.$$

In the LQ problem without learning, we saw that the value function also has this form, but in that case  $\xi$ ,  $\gamma$ , and  $\rho$  are numbers. Here, with learning, they are functions of the information,  $\chi$ . These functions can be calculated numerically.

However, we can obtain some insight even without performing this calculation. In particular, consider the problem with learning where our current subjective expectation of  $G^*$  is  $G_t = G$ , where  $G$  is just some number. Compare this to the problem in which we never expect to learn (perhaps because we are certain that we know the value, or because we are naive, or simply because we never do learn anything). In this problem without learning, our expectation of  $G^*$  is also  $G$ . Obviously, the two problems are identical except that we expect to change our mind about the value of  $G^*$  under learning, but not in the absence of learning.

We can show analytically that for any value of  $G$  and  $S$ , the optimal level of emissions is higher when we anticipate learning, compared to when we

expect never to learn. This result does not depend on the updating rule for information. Anticipated learning makes us less cautious because we know that if we get bad news in the future (i.e., learn that  $G^*$  is probably greater than we originally thought) we can modify our actions in the future. This option to act on future information makes bad news "less bad".<sup>3</sup>

**The assumption of normality** In order to calibrate a model for greenhouse gasses, we need an explicit learning rule. One alternative is to assume that the distribution of the damage shock is lognormal:

$$\omega_t \sim i.i.d. \text{ lognormal} \left( -\frac{\sigma_\omega^2}{2}, \sigma_\omega^2 \right). \quad (10)$$

We express the subjective moments in terms of  $g \equiv \ln G$ . The regulator begins in period  $t$  with normal priors on  $g^* = \ln G^*$ , with mean  $g_t$  and variance  $\sigma_{g,t}^2$ :

$$g^* \sim N(g_t, \sigma_{g,t}^2). \quad (11)$$

Given distribution (11), the subjective distribution of  $G^*$  is log-normal with

$$E_t G^* \equiv G_t = \exp \left( g_t + \frac{1}{2} \sigma_{g,t}^2 \right) \quad (12)$$

$$\sigma_{G,t}^2 \equiv \text{var}_t(G^*) = \exp(2g_t + \sigma_{g,t}^2) (\exp(\sigma_{g,t}^2) - 1).$$

Since damages are a product of independent log-normally distributed variables, the regulator has log-normal priors on damages. After observing damages and the current stock, the Bayesian regulator updates his belief about  $g^*$ . The moment estimator of  $g^*$ , denoted  $\hat{g}_t$ , is

$$\hat{g}_t = \ln \frac{2D_t}{(S_t - \bar{S})^2} + \frac{\sigma_\omega^2}{2} \quad (13)$$

with variance  $\sigma_{\hat{g}}^2 = \sigma_\omega^2$ . The posterior for  $g^*$  is normally distributed with the posterior mean  $g_{t+1}$  and posterior variance  $\sigma_{g,t+1}^2$ :

$$g_{t+1} = \frac{\sigma_\omega^2}{\sigma_\omega^2 + \sigma_{g,t}^2} g_t + \frac{\sigma_{g,t}^2}{\sigma_\omega^2 + \sigma_{g,t}^2} \hat{g}_t, \quad (14)$$

---

<sup>3</sup>In the problem with asymmetric information about abatement costs, anticipated learning about damages favors the use of taxes.

$$\sigma_{g,t+1}^2 = \frac{\sigma_{g,t}^2 \sigma_\omega^2}{\sigma_\omega^2 + \sigma_{g,t}^2} \Rightarrow \sigma_{g,t}^2 = \frac{\sigma_{g,0}^2 \sigma_\omega^2}{\sigma_\omega^2 + t \sigma_{g,0}^2}, \quad (15)$$

where  $\sigma_{g,0}^2$  is the prior at the beginning of the initial period,  $t = 0$ . (You can find these formulae in Greene's text on Econometrics, pages 407-410. Or look them up in Wikipedia.)

A smaller value of  $\sigma_\omega^2$  is equivalent to greater precision of future information. Using equation (14), greater precision of information implies that this period's posterior mean,  $g_{t+1}$ , is more responsive to information obtained in the current period. Using equation (15), greater precision of information means that the posterior variance decreases over time more rapidly. Thus, greater precision of information increases the speed and amount of learning.

The subjective distribution for the unknown damage parameter  $G^*$  collapses to the true value of this parameter as the number of observations approaches infinity. Appendix B1, available through *JEEM*'s online archive for supplementary material at <http://www.aere.org/journal/index/html>, proves this result.

If the regulator begins with too optimistic a prior ( $g_0 < g^*$ )  $g_t$  increases over time, on average. This increase can be enough to offset the decrease in  $\sigma_{g,t}^2$ , leading to an increase in  $var_t(G_t)$  (using equation (12)). In this case, during a phase of the learning process the regulator becomes less certain about the value of  $G^*$ , although he eventually learns the correct value with probability 1. It is also straightforward to show that the regulator's current expectation of  $G^*$  is an unbiased estimate of the future expectation:  $E_t G_{t+\tau} = G_t, \forall \tau \geq 0$ .

In the absence of anticipated learning, the regulator solves the control problem treating  $G_t$  as a constant. In this case the constant  $\bar{G} \equiv G_t = \exp\left(g_t + \frac{\sigma_{g,t}^2}{2}\right)$  is the certainty equivalent value of  $G^*$ .

### 1.4.3 The role of conjugates

In both of these examples I chose "conjugate priors", i.e. I picked a prior and a likelihood function so that the posterior has the same form as the prior. This choice greatly simplifies the problem because it makes it possible to write, in closed form, the equations of motion for the parameters that define current beliefs.

I think that in principle it is fairly straightforward to drop the assumption of conjugates, at the cost of some additional work on numerical approxima-

tion. I have given this topic a bit of thought but I have not seen any research on it. (I also have not looked for it.) Let me know if this strikes you as an interesting research topic.

#### **1.4.4 Criticism of these examples**

These examples show how to model anticipated learning about climate change. Arguably, the examples miss an important feature of the real world problem, because they do not take into account abrupt and irreversible changes. Later in this course we will discuss two kinds of irreversibilities, in the sections on sudden, catastrophic events, and in the section on nonconvex problems. Both of these modeling frameworks – particularly the nonconvex environment – offer the possibility of studying anticipated learning in a much more interesting and realistic setting. I am not aware of any research on this topic – and it seems a fruitful area of research.