Climatic Change (2008) 89:67–85
DOI 10.1007/s10584-008-9406-0

Learning about climate change and implications for near-term policy

Mort Webster · Lisa Jakobovits · James Norton

Received: 12 January 2007 / Accepted: 13 December 2007 / Published online: 18 April 2008

Abstract Climate change is an issue of risk management. The most important causes for concern are not the median projections of future climate change, but the low-probability, high-consequence impacts. Because the policy question is one of sequential decision making under uncertainty, we need not decide today what to do in the future. We need only to decide what to do today, and future decisions can be revised as we learn more.

In this study, we use a stochastic version of the DICE-99 model (Nordhaus WD, Boyer J (2000) Warming the world: economic models of global warming. MIT Press, Cambridge, MA, USA) to explore the effect of different rates of learning on the appropriate level of near-term policy. We show that the effect of learning depends strongly on whether one chooses efficiency (balancing costs and benefits) or cost-effectiveness (stabilizing at a given temperature change target) as the criterion for policy design. Then, we model endogenous learning by calculating posterior distributions of climate sensitivity from Bayesian updating, based on temperature changes that would be observed for a given true climate sensitivity and assumptions about errors, prior distributions, and the presence of additional uncertainties. We show that reducing uncertainty in climate uncertainty takes longer when there is also uncertainty in the rate of heat uptake by the ocean, unless additional observations are used, such as sea level rise.

M. Webster (✉)
MIT Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, E40-408, 77 Massachusetts Avenue, Cambridge, MA 02138, USA
e-mail: mort@mit.edu

L. Jakobovits
Technology and Policy Program, Massachusetts Institute of Technology, Cambridge, MA, USA

J. Norton
Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

1 Introduction

Climate change is a problem with a time scale of centuries and significant uncertainties. Current debates over appropriate greenhouse gas emission reductions, both over potential US domestic action and future international accords after the Kyoto commitment period ends in 2012, focus on the overall level of reductions needed. Arguments for undertaking only very minimal steps (e.g., \$7–12 per ton CO_2)[1] or delaying reductions altogether remain prominent in these debates. Among the justifications for the delayed approach is that we should wait until the uncertainty about the climate system is reduced.

While the actual policies implemented will ultimately be political decisions, to the extent that they are informed by science, the level of near-term emissions reductions should be guided by a long-term risk[2] management perspective. In the absence of uncertainty, efficient emissions reductions today would be determined by what is needed to meet the long run target, however that target is chosen. Under uncertainty, the level of reductions today should be a hedge between the policies that would be chosen under certainty for all possible states of the world. We will show that the optimal level of this hedge depends critically on both the amount of uncertainty and on whether the uncertainty can be expected to be reduced (see also Manne and Richels 1995; Kelly and Kolstad 1999).

In this study, we explore the effect of reducing uncertainty about the climate on near-term optimal greenhouse gas reductions. We adopt a Bayesian framework for updating knowledge and beliefs about uncertainty. We start with the simple case of observations only, and assume no cost to this learning. In reality, there is a cost to continuing observations, and to improving the resolution of the observational network. This study can provide a basis for value-of-information calculations. Hopefully scientific research will make progress faster than can be made by passively observing the evolution of climate, but the value of this additional information should be calculated as incremental to the learning from observations. Thus, as a first step, we explore the rate of costless learning from observing temperature change over the next several decades (for an estimate of the potential costs of climate observations, see Keller et al. 2007; Baehr et al. 2007). In reality, learning does not necessarily reduce the variance or even narrow in immediately on the "truth" (Oppenheimer et al. 2008; Henrion and Fischoff 1986). In this paper, we will use the term 'learning' to refer only to the idealized process of narrowing uncertainty.

Our approach to estimating future learning builds on several previous studies. In a seminal study of the time needed to learn about climate sensitivity, Kelly and Kolstad (1999) use a first-order autoregressive model of temperature change to solve both analytically and numerically for the number of decadal observations necessary to reject alternative sensitivities with high confidence. They estimated that 9 to 16 decades are needed for 95% confidence, and 11–20 decades are needed for 99% confidence. Leach (2007) performed a similar calculation, and showed that adding an additional uncertainty in the persistence of shocks could extend the time required to learn the climate sensitivity with 95% confidence to several centuries. Kolstad (1996) explored the effect of the rate of learning in an optimal growth model with abatement and damage costs. This study showed that if abatement capital is fully reversible, the learning rate has virtually no effect on near-term abatement, but if capital is not reversible, faster rates of learning make lower control rates optimal in the near-term.

[1] Nordhaus 1994; Nordhaus and Boyer 2000, Yohe et al, 2004, U.S. Senate 2007.

[2] Here, we define risk as consequence times probability.

In the current paper, we build on the results of these earlier studies by showing the effect of an additional uncertainty on the rate of learning. First, we review the effects of learning on near-term optimal policy as compared with the case where learning will not occur. We show in this context that whether learning matters depends in part on the framing of the policy questions; expectation of future learning has a stronger effect if the goal is temperature stabilization as opposed to economic efficiency (see also Keller et al. 2008). We also show that what matters for policy is how soon we will reduce uncertainty, and how much we will reduce it. We then calculate the rate of learning that would occur based on observations of temperature change. Bayesian updating calculations are performed using a two-dimensional climate model of intermediate complexity that explicitly represents feedbacks between different climate processes: atmosphere, ocean, biosphere, sea–ice, the carbon cycle, and heat fluxes. We show with our analysis that the presence of additional uncertainties in these physical processes increase the time needed to reduce uncertainty, but also show that observations of additional climate variables, such as sea level rise, can decrease the time to learn.

To decide what we should do today, conditional on what we expect to learn, it is useful to decompose this issue into two distinct questions. The first question is: how much do we need to learn and by when for today's policy to differ from what we would do without learning? The second question is: how much and by when can we expect to reduce uncertainty about climate sensitivity by observing temperature change? After addressing each of these questions, we synthesize the results to determine whether and how expected learning should affect the stringency of today's policy.

In Section 2, we describe the stochastic optimization model used to explore the effect of reducing uncertainty on near-term optimal emissions reductions. We compare the effect of learning on optimal carbon taxes under cost–benefit optimization and under temperature stabilization for several temperature targets. Calculations of how much uncertainty in climate sensitivity will be reduced under various assumptions are given in Section 3. Discussion of the results and conclusions are in Section 4.

2 Effect of learning on near-term emissions reductions

We first examine how policy should adjust to the expectation of learning. The effect of learning on optimal policy choice is calculated using a stochastic version of the DICE-99 model (Nordhaus and Boyer 2000). The DICE-99 model is a Ramsey growth model augmented with equations for CO_2 emissions as a function of economic production, the carbon-cycle, radiation, heat balance, and abatement cost and climate damage cost functions. The model solves for the optimal path over time of the savings/consumption decision, and also the emissions abatement decision that balances the cost of emissions abatement against damages from increased temperatures.

We modify DICE-99 to perform stochastic optimization[3] (Yohe et al (2004); Nordhaus and Popp 1997; Kolstad 1996) by simultaneously representing different states of the world (SOW), each with a different parameter value and probability of being obtained. Decisions before learning are simulated by constraining policy to be the same across all SOWs. Decisions after learning are simulated by removing this constraint. Using stochastic

[3] In operations research, stochastic optimization refers to the choice of control variables that maximize the expected value of an objective function, given a probability distribution over one or more parameters in the model. In this usage, "stochastic" does not mean that the model output will vary for the same input.

programming, we solve for the optimal intertemporal path of emissions reductions that maximize the expected net present value of utility.

We assume for this initial example that the only uncertainty is the climate sensitivity. This uncertainty is represented with the discrete four-valued prior probability distribution given in Table 1. The assumed prior is a discrete distribution based on current estimates of the probability distribution of climate sensitivity using historical temperature change and radiative forcing (Andronova and Schlesinger 2001; Forest et al. 2001);

2.1 Effect of learning on efficient policy choice

When considering the effect of reducing uncertainty on the choice of optimal policy, it is important to first compare the policy chosen under three simpler cases:

- Perfect Information – the true state of nature is known with certainty. A different optimal policy path may be chosen for each SOW.
- Never Learn – knowledge of the true state of nature is never obtained. A single policy path, which maximizes the expected utility under the prior probability distribution, is chosen for all SOWs.
- Learn Perfectly – The true state of nature becomes known with certainty in a specified time period. Before learning occurs, decisions are made based on the prior distribution.

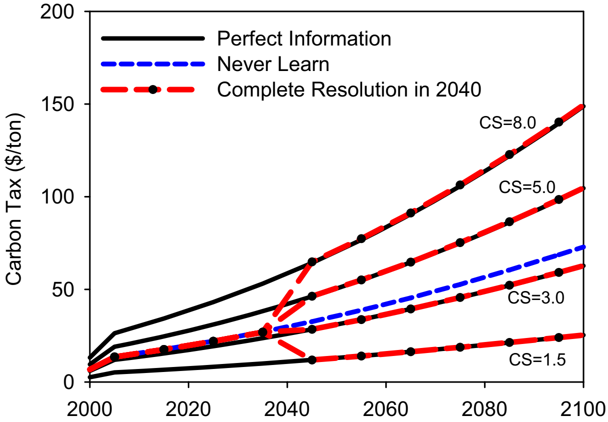
None of these assumptions are realistic, but they bound the range of policy choice across all possible assumptions about learning, and thus are a useful departure point for analysis.

We first examine these three scenarios under the efficient policy choice in DICE, which minimizes the sum of abatement and damage costs. The results are described below and shown in Fig. 1. If the climate sensitivity is known with certainty (Perfect Information) all along, different carbon taxes and emissions abatement will result for each value of climate sensitivity, with greater carbon taxes/abatement for higher climate sensitivities (black solid lines in Fig. 1). If the uncertainty in climate sensitivity as given by the prior in Table 1 never changes (Never Learn), then the optimal taxes will be slightly higher than for the case with climate sensitivity of 3° (blue dashed line in Fig. 1). This case is also sometimes called the certainty equivalent case. If decisions are made under the prior until 2040, after which the true climate sensitivity is revealed with certainty (Learn Perfectly), the optimal taxes before learning follows almost exactly the never learn case, and after revelation the policy reverts to the perfect information case (red dashed lines with circles). The carbon tax path

Table 1 Prior and posterior probabilities for partial reduction of uncertainty, assuming a 14% reduction in the coefficient of variation from the prior

True state	Most likely sensitivity after observation				Prior probability
	1.5	3	5	8	
1.5	0.75	0.11	0.07	0.07	0.25
3	0.15	0.75	0.13	0.13	0.45
5	0.05	0.07	0.75	0.04	0.15
8	0.05	0.07	0.04	0.75	0.15
Coefficient of variation	0.789	0.499	0.384	0.380	0.596
Average	0.513				
Reduction from prior	13.9%				

Fig. 1 Efficient carbon taxes under perfect information (*solid black lines*), never learning (*dashed blue line*), and complete resolution of uncertainty in 2040 (*dashed red lines with circle*)



before learning occurs (2000–2040) is the "hedging" policy under the uncertainty, as demonstrated by Manne and Richels (1995).

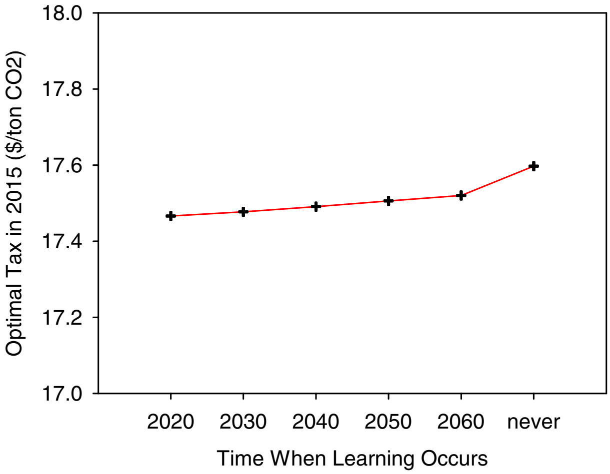
The relevant question for policy is: what should the carbon tax be today given what we expect to learn? The effect of learning can be measured by the difference between the optimal policy if learning is expected to occur in the future, and the optimal policy if learning will never occur. Although the efficient level of taxes for a near-term period, 2015 for example, do depend on whether and when learning will occur, this dependence is negligible (Fig. 2, Table 2). The reason for the weakness of the effect of learning on near term policy has been shown elsewhere to be a result of the absence from models such as DICE of a strong relationship between marginal costs or damages in one period and policy choices in previous periods (Webster 2002), or equivalently of the absence of a binding irreversibility (Kolstad 1996). In the absence of such inter-period dependencies, the expectation of future learning has no significant effect because decisions before and after learning are essentially independent in terms of the marginal costs and benefits. This contrasts with the results of Keller et al. (2004), where learning had a strong effect in the presence of a threshold. Because a partial reduction in uncertainty would necessarily result in an optimal policy level between those chosen under the "never learn" and the "learn perfectly" cases, there is no reason to further explore the effect of learning under the economic efficiency framework in this model[4].

2.2 Cost-effectiveness and avoiding temperature thresholds

An alternative decision-making framework to the economic efficiency approach used above is to choose a minimum cost path of abatement that avoids some threshold level of climate impact. The concept of stabilization is established in the Framework Convention on Climate Change (UN 1992), and continues to be a focus in policy discussions of long-term mitigation targets. One common proposal for a threshold is to stabilize global mean surface temperature at some specified level above preindustrial levels, such as 2°C or 3°C (Yohe et al. 2004; Toth et al. 2003). The proposed target of the European Union is stabilization of 2° (European Council 2005).

[4] In a different model with processes not represented in DICE-99, such as endogenous technical change (Popp 2004) or nonlinear ocean dynamics (Keller et al. 2004), learning will have a greater effect on policy choice.

Fig. 2 Efficient carbon tax in 2015 as a function of when uncertainty in climate sensitivity is completely resolved (Perfect Learning)



Under the cost-effectiveness framework[5], the expectation of learning can have a significant impact on near-term policy choice. Figure 3 and Table 3 show the optimal carbon taxes required to remain below a threshold of 2° warming under perfect information, along with the hedging policies that would result from learning perfectly in different periods. In these simulations, we assume that the temperature constraint must be met with probability 1.0. If the uncertainty in climate sensitivity could be completely resolved by 2020, the optimal path would be between those under certainty for 3° and 5° climate sensitivity. If learning is expected in 2030, the optimal path is between those for 5° and 8° climate sensitivity. If the learning does not occur until 2040 or later, then until learning occurs, the optimal hedging path is roughly the same as when climate sensitivity is known to be 8°. If we discover after 2040 that the true climate sensitivity is less than 8°, the tax level will thereafter be lowered.

The cause of this behavior is the threshold target. If the information can be obtained far enough ahead of time to change course and still avoid the threshold, then a less stringent policy will be optimal in the near term. However, if the learning occurs too late to avoid the threshold in the worst case, then the optimal hedge will be chosen as if the true state of nature is the worst case. To do any less would make it impossible to stay below the threshold with probability one. If it later turns out that climate sensitivity is lower and the abatement was excessive, the foregone growth is irreversible. This result is consistent with studies that have shown the effect of learning in the presence of irreversibilities (Arrow and Fisher 1974; Kolstad 1996; Ha-Duong 1998). Keller et al. (2004) and O'Neill et al. (2006) have also shown the effect of a threshold on learning with similar results to those shown here.

Figure 4 shows the cost-effective tax in 2015 for several temperature targets with learning occurring in different periods. Note that less stringent temperature targets allow learning to matter if it occurs later than 2040, but conversely that under more stringent temperature targets the ability to avoid high carbon taxes by learning is greater.

[5] Note that we neither advocate nor reject a cost-effectiveness framework with a temperature change stabilization target. Our goal here is to illustrate the effect of learning under different analysis frameworks. While stabilization at 2 degrees is here shown to have costs which outweigh the benefits, proponents of threshold targets justify them on the large uncertainties in abatement costs and climate damages.

Table 2 Value of Information from learning in different periods under efficient case

	Optimal Tax in 2015 ($/ton CO2)	NPV Consumption (in 2005 trillion $)	Value of Information (in 2005 billion $)	Value of Information (% NPV Consumption)
Learn in 2020	17.47	208.0593	23.937	0.0115
Learn in 2030	17.48	208.0548	19.387	0.0093
Learn in 2040	17.49	208.0494	14.037	0.0067
Learn in 2050	17.51	208.0433	7.909	0.0038
Learn in 2060	17.52	208.0368	1.392	0.0007
Never learn	17.60	208.0354		

Temperature targets above 4°C result in the same optimal taxes as in the cost-benefit version above, since in this model the efficient solution keeps temperature change below 4° for all climate sensitivities. Thus, the expectation of learning does change the choice of cost-effective policy under a temperature stabilization target, but only if the target is fairly stringent and only if the learning will occur quickly; i.e., within two to four decades. The effect of learning in the presence of a threshold target is greatest during the decades just before the threshold is reached. When the threshold is several decades away or it is too late, learning has almost no effect.

2.3 Partial reduction in uncertainty under a temperature target

As shown above, optimal near-term carbon taxes will depend on whether we expect to resolve uncertainty if the goal is to stabilize temperature change. The question then becomes whether it is also optimal to set a lower near-term carbon tax if we only expect to reduce, but not completely resolve, the uncertainty. How much do we need to learn in order for near-term policy to be less stringent than if we would never learn?

To model partial reduction in uncertainty, we modify the stochastic DICE model. Here we maximize the expected utility within the temperature change constraint across 16 SOWs, which consist of four subsets of four SOWs each. Each subset represents one possible observation that results in a new posterior distribution over climate sensitivity, and within a subset, each of the four states has a different climate sensitivity as its true state and

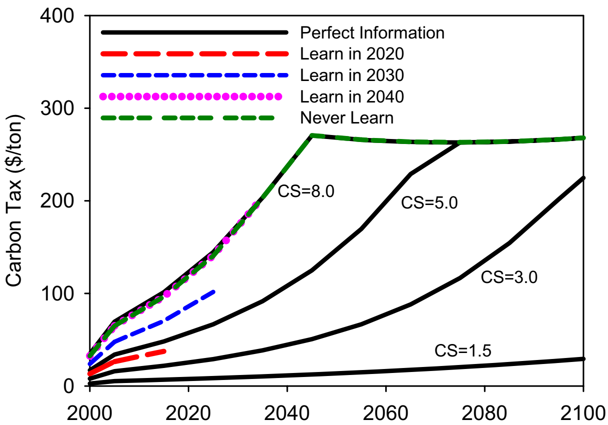


Fig. 3 Cost-effective carbon taxes to stabilize global temperature change at 2° above preindustrial. Perfect information cases shown as *solid lines* and *dashed lines* show the hedging policy before learning occurs in each of several different periods. Carbon taxes after learning occurs not shown (see Table 3)

Table 3 Optimal carbon prices with learning under cost-effective 2° target

	Optimal carbon price ($/ton)											
	CS	2000	2010	2020	2030	2040	2050	2060	2070	2080	2090	2100
Perfect info	1.5	5.2	6.8	8.5	10.3	12.4	14.8	17.4	20.4	23.6	27.3	31.4
	3	15.9	21.8	29.1	38.4	50.6	66.7	88.0	116.6	154.4	201.7	247.8
	5	33.9	48.0	66.6	91.5	124.9	169.5	228.8	263.0	264.0	266.3	269.7
	8	69.4	101.3	144.4	203.4	270.4	265.8	263.5	263.0	264.0	266.3	269.7
Learn 2020	1.5	26.3	37.5	8.4	10.3	12.4	14.7	17.4	20.3	23.6	27.2	31.3
	3	26.3	37.5	28.8	38.0	50.0	65.9	86.9	115.0	152.2	198.7	244.8
	5	26.3	37.5	67.6	92.9	126.9	172.5	233.0	263.0	264.0	266.3	269.7
	8	26.3	37.5	180.0	255.3	270.4	265.8	263.5	263.0	264.0	266.3	269.7
Learn 2030	1.5	47.6	70.2	101.6	10.1	12.2	14.6	17.2	20.2	23.4	27.0	31.0
	3	47.6	70.2	101.6	35.9	47.1	61.8	81.1	106.7	140.4	183.3	229.1
	5	47.6	70.2	101.6	85.1	115.6	156.3	210.2	263.0	264.0	266.3	269.7
	8	47.6	70.2	101.6	277.8	270.4	265.8	263.5	263.0	264.0	266.3	269.7
Learn 2040	1.5	64.8	96.5	140.8	203.7	11.8	14.2	16.9	19.8	23.1	26.6	30.4
	3	64.8	96.5	140.8	203.7	41.1	53.3	69.2	89.8	116.8	151.9	195.5
	5	64.8	96.5	140.8	203.7	93.7	125.2	166.5	219.6	264.0	266.3	269.7
	8	64.8	96.5	140.8	203.7	270.4	265.8	263.5	263.0	264.0	266.3	269.7
Never learn	All	64.9	96.4	140.5	203.1	270.4	265.8	263.5	263.0	264.0	266.3	269.7

a posterior probability that this state obtains. Table 1 shows the posterior probabilities for one case in which the standard deviations of the posteriors are on average reduced by 14% from the prior. Each column shows one posterior distribution, and the total probability of each state is the product of the prior and posterior probability.

We solve for the optimal carbon tax in 2015 for three different temperature targets (2, 2.5, 3) and for two different assumptions per target about the period in which learning occurs. The chosen learning periods differ for each temperature target, and reflect the time

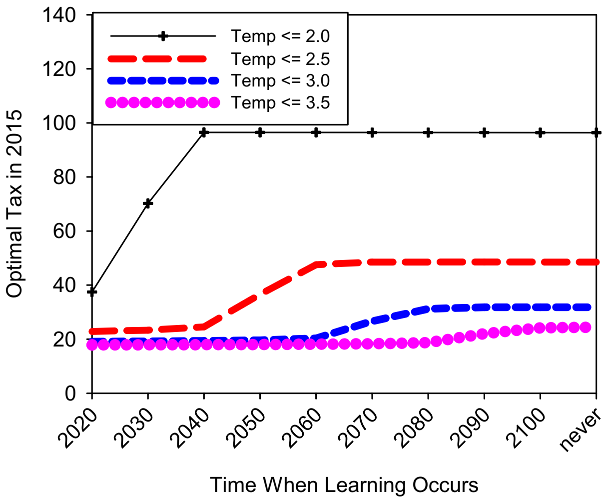


Fig. 4 Cost-effective carbon tax in 2015 to achieve temperature targets and the effect of complete resolution of uncertainty in different periods

period over which learning will influence the optimal hedging policy (see Fig. 5). To achieve a 2° target, learning matters only if it occurs before 2040; for a 2.5° target, learning must occur before 2060; for a 3° target, learning must occur before 2080. We assume for all temperature targets that perfect information is obtained at this time (i.e. 2040, 2060, 2080 respectively), and assess the effect of obtaining revised posteriors 20 years or 10 years prior to this date.

The revised posteriors assume a constant level of confidence across all four SOWs (e.g., 75%), with the remaining probability (e.g., 25%) divided across states in accordance with the ratio in the prior. The reduction in the coefficient of variation from the prior is then assessed for all states and averaged, to provide a measure of the reduction in uncertainty (as in Table 1). The results are given in Table 4 and Fig. 5. If the uncertainty in climate sensitivity is reduced by 20% or less, there is relatively little change in the optimal carbon tax relative to the "Never Learn" solution. If the reduction in uncertainty is greater than 40%, however, the optimal tax is nearly as low as the case in which all uncertainty is resolved in 2020. There appears, in this model, to be a critical reduction in the level of uncertainty necessary to affect policy, a reduction of at least 20–40% in the coefficient of variation.

In summary, the results from the stochastic optimization model indicate that while the economically efficient policy is not significantly influenced by the expectation of future learning, the cost-effective policy to achieve temperature stabilization is influenced by the expectation of learning. In the latter case, the effect on policy depends critically on how soon the learning will occur, and on how much the uncertainty will be reduced. If the uncertainty in climate sensitivity can be reduced by 20–40% within the next 20 years, a lower carbon tax and lower carbon abatement will be cost-effective in achieving temperature targets of 2° to 3°. However, if this much learning is not expected within a few decades, the cost-effective target for temperature stabilization is the same as if we never expected to learn, which entails a significantly higher level of abatement.

3 Bayesian learning about climate sensitivity

The previous section has shown that to remain below an increase of 2°C to 2.5°C global mean temperature change, the cost-effective level of carbon taxes over the next decade depends on whether we can reduce the uncertainty in climate sensitivity by 20–40% within the next few decades. We now turn to a more detailed model of the climate system to estimate the amount of uncertainty that could be expected to be reduced by hypothetical future observations of climate.

3.1 Methods and models

Unlike the 4-state discrete distribution in the previous section, here we approximate the continuous uncertainty in climate sensitivity with a 100 bin discrete distribution. We measure the amount of learning for each posterior in terms of the reduction in the coefficient of variation, which is equal to the standard deviation normalized by the mean and provides a measure of variance independent of the mean of the distribution. Since different posteriors result from different sets of observations, each conditional on a different sensitivity being the "true state of nature", we use the average reduction in the coefficient of variation, relative to the prior, of four possible true states: 1.5°C, 3.0°C, 5.0°C, and 8.0°C.

Table 4 Cost-effective carbon taxes for temperature stabilization when uncertainty is reduced

		% Reduction in coefficient of variation from prior											
Temperature target (°C)	Learning period	0%	13.9%	17.0%	20.8%	25.6%	28.1%	32.5%	34.2%	36.0%	38.0%	40.3%	100%
2	2020–2040	$96.47	$92.76	$90.96	$87.83	$81.14	$76.25	$64.66	$59.86	$54.48	$48.87	$43.14	$37.46
	2020–2040	$96.47	$92.90	$91.21	$88.33	$82.39	$78.16	$72.18	$71.90	$71.55	$71.15	$70.70	$70.21
2.5	2040–2060	$47.55	$46.34	$45.73	$44.73	$42.70	$41.24	$37.46	$35.65	$33.38	$30.75	$27.76	$24.54
	2050–2060	$47.55	$46.39	$45.81	$44.87	$43.05	$41.77	$38.59	$38.15	$37.85	$37.49	$37.06	$36.57
3	2060–2080	$31.18	$30.80	$30.55	$30.15	$29.35	$28.77	$27.20	$26.34	$25.24	$23.85	$22.19	$20.29
	2070–2080	$31.18	$30.82	$30.58	$30.20	$29.47	$28.96	$27.65	$27.53	$27.38	$27.17	$26.92	$26.63

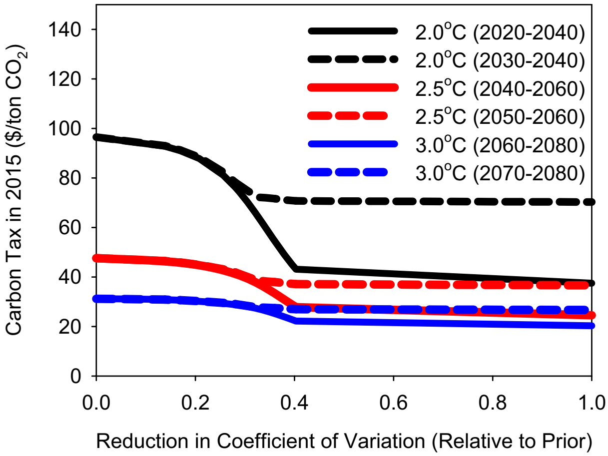


Fig. 5 Cost-effective carbon tax in 2015 to achieve temperature targets as a function of the amount of reduction in the standard deviation of the distribution of climate sensitivity. Results shown for three different temperature targets (2°C, 2.5°C, and 3°C) and for two different assumptions per target about the period under which revised posteriors are available before perfect information arrives

Our model for projecting future temperature changes and sea level rise is a reduced-form model that has been statistically fitted to the MIT two-dimensional climate model (Sokolov and Stone, 1998; Webster et al. 2003). The MIT climate model consists of a two-dimensional (2D) zonally-averaged land–ocean resolving atmospheric model, coupled to an atmospheric chemistry model, a 2D ocean model consisting of a surface mixed layer with specified meridional heat transport, diffusion of temperature anomalies into the deep ocean, an ocean carbon component, and a thermodynamic sea–ice model (Sokolov and Stone 1998; Wang et al. 1998, 1999; Prinn et al. 1999).

The reduced-form model reproduces the global and zonal temperature changes and sea level rise from the MIT model, based on over 1500 simulations with the 2D model, and exhibits less than 1% error over a wide range[6] of its input parameters: greenhouse gas emissions over time, climate sensitivity, deep ocean heat uptake, and aerosol forcing. The reduced-form model consists of third-order expansions for decadal average temperature change and decadal sea level rise, as a function of the above uncertain input parameters. The reduced-form model has been documented in detail elsewhere (Webster et al. 2003; Webster 2002). Use of the MIT climate model enables us to consider two additional factors: the effect of uncertainty in the deep ocean heat uptake, and the additional learning from combining sea level rise observations with temperature observations.

Learning is modeled as Bayesian updating according to Bayes' Law:

$$P(\mathrm{CS}|\Delta T) = \frac{P(\Delta T|\mathrm{CS})P(\mathrm{CS})}{\sum_{i=1}^{n}\left(P(\Delta T|\mathrm{CS}_i)P(\mathrm{CS}_i)\right)} \tag{1}$$

Equation 1 expresses Bayes' Law in terms of the specific problem here: what is the posterior distribution for climate sensitivity (CS) given an observed temperature change trajectory (ΔT)?

The Bayesian updating calculation is performed by Monte Carlo Integration. First, a value for the "true" climate sensitivity is assumed. Then the MIT climate model is used to generate a time series of observations for the years 2000–2100 that could result from that

[6] Reduced-form model is fitted to results for climate sensitivity varying from 0 to 10K, deep ocean heat uptake from 0 to 10 cm^2/s, aerosol forcing strength from −1.5 to 0 w/m^2, and rates of increased CO_2 forcing from 0.2%/year to 1.8%/year.

true sensitivity. The third step is to calculate the likelihood of the observation $P(\Delta T|\mathrm{CS})$ for successive intervals of 0.1°C over the range of possible values for sensitivity based on a conditional Monte Carlo simulation with 4,000 samples. The final step is to calculate the posterior probability distribution $P(\mathrm{CS}|\Delta T)$, by multiplying the prior distribution and the likelihood, and renormalizing to integrate to one. We repeat this calculation for each of four possible "true" values for climate sensitivity: 1.5°C, 3.0°C, 5.0°C, and 8.0°C.

A critical assumption in these calculations is the level of error or "noise" between observations and the model. In the calculations here, the error is modeled as independently and identically normally distributed with mean zero and a standard deviation σ_e, which we vary in sensitivity testing. The error is discussed further in section 3.3 below.

3.2 Updating climate sensitivity from observations of temperature change

We begin with the simplest case, using observations of decadal average global mean surface temperature change to revise the probability distribution of climate sensitivity. Figure 6 shows the revised probability distributions for climate sensitivity from the updating calculations outlined above, using one possible set of assumptions about the prior distribution and magnitude of error. The figure shows the posterior distributions for four possible "true" values of climate sensitivity that would be learned by 2020 and 2040, respectively. Note that by 2020 under these particular assumptions, the change is small, mainly reducing the likelihood of very high sensitivities in the 1.5 case, and of very low sensitivities in the other cases. By 2040, in contrast, the posterior PDFs are substantially different from each other, except for the difference between the 5.0 and 8.0 cases, as will be explained below.

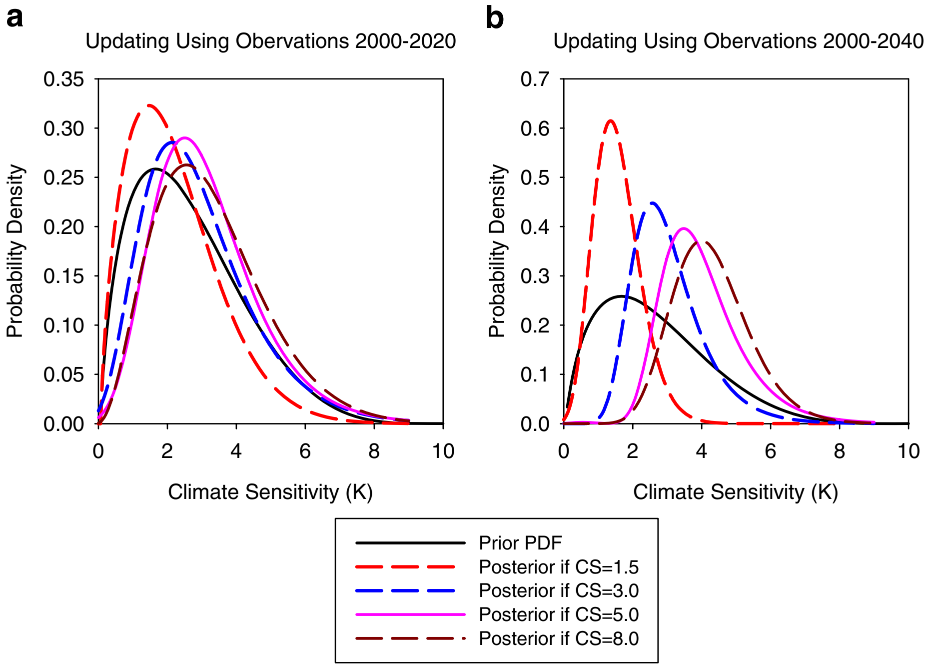


Fig. 6 Prior and posterior probability distributions for climate sensitivity, based on observations of decadal average temperature change in **a** 2020 and **b** 2040, assuming an error of 0.3*K*

The rate at which uncertainty is reduced will also depend on the "true" climate sensitivity that generates the observations. Figure 7 shows the reduction in the coefficient of variation by decade, for the case where the errors have a standard deviation of 0.3*K*. Note that if the true sensitivity is low (1.5*K*), the rate of learning is slower.

3.3 Sensitivity of learning rate to error and prior distribution

One of the key determinants of the rate of learning is the assumed error or noise in matching observation to model. If there were no sources of error or variability at all, and if climate sensitivity was the only uncertainty, then one observation would be enough to know the climate sensitivity perfectly. There are several different sources of the divergence between observations and model. The first and best known source is the natural variability of the climate system (Folland et al. 2001), estimated at 0.09°C/decade. However, natural variability is not the only or even necessarily the largest source of error. A second source for which there are less precise estimates is the error in a model's ability to match observations. This second source includes the effects of both model bias and model variability. A third source of error is observational or measurement error. While not negligible, this error is likely not as large as the other two sources.

The sum of all these sources of error must be accounted for in the Bayesian updating calculations. Figure 8 shows the effect of different magnitudes of total error on the rate of learning, measured as the average reduction in the coefficient of variation. For example, the time until the uncertainty is reduced by 40% ranges from 2020 if the error is 0.1°C (natural variability only) to 2065 if the error is 0.8°C. The greatest effect of the error is on the amount of learning in the first few decades.

In order to estimate the total error between model and observations, we compare 500 model simulations of 20-century climate from the MIT 2D model (Forest et al. 2006) to the Hadley Center observation dataset (Brohan et al. 2006; Jones et al. 1999; Rayner et al. 2003, 2006). The deviations in decadal average global mean surface temperatures over 1860–2000 have standard deviations that range from 0.2*K* to 0.3*K*.

For the calculations shown here, we assume that the error is independent and identically distributed (iid). It has been shown that the error in surface temperature change is autocorrelated (Andronova and Schlesinger 2001) and that autocorrelated errors will slow the rate of learning (Keller and McInerney 2007; Zellner and Tian 1964). Thus, all results

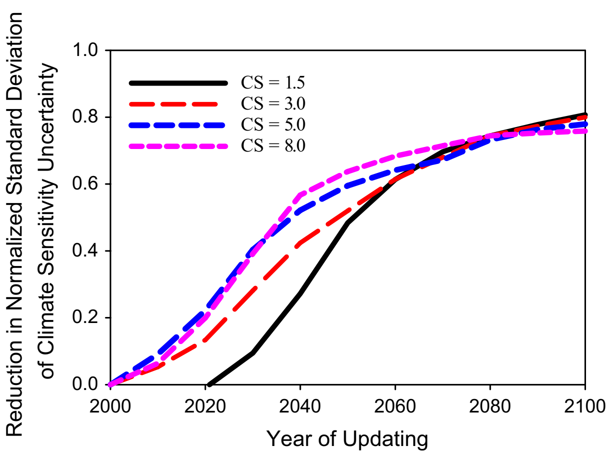


Fig. 7 Coefficient of variation (standard deviation/mean) for posterior distributions from each additional decade of observation, assuming an error of 0.3*K*

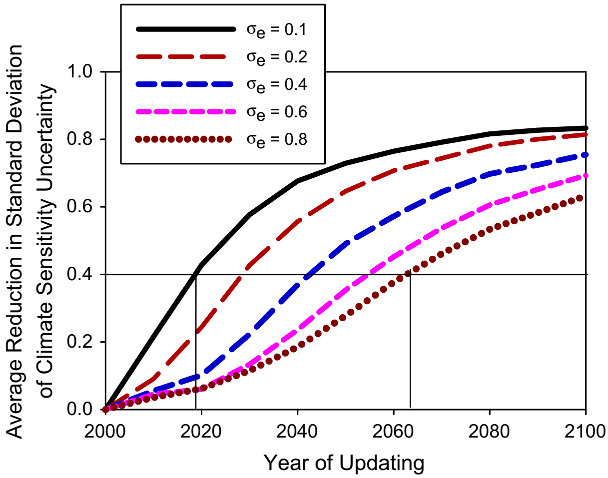


Fig. 8 Average reduction in the coefficient of variation for different magnitudes of error

here are likely biased towards more rapid learning than might be expected. However, the relative rates of learning for one uncertain parameter vs. two parameters below, would likely remain qualitatively unchanged with autocorrelated errors.

A second assumption that significantly affects the rate of learning is the prior probability distribution. We consider here three alternative priors (Fig. 9). The simplest assumption is a uniform distribution between 0.0 and 10.0*K*. Two other priors from the literature are from Forest et al. (2001, 2002), one developed from a uniform prior before 20-century observations were used to constrain the distribution, and one based on an expert prior. The expert prior from Forest et al has the smallest variance of the three probability distributions.

There are two distinct effects of using a prior with a smaller variance. In general, a lower variance for the prior will result in more rapid decreases in posterior variances (Fig. 10a). Thus, one might expect that the rate of learning, measured by the reduction in the standard deviation, can be accelerated by assuming more confident priors.

However, narrower priors increase the risk of overconfidence which may slow the convergence of posteriors to the true value. This problem occurs when the true value in

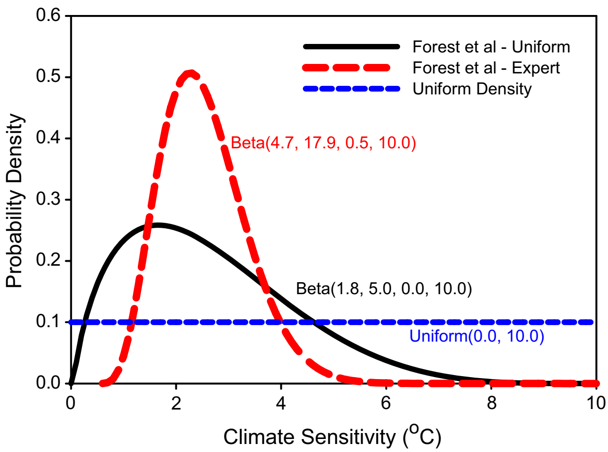


Fig. 9 Three alternative prior probability distributions for climate sensitivity

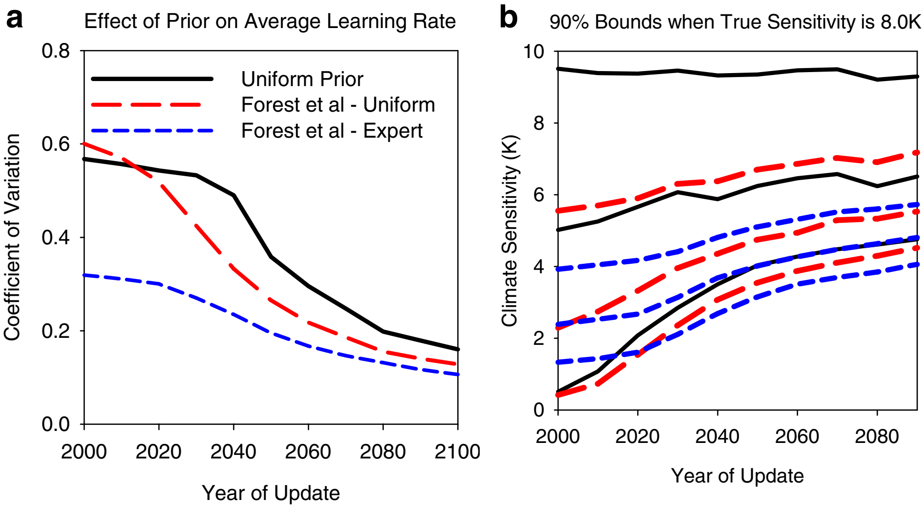


Fig. 10 Effect of three different priors on the rate of reducing the uncertainty: **a** the average coefficient of variation (standard deviation/mean); **b** estimated median and 90% bounds of climate sensitivity when true climate sensitivity generating observations is 8.0*K*

nature generating the observations is given a very low likelihood in the prior. As an example, consider the case where the true climate sensitivity is 8.0*K*. Although the uniform prior gives equal weight to this possibility as all others, the other two priors assign a very low likelihood. As a result, many more observations are required before the 90% bounds of the posterior even include the true value. In this example, even after a century of observations, the true value lies outside the 90% bounds (Fig. 10b).

3.4 Updating joint distribution of climate sensitivity and ocean uptake

There is a third critical factor that determines the rate at which we will reduce uncertainty about climate sensitivity: the presence of other uncertainties about the climate system. If climate sensitivity were the only uncertainty, learning would be more straightforward and might progress more rapidly. Unfortunately, there are several critical uncertainties[7] about the system that are not independent of each other (Forest et al. 2001, 2002, 2006). One of the most important of these is the rate of heat uptake by the deep ocean (Sokolov and Stone 1998; Forest et al. 2006). The combination of the uncertainties in ocean uptake and climate sensitivity presents a difficulty for reducing uncertainty, because for any observed temperature change, there are many combinations of climate sensitivity and heat uptake that are consistent with it.

Here we extend the above procedure to perform Bayesian updating on the joint distribution of climate sensitivity (CS) and deep ocean heat uptake[8] (*K*v). The procedure is

[7] Another important uncertainty is the radiative forcing strength of sulfate aerosols. Including this third uncertainty is beyond the scope of this study.

[8] The version of the MIT climate model used in this study has a 2D ocean: zonal and vertical layers. The *K*v parameter does not correspond to the vertical eddy diffusivity in a 3D ocean model. Rather, it represents global scale heat diffusivity between the mixed layer and deeper layers. Specifically, the vertical diffusivity is proportional to the square root of *K*v.

the same as above except that now the likelihood function is developed for all possible combinations of CS and *K*v. When temperature change is used as the set of observations to update the joint distribution, learning is slower than it is for climate sensitivity only (Fig. 11). The reason for this is that the range of climate sensitivities that can result in a given temperature change is wider if heat uptake is simultaneously adjusted. A higher climate sensitivity could be consistent with an observation if the additional heat was absorbed into the deep ocean (faster heat uptake). Conversely, a lower climate sensitivity could be consistent with the same temperature observation if less heat is mixed into the deep ocean (slower heat uptake). If the posterior distribution resulting from updating on temperature alone is overlaid on contours of temperature change as a function of sensitivity and heat uptake, one can see how the high likelihood portion of the posterior clusters along the observed temperature change contour (Fig. 12b).

Thus, an additional observation is needed to distinguish between different combinations of CS and *K*v that give similar surface temperatures. The difference between such pairs will be reflected by the amount of heat stored in the ocean, and one measure of this is the sea level rise from thermal expansion. Therefore, we repeat the calculations assuming true values for climate sensitivity and for heat uptake, and using joint observations of surface temperature change and sea level rise. Sea level rise from thermal expansion is similarly determined by the MIT climate model, and we assume an independent and identically distributed error of ±30% based on Church et al. (2001). Using the combined observations more effectively constrains the posterior distributions and results in more rapid learning (Figs. 11 and 12c).

4 Conclusions

From the calculations in this study, it appears that a substantial reduction in uncertainty in climate sensitivity, of up to 20–40%, could be possible within the next two to five decades as a result of the additional climate observations. The rate of learning would be accelerated by considering multiple climate variables simultaneously. In addition to global mean surface temperature and sea level rise, temperatures at various heights in the atmosphere and depths in the ocean and zonal mean surface temperatures may further constrain possible

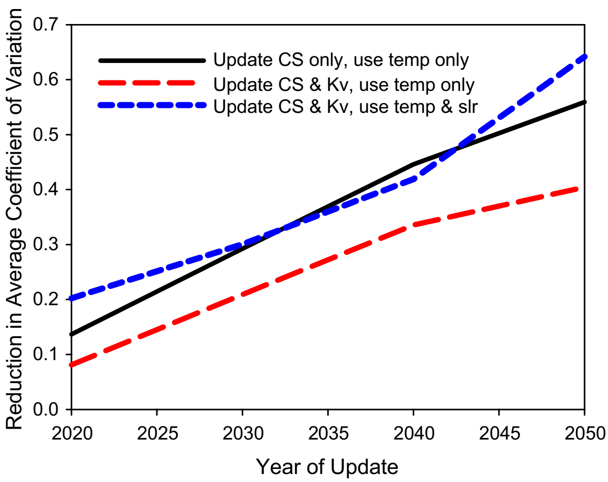


Fig. 11 Rate of learning from updating climate sensitivity only, joint distribution of climate sensitivity and heat uptake using temperature observations only, and joint distribution of climate sensitivity and heat uptake using temperature and sea level rise observations. These updates assume an error of 0.3*K* and use the Forest et al. uniform prior

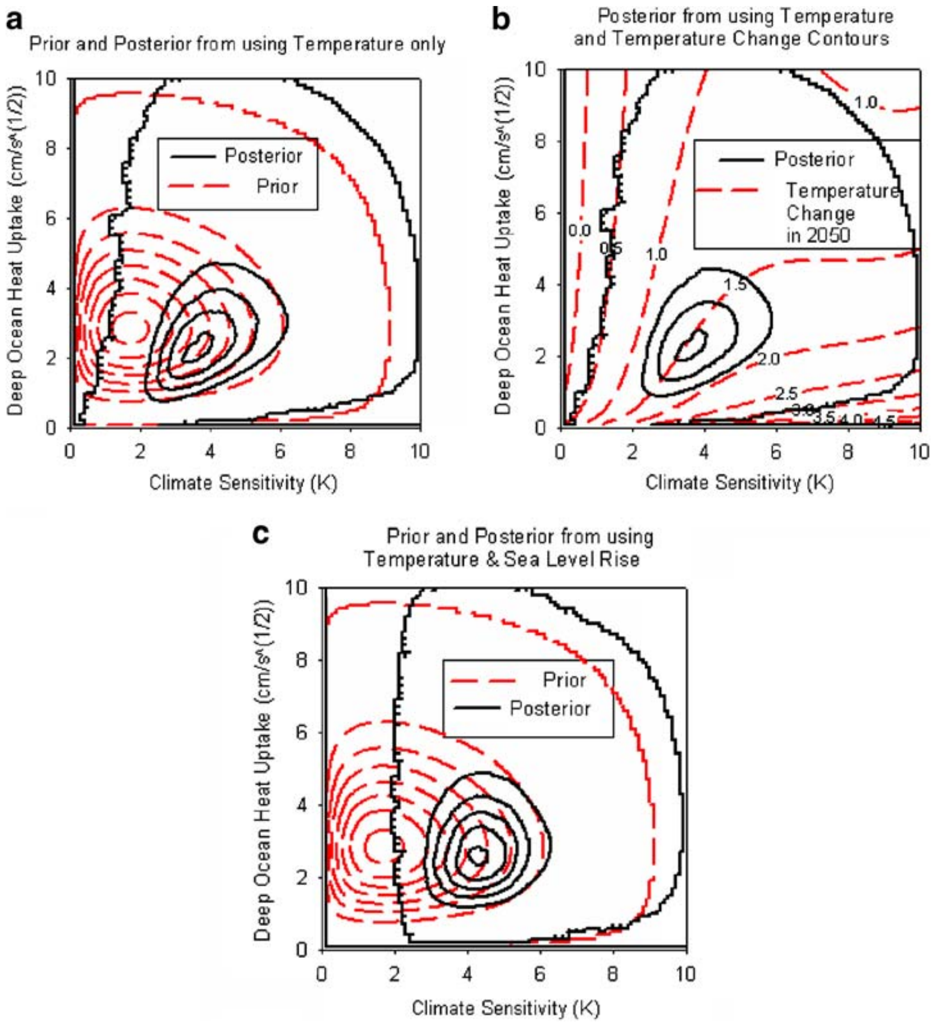


Fig. 12 Joint prior and posterior probability distributions for climate sensitivity and deep ocean heat uptake; **a** prior and posterior uptake from temperature observations 2000–2040 when true climate sensitivity is 5.0K and true heat uptake is 2.5 cm/s$^{1/2}$; **b** posterior from temperatures only through 2040 and contours of temperature change from MIT 2D climate model; **c** prior and posterior joint distributions, updated on observations of temperature change and sea level rise 2000–2040

values, although smaller scales of aggregation will also be associated with larger errors. In general, the rate of learning will depend strongly on the magnitudes of natural variability, model errors, and observational errors. A misspecification of these errors could even lead beliefs temporarily away from the true state of nature (Oppenheimer et al. 2008).

In terms of the effect of expected learning on the appropriate stringency of greenhouse gas reductions in the near-term, the estimates in this study suggest that uncertainty could be reduced enough in the next few decades to justify a somewhat lower carbon tax (or less stringent emissions cap) than if we did not expect to learn, but not as low as if we expected to know for certain within a few decades. However if the goal is to meet a stabilization target, the delay in reducing these uncertainties argues for a somewhat precautionary approach with greater near-term emissions reductions. Under a wide range of alternative assumptions with the model shown here, the expectation of future learning never justifies a carbon tax below the range of \$15–20 per ton of CO_2 for the next decade or two, and may justify higher. In addition to climate thresholds, there are other reasons that future learning might affect near-term policy, such as the possibility of significant endogenous technical change within the energy system.

Much work remains to refine estimates of how quickly we might reduce uncertainty, and even more importantly, to better identify what observations and what research could accelerate the learning. Bayesian learning techniques, like the ones used here, can be used to focus on both what climate variables and at what locations and resolution would most efficiently reduce the uncertainty about the climate system. The value of information of such additional observations would likely more than justify their costs if they can reduce the uncertainty enough to avoid either unnecessary climate damage or unnecessary regulation.

Acknowledgments The authors are grateful for helpful comments and suggestions from one anonymous referee, Natasha Andronova, Karen Fisher-Vanden, Chris Forest, Klaus Keller, Charlie Kolstad, Michael Schlesinger, and Andrei Sokolov, and the participants of the IIASA Conference on Learning and Climate Change.

References

Andronova NG, Schlesinger ME (2001) Objective estimation of the probability density function for climate sensitivity. J Geophys Res-Atmos Res 106:22605–22611

Arrow KJ, Fisher AC (1974) Environmental Preservation, Uncertainty, and Irreversibility. Q J Econ 88: 312–319

Baehr J, Keller K, Marotzke J (2007) Detecting potential changes in the meridional overturning circulation at 26°N in the Atlantic. *Climatic Change*, Published online: 10 January 2007 DOI 2010.1007/s10584-10006-19153-z

Brohan P, Kennedy JJ, Haris I, Tett SFB, Jones PD (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. J Geophys Res 111:D12106

Church JA, Gregory JM, Huybrechts P, Kuhn M, Lambeck K, Nhuan MT, Qin D, Woodworth PL (2001) Changes in sea level,' Chapter 11. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) Climate change 2001: the scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, p 881

European Council (2005) Presidency conclusions. European Council, Brussels

Folland CK, Karl TR, Christy JR, Clarke RA, Gruza GV, Jouzel J, Mann ME, Oerlemans J, Salinger MJ, Wang S-W (2001) Observed climate variability and change,' Chapter 2. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) Climate change 2001: The Scientific Basis Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, p 881

Forest CE, Allen MR, Sokolov AP, Stone PH (2001) Constraining climate model properties using optimal fingerprint detection methods. Clim Dyn 18:227–295

Forest CE, Stone PH, Sokolov AP, Allen MR, Webster MD (2002) Quantifying uncertainties in climate system properties with the use of recent climate observations. Science 295:113–117

Forest CE, Stone PH, Sokolov AP (2006) Estimated PDFs of climate system properties including natural and anthropogenic forcings. Geophys Res Lett 33(L01705):1–4

Ha-Duong M (1998) Quasi-option value and climate policy choices. Energy Econ 20:599–620

Henrion M, Fischoff B (1986) Assessing uncertainty in physical constants. Am J Phys 54(9):791–798

Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperature and its variations over the last 150 years. Rev Geophys 37:173–199

Keller K, McInerney D (2007) The dynamics of learning about a climate threshold,' Climate Dynamics, in review, available at: http://www.geosc.psu.edu/~kkeller/publications.html
Keller K, Bolker BM, Bradford DF (2004) Uncertain climate thresholds and optimal economic growth. J Environ Econ Manage 48:723–741
Keller K, Deutsch C, Hall MG, Bradford DF (2007) Early detection of changes in the North Atlantic meridional overturning circulation: Implications for the design of ocean observation systems. J Climate 20:145–157
Keller K, Kim S-R, Baehr J, Bradford DF, Oppenheimer M (2007) What is the economic value of information about climate thresholds? Chapter in: Integrated Assessment of Human Induced Climate Change, Chief Editor: Michael Schlesinger, Cambridge University Press, pp. 343–354
Kelly DL, Kolstad CD (1999) Bayesian learning, growth, and pollution. J Econ Dyn Control 23:491–518
Kolstad CD (1996) learning and stock effect in environmental regulation: the case of greenhouse gas emissions. J Environ Econ Manage 31:1–18
Leach AJ (2007) The climate change learning curve. J Econ Dyn Control 31(5):1728–1752
Manne AS, Richels RG (1995) The greenhouse debate: economic efficiency, burden sharing and hedging strategies. Energy J 16(4):1–37
Nordhaus WD (1994) Managing the global commons: the economics of climate change. MIT Press, Cambridge, MA, USA
Nordhaus WD, Boyer J (2000) Warming the world: economic models of global warming. MIT Press, Cambridge, MA, USA
Nordhaus WD, Popp D (1997) What is the value of scientific knowledge? An application to global warming using the PRICE Model. The Energy J 18(1):1–45
O'Neill B, Ermoliev Y, Ermolieva T (2006) Endogenous risks and learning in climate change decision analysis. In: Marti K, Ermoliev Y, Pflug G, Makovskii M (eds) Coping With Uncertainty: Modeling and Policy Issues. Springer Verlag, Berlin, pp 283–299
Oppenheimer M, O'Neill B, Webster MD (2008) Negative learning' climatic change. Clim Change, (in press) DOI 10.1007/s10584-008-9405-1
Popp D (2004) ENTICE: endogenous technological change in the DICE model of global warming. J Environ Econ Manage 48:742–768
Prinn R, Jacoby H, Sokolov A, Wang C, Xiao X, Yang Z, Eckaus R, Stone P, Ellerman D, Melillo J, Fitzmaurice J, Kicklighter D, Holian G, Liu Y (1999) Integrated global system model for climate policy assessment: feedbacks and sensitivity studies. Clim Change 41(3–4):469–546
Rayner NA, Brohan P, Parker DE, Folland CK, Kennedy JJ, Vanicek M, Ansell T, Tett SFB (2006) Improved analyses of changes and uncertainties in marine temperature measured in situ since the mid-nineteenth century: the HadSST2 Dataset. J Climate 19:446–469
Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Globally complete analyses of sea surface temperature, sea ice and night marine air temperature, 1871–2000. J Geophysical Research 108:4407
Sokolov A, Stone P (1998) A flexible climate model for use in integrated assessments. Clim Dyn 14: 291–303
Toth FL, Bruckner T, Füssel H-M, Leimbach M, Petschel-Held G (2003) Integrated assessment of long-term climate policies: Part 1 – Model presentation. Clim Change 56(1–2):37–56
United Nations (1992) Framework Convention on Climate Change. International Legal Materials 31:849–873
U.S. Senate (2007) The low carbon economy act. Senate Draft Proposal. http://energy.senate.gov/public_new/_files/END07842_xml1.pdf
Wang C, Prinn RG (1999) Impact of emissions, chemistry and climate on atmospheric carbon monoxide: 100-year predictions from a global chemistry-climate model. Chemosphere Glob Chang Sci 1(1–3): 73–81
Wang C, Prinn RG, Sokolov AP (1998) A global interactive chemistry and climate model: formulation and testing. J Geophys Res 103(D3):3399–3417
Webster MD (2002) The curious role of "learning" in climate policy: should we wait for more data? Energy J 23(2):97–119
Webster MD, Forest CE, Reilly JM, Babiker MH, Kicklighter DW, Mayer M, Prinn RG, Sarofim M, Sokolov AP, Stone PH, Wang C (2003) Uncertainty analysis of climate change and policy response. Clim Change 61(3):295–320
Yohe G, Andronova N, Schlesinger M (2004) To hedge or not against an uncertain climate future? Science 306:416–417
Zellner A, Tian GC (1964) Bayesian analysis of the regression model with autocorrelated errors. J Am Stat Assoc 59:763–778