# MIT CEEPR

**MIT Center for
Energy and Environmental
Policy Research**

# An Approximate Dynamic Programming Framework for Modeling Global Climate Policy under Decision-Dependent Uncertainty

Mort Webster, Nidhi Santen and
Panos Parpas

# An approximate dynamic programming framework for modeling global climate policy under decision-dependent uncertainty

**Mort Webster · Nidhi Santen · Panos Parpas**

September 14, 2011

**Abstract** Analyses of global climate policy as a sequential decision under uncertainty have been severely restricted by dimensionality and computational burdens. Therefore, they have limited the number of decision stages, discrete actions, or number and type of uncertainties considered. In particular, other formulations have difficulty modeling endogenous or decision-dependent uncertainties, in which the shock at time $t+1$ depends on the decision made at time $t$. In this paper, we present a stochastic dynamic programming formulation of the Dynamic Integrated Model of Climate and the Economy (DICE), and the application of approximate dynamic programming techniques to numerically solve for the optimal policy under uncertain and decision-dependent technological change. We compare numerical results using two alternative value function approximation approaches, one parametric and one non-parametric. Using the framework of dynamic programming, we show that an additional benefit to near-term emissions reductions comes from a probabilistic lowering of the

Mort Webster
Engineering Systems Division
Massachusetts Institute of Technology
77 Massachusetts Avenue Cambridge, MA 02139
E-mail: mort@mit.edu

Nidhi Santen
Engineering Systems Division
Massachusetts Institute of Technology
77 Massachusetts Avenue Cambridge, MA 02139
E-mail: nrsanten@mit.edu

Panos Parpas
Department of Computing
Imperial College London
180 Queen's Gate
London, SW7 2AZ
United Kingdom
E-mail: p.parpas@imperial.ac.uk

costs of emissions reductions in future stages, which increases the optimal level of near-term actions.

# 1 Introduction

Responding to the threat of global climate change is one of the most difficult risk management problems that society faces. An optimal path of greenhouse gas emissions reductions in principle should be the path that balances the costs of emissions reductions, or abatement, with the climate-related damages from emissions. However, both the costs of emissions reductions and the damages from climate change are uncertain, and neither will be known with certainty for a long time. Nevertheless, information about the uncertainties will be revealed gradually, and policies will be continually responding to new information and other changing conditions.

Models that represent the complete causal chain from economic activity to ultimate physical impacts of climate change are referred to as "integrated assessment models (IAMs)." They simulate both the economic and the biogeophysical systems and their interactions in a single model. The majority of analyses with integrated assessment models are deterministic, and are focused on understanding and improving representations of the integrated system. There has been some work applying probabilistic uncertainty analysis to IAMs, usually in the form of Monte Carlo simulation, e.g., [17, 26, 27, 31, 32]. Studies that have explicitly modeled sequential decision under uncertainty have represented the problem in a highly simplified and stylized manner, often as a two-stage problem with a small number of discrete actions and uncertainties (e.g., [2, 10, 14, 28–30, 36].

An appropriate framing of this problem is as a dynamic stochastic optimization model. This general class of problems can be formulated and solved with either stochastic programming with recourse or dynamic programming methods. There are several special challenges to the climate problem that make it difficult to solve with existing numerical methods. First, the long time-lags in the earth system make it necessary to simulate at a minimum a few centuries ahead. Policy decisions can be revised at any time, making this a decision problem with many stages. The dimensionality of the problem is increased further by the number of uncertainties inherent in projecting global economic and technological change over several centuries and in projecting the response of the earth's climate system to greenhouse gas emissions. The action space (emissions reductions) and state space (all variables required to describe the evolution of the system over time) are continuous variables that may need to be discretized. The dimensionality of this problem, even in a highly simplified form, is extremely large.

One important complication associated with current integrated assessment models is that arguments for near-term emissions reductions are motivated less by the value of the emissions they avoid in this long-term problem and more by the possibility that policies today will encourage technological change which will lower future abatement costs. To explore this argument in a rigorous framework requires modeling endogenous or decision-dependent uncertainties, since the decision to abate today will change the probability distribution of next period's abatement costs. Modeling decision-dependencies of this type poses a unique challenge to conventional stochastic programming methods, which typically use exogenous scenario trees. Existing stochastic programming methods that attempt to model decision-dependent uncertainties cannot be applied to this model because they apply only under very specific circumstances. For example, Goel and Grossman [9] present a framework in which decisions affect the time in which the uncertainties will be resolved. Baker and Solak [2] introduce a stochastic programming version of an IAM with endogenous decision-dependent probabilities, but one that uses a customized deterministic mapping function to assign outcomes to decisions. Decisions in climate policy analysis influence the probability of different outcomes, and the need to have a flexible way to capture endogenous uncertainties means that Stochastic Dynamic Programming (SDP) is an appropriate framework for climate policy analysis. Unfortunately, classical SDP algorithms (e.g., value iteration, policy iteration [4]), suffer from the curse of dimensionality; i.e., the complexity of the problem grows exponentially with the number of states.

There have been a few studies that have formally framed the climate decision problem under uncertainty as a multi-stage stochastic dynamic program, using a variety of approaches to overcome the dimensionality challenge. Gerst et al. [8] use discrete sampling via experimental design and a very large number of iterations to learn about the solution space, which can be computationally expensive. Kelly and Kolstad [13] and Leach [15] approximate the value function associated with the Bellman equation using neural networks to estimate a functional form with 16 terms, but use discrete gridded samples in state-space to iteratively improve the approximation. Crost and Traeger [6] and Lemoine and Traeger [16] statistically estimate relationships between state variables offline in order to reduce the dimensions of the state vector, and then use conventional backward induction on the reduced state-space. All of these approaches rely on discretizing a (possibly reduced) state-space into intervals, and therefore require difficult tradeoffs between resolution/accuracy and computation time.

Here we present an alternative efficient solution method for multi-stage, multi-dimensional stochastic dynamic programs, based on Approximate Dynamic Programming (ADP)[4,25]. In this approach, we approximate the value function with a continuous function, which avoids the resolution and computational issues of discretized approaches. A key challenge associated with the successful application of the ADP methodology is the specification of the set of basis functions used to construct an approximate value function. The solution obtained via ADP methods is known to be sensitive to the choice of basis

functions that are used to build the value function. If the true value function is not spanned by this basis then the ADP algorithm will converge to the wrong solution. This false convergence is difficult to detect in practice.

We address this issue using two alternative approaches for value function approximations, one parametric, using global regression, and one non-parametric, using a mesh-free moving least squares approach. The parametric method is in principle faster but may exhibit the false convergence issue discussed above. The non-parametric method may be slower but can be used to detect errors in the choice of basis functions. We develop and test our algorithm using a stochastic dynamic programming version of the Dynamic Integrated model of Climate and the Economy (DICE) [21]. We demonstrate that for this application, ADP has several advantages over alternative solution methods including the ability to model decision-dependent uncertainties, manage a high-dimensional state space over a multi-stage stochastic decision problem, and converge in a fraction of the computational time. The results of the analysis show that an increase in uncertainty in future abatement costs results in a slight reduction in the optimal level of near-term emissions reductions in the standard model. However, once a probabilistic decision-dependent effect is included, the optimal near-term emissions reductions are greater than the expected value case.

We describe the DICE model, the formulation of the stochastic version, and the algorithms for solution using ADP in Section 2. Section 3 validates the new algorithms. In Section 4, we present the results of simulations of the base model using both parametric and non-parametric value function approximations, as well as the results of the decision-dependent variation. Section 5 gives a concluding discussion and suggests directions for future research.

## 2 Methods

Integrated assessment models [18, 33] are a general class of models that couple economic growth equations with differential equations that describe the transient evolution of the biogeophysical earth system. IAMs fall into two broad subgroups, policy evaluation models which simulate exogenous emissions policies, and policy optimization models which are optimal control models. The model described below falls into the latter category. We illustrate our computational solution algorithm on one such model, but the techniques are broadly adaptable to other models in the same class as well as other intertemporal optimization models.

2.1 The DICE Model

The effect of learning on optimal policy choice is calculated using a stochastic version of the DICE-99 model [21][1]. 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. Specifically, DICE is a deterministic, constrained non-linear program which chooses investment $I(t)$ and abatement $\mu(t)$ in order to maximize the sum of discounted utility:

$$\max_{I(t),\mu(t)} \sum_{t=0}^{T} U(c(t), L(t))(1 + \rho(t))^{-1}, \tag{1}$$

where $U(\cdot, \cdot)$ is the utility function, $c(t)$ is the per capita consumption, $L(t)$ is the population, and $\rho(t)$ is the social rate of time preference. The maximization is subject to a set of constraints, including the production function for the economy, the relationship between economic output and emissions, relationships for concentrations, radiative forcing, temperature change, and the reduction in output from both abatement costs and damage costs. The full set of equations for the model is given in [21]. The time horizon of the model is 350 years in 10 year steps.

2.2 Formulation of the Decision under Uncertainty Problem

Parameters in the DICE model are uncertain, as clearly we do not have perfect information about future economic growth and technological change or a complete understanding of the earth's climate system. But the uncertainty in some parameters are more important than others in terms of their effect on optimal abatement decisions. Nordhaus and Popp [22] performed uncertainty analysis of the DICE model and concluded that the most critical parameters are those that determine the costs of emissions reductions and those that determine the ultimate damages from temperature changes (as opposed to uncertainties in baseline projections). Decision under uncertainty in climate damages have been more fully examined by others [6,10,14,28,36]. However, Kelly and Kolstad [13] and Webster et al. [30] have shown that it may take a very long time before the uncertainty in damages is reduced. In contrast, some of the uncertainty in the costs of emissions reductions may be reduced sooner. However, the time in which new information is obtained depends on the level of abatement attempted. In this analysis, we focus on the uncertainty in the cost of abatement. The problem becomes one of choosing a level of abatement in each

---

[1] Newer versions of DICE exist [20], however we use DICE-99 for its relative simplicity and because the subsequent updates do not change the qualitative points being made here.
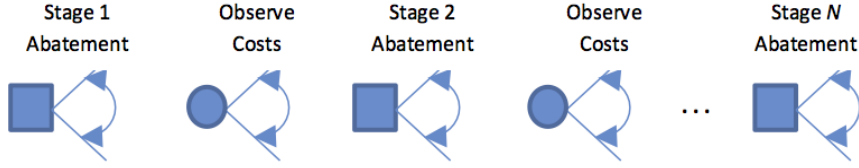
**Fig. 1** Schematic of Sequential Abatement Decision under Uncertainty

decision stage under uncertainty in abatement costs, after which information is received about costs that may shift the expected future costs, which are still uncertain. In the next stage, abatement levels are chosen again after observing the realized costs in the previous period. The decision problem under uncertainty is illustrated in Figure 1, using a skeleton of a decision tree, assuming a finite set of $N$ decision stages. Mathematically, the stochastic problem is that in equation (2), where $\mu_t$ is the abatement level in stage $t$, $\theta_t$ is the cost shock in stage $t$, and $R_t$ is the discounted utility in stage $t$.

$$\max_{\mu_1} \left\{ R_1 + \max_{\mu_2} E_{\theta_1} \left[ R_2 + \ldots \right] \right\}. \tag{2}$$

We implement and solve the stochastic version of the model using the framework of stochastic dynamic programming. Dynamic programming uses the Bellman equation [3] to decompose the problem in (2) into the relatively simpler condition that must hold for all decision stages $t$:

$$V_t = \max_{\mu_t} \left[ R_t + E \left\{ V_{t+1}(\mu_t, \theta_t) \right\} \right]. \tag{3}$$

The expression in (3) clarifies the nature of the stochastic optimization problem; just as the deterministic DICE finds the intertemporal balance between the costs of reducing emissions now and the future benefits of avoided climate damage, the stochastic problem is to find the optimal balance between near-term costs and expected future costs. As an illustration, Figures 2 and 3 show the near-term costs and expected future costs, respectively, that are balanced in the first of the two numerical experiments presented below.

In the deterministic version of DICE, abatement cost as a percentage of output (GDP) is a function of the abatement decision variable,
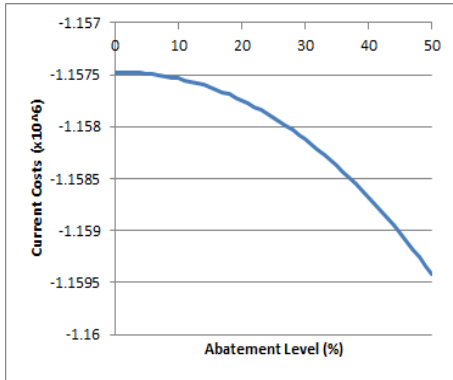
$$AC = 1 - c_1 \mu^{c_2}.$$

Where $c_1$ and $c_2$ are calibrated [21] to obtain values of $c_2 = 2.15$ and $c_1$ starts at a value of 0.03 and grows over time. The growth rate of the cost coefficient declines over time as
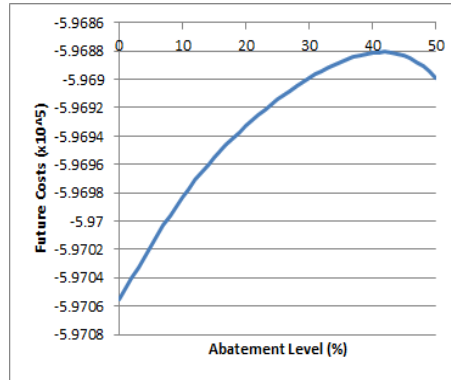
$$g_c(t) = -0.08 e^{-0.08t}.$$

The cost coefficient grows as,

$$c_1(t) = \frac{c_1(t-1)}{(1 - g_c(t))}.$$

**Fig. 2** Near-term costs as a function of control rate (%) in first decision stage when abatement cost is uncertain.



**Fig. 3** Expected future costs as a function of control rate (%) in first decision stage when abatement cost is uncertain.

We represent uncertainty in future abatement costs as a multiplicative shock in each period to the reference growth rate of costs $g_c(t)$. In the results presented below, the reference distribution for the cost growth rate shock is assumed to be Normal with a mean of 1.0 and a standard deviation of 0.4. The uncertainty in abatement cost is based on a detailed uncertainty analysis of a higher resolution economic model of emissions [31]. We also show results for a range of standard deviations for the cost shock.

For ease of exposition, we have made a few simplifying assumptions in the stochastic decision model. First, we fix the investment in capital stock in the economy $I(t)$ to the optimal trajectory from the deterministic model, because it is largely unresponsive to changes in abatement decisions or in abatement cost assumptions. The DICE model is defined in 10-year steps over a 350-year time horizon (35 model periods). Rather than define each decade as a decision stage, each decision stage in the stochastic model consists of 5 DICE model periods (50 year steps), for a total of seven decision stages ($N = 7$). Fewer multi-decade decision stages make the communication of results easier but more importantly better characterize the stochastic process being modeled. The lifetime of many of the large capital investments that are affected by the abatement decision, such as a coal-fired power plant, typically have lifetimes of $30 - 50$ years. Similarly, information about technological breakthroughs that drive the changes in the abatement costs may not occur every 10 years, nor do large policy shifts. The seven-stage model presented here approximates the time scale of the problem, while having significantly higher resolution than the two-stage model approaches that are most common in the literature.

In addition to exploring the effect of abatement cost uncertainty on near-term decisions, we also wish to explore the impacts of decision-dependence. For this purpose, we developed a second version of the stochastic model. In this version, we assume that abatement $\mu$ in one decision stage lowers the

mean of the cost distribution as

$$\widetilde{c}_1(t) = \bar{c}_1(t)(1 - \alpha\mu(t-1)).$$

Where $\widetilde{c}_1$ is the random variable for the cost coefficient in the decision-dependent version and $\bar{c}_1$ is the cost coefficient in the reference model, and $\alpha > 0$ is a scaling constant that alters the magnitude of the decision-dependent effect.

## 2.3 Approximate Dynamic Programming Implementation

The finite-horizon stochastic dynamic programming problem formulated above is traditionally solved as a Markov Decision Problem [4], using a backward induction algorithm. The algorithm iterates over the state, action, and uncertainty spaces for each decision stage to calculate the exact value function and corresponding policy function in each decision stage. Because the action and state spaces are all continuous, this would require discretization for each variable. For the DICE model, there are seven state variables that must be known; the capital stock ($K(t)$), three carbon concentration variables for a three-box model of the carbon cycle ($MAT(t), MU(t), ML(t)$), two temperature variables for a two-box energy-balance model ($TE(t), TL(t)$), and the evolving abatement cost coefficient ($c_1(t)$). All of these variables require knowledge of the previous value to calculate the next value (see [21]).

In addition to the state variables, conventional dynamic programming would also iterate over discretized values of the action $\mu(t)$ and the cost growth shock $\theta(t)$, resulting in at least a 9-dimensional problem in each of 7 decision stages. This is an extremely large problem even if the discrete intervals are at unsatisfyingly coarse resolution.

Instead of traditional backward induction, we have developed an approximate dynamic programming (ADP) algorithm for solving this problem, shown in Algorithm 1 (reference to only one of two value function approximations is made). ADP is a class of methods (e.g.,[5,25]) that approximates the value function in each stage by adaptively sampling the state space to focus on higher expected value states until the value function converges. One critical advantage of forward sampling is that this enables a straightforward representation of decision-dependency. Two key elements in any efficient ADP algorithm are 1) the sampling strategy, and 2) the value function approximation.

Our solution algorithm consists of two phases. In phase I, the bootstrap phase, we use Latin Hypercube Sampling [19] to explore both the action space over all stages and the cost shock space. These sample paths are simulated forward, and the resulting Bellman values for the sample states and actions are saved for each decision stage. The full set of these samples of the value function are used to produce the first estimate of the value function approximation for each decision stage, using either of the two methods described below.

In phase II, we randomly sample the cost shock in each period to obtain a sample path, and choose the optimal action in each stage using the current value function approximations for the value of the next state, and the simulated

---

**Algorithm 1:** DICE Approximate Dynamic Programming Algorithm

---

**Input**: Decision stages $N$, bootstrap iterations $bs$, possible controls $\mu$, uncertainty variable $\theta \sim N(1, \sigma)$, system state $s_0 \in S$ at time $t_0$, system state transition equations $F(\mu, \theta)$, convergence criterion, $\bar{\epsilon}$

**Phase I Initialization-Bootstrap**: While $i \leq bs$,

**1. Forward Pass**
**Loop** over $t$ from **1 to** $N$, Latin Hypercube Sampling from $\theta$ and $\mu$ and set current reward as:

$$R_t(s_i) = U(c_t, L_t)(1 + \rho_t)^{-1}.$$

**2. Backward Pass**
**Loop** over $t$ from $N$ **to 1**, setting the Bellman Value as:

$$v_t(s_i) = (R_t(s_i) + v_{t+1}(y_i|s_i))$$

where $y_i$ is the sampled next system state resulting from $\mu_t$ and $\theta_t$, and $v_N$ is a pre-defined terminal value.

**3. Construct First Estimate of Value Function**: When $i = bs$, use OLS to set:

$$\widehat{v}_t(s) = \Phi(s)r_0,$$

where $\Phi$ is a row vector of basis functions and $r_0$ is a column vector of coefficients that solves:

$$\min_{r_0} \sum_{s_i} (\widehat{v}_t(s_i) - \Phi(s_i)r_0)^2.$$

for all sample states $s_i$.

**Phase II Main Loop-Optimization**: While $i > bs$,

**1. Forward Pass**
**Loop** over $t$ from 1 to $N$, sampling $\theta$ randomly and sampling controls $\mu$ that achieve:

$$\max_{\mu} \; [R_t(s_i) + E\{v_{t+1}(y_i|s_i)\}]$$

where

$$E\{v_{t+1}(y_i|s_i)\} = \widehat{v}_{t+1}(\mu_t, \theta_t).$$

Set current reward, $R_t(s_i)$, as in Phase I.

**2. Backward Pass**
**Loop** over $t$ from $N$ **to 1**, setting the new Bellman Value as:

$$v_t(s_i) = (R_t(s_i) + \widehat{v}_{t+1}(y_i|s_i))$$

where $y_i$ is the sampled next system state.

**Update** $r_i$ using a Bellman Error routine:

$$r_{i+1} = r_i - \gamma_i \epsilon_i \nabla r_i$$

where $\gamma_i$ is a predefined smoothing parameter and

$$\epsilon_i = v_t(s_i) - \widehat{v}_t(s_i).$$

**Exit** when:

$$\bar{\epsilon} = |\bar{v}_{1,i} - \bar{v}_{1,i-1}|$$

where $\bar{\epsilon}$ represents the change in the moving average of the total Bellman value in the initial stage.

**Output**: Optimal first-stage control, $\mu_1^*$, value function approximations, $v_t^*(s)$

---

DICE equations to obtain the current reward. The overall sampling approach is an efficient (stratified) pure explore strategy in Phase I and a pure exploit strategy in Phase II.

In this study, we compare two alternative approaches to value function approximation, one parametric and one non-parametric. Both approaches approximate the expected value of being in any state as a reduced-form function of key features. Because of the forward sampling, not all state variables required for backward induction are needed as the key features or basis functions [5]. For this application, the fundamental structure is one of balancing near-term costs of abatement (reducing the size of the economy) against long-term costs from climate change. In terms of the state variables described above, the key features needed to approximate the value function are the capital stock $K(t)$ and the global surface temperature change $TE(t)$. The parametric approach employed is an iterative least squares regression method [4], approximating the value function as a nonlinear function of capital stock and temperature change. That is, the approximation of the value function is

$$\widehat{v}_t(s) = \Phi(s)r \tag{4}$$

where $\Phi$ is a row vector of basis functions and $r$ is a column vector of coefficients that solves,

$$\min_r \sum_{s_i} (\widehat{v}_t(s_i) - \Phi(s_i)r)^2.$$

for all sample states $s_i$. Given an initial estimate of the coefficient vector $r$ from the bootstrap phase, we iteratively improve the estimate using a Bellman error approach [4].

We compare an iterative regression approach with a non-parametric alternative. In this second approach, we apply moving least squares (MLS) [7] to interpolate the value function at a given state within a neighborhood. Meshfree methods such as MLS have been applied to other problems requiring interpolation in high dimensional space such as scattered data modeling, the solution of partial differential equations, medical imaging, and finance [7]. In the context of stochastic optimization, MLS was applied in [23] in an iterative algorithm that solves for the stochastic maximum principle. Here we apply the method in the context of the dynamic programming principle.

The approximate value of a state $s$ is:

$$\widehat{v}_t(s) = \bar{\Phi}(s)\bar{r}(s) \tag{5}$$

The difference between equations (4) and (5) are that the coefficient vector $\bar{r}$ depends on the state s. Note that $\bar{r}(s)$ is obtained by solving

$$\min_{\bar{r}} \sum_{s_i} (\widehat{v}_t(s_i) - \bar{\Phi}(s_i)\bar{r}(s_i))^2.$$

for all sample states $s_i$ within some neighborhood of the state $s$. This requires solving many regressions, one for each point to be interpolated, as compared

with the parametric approach. However, these regressions are generally for a small number of samples in the immediate neighborhood, whereas the parameteric approach is global and uses all samples which can grow to a large number. Thus, the tradeoff is between many small regressions (MLS) versus fewer large regressions (parametric). Furthermore, by using linear basis functions and relatively small neighborhoods, this approach can approximate a large class of value functions, which may not be true for global approximations with any fixed set of basis functions. To store samples from all previous iterations and efficiently search for samples within a given neighborhood, we use a kd-tree data structure [7].

## 3 Model Validation

### 3.1 Comparison with Deterministic Version

We first validate both ADP implementations against the results of the original DICE model as a deterministic NLP. For ease of comparison, we modify the deterministic version of DICE to choose a single emissions control rate for each 50-year decision stage, rather than a distinct control rate for each decade. For the ADP version, we eliminate the uncertainty in abatement cost growth rates in order to approximate the deterministic solution. Figure 4 shows the resulting optimal decisions, as the fractional reduction below the reference emissions. Both implementations consistently converge to within 5% of the NLP results. To test for convergence, we use the change in the moving average of the total Bellman value in the initial stage:

$$\bar{\epsilon} = |\mu_V^i - \mu_V^{i-1}|. \tag{6}$$

Because convergence is stochastic, we use a running average of 1000 sample values. The evolution of the convergence criterion over iterations is shown in Figure 5 for representative solutions of the global regression and the moving least squares algorithms. In general, the MLS algorithm converges in fewer iterations, typically 2000 iterations beyond the bootstrap for MLS as compared with over 10000 iterations for global regression to achieve a convergence criterion below 1e-7. The tradeoff in computation speed is less clear, because each iteration of the MLS algorithm is more expensive, including augmenting kd-tree, numerous searches of the kd-tree for nearest neighbors, and numerous regressions. The total computation time to achieve convergence to less than 1e-7 is roughly equivalent, the results are comparable. For the remainder of this paper, we present only the results from the global regression version.

### 3.2 Comparison with Three-Period Stochastic Version

Next, we validate both ADP implementations against the results of a stochastic dynamic program (SDP) version of DICE solved using a traditional backward
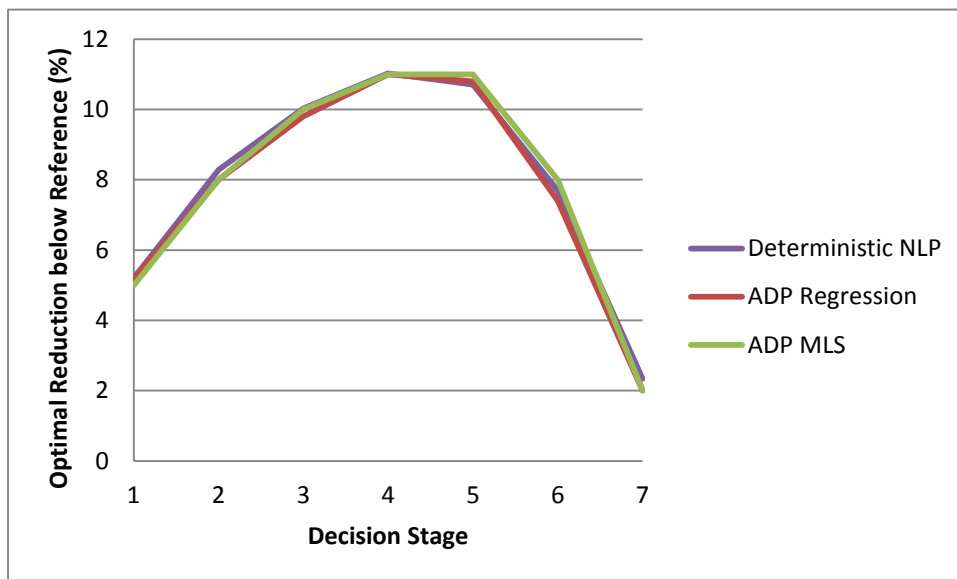
**Fig. 4** Optimal control rate (%) in all decision stages when abatement cost is deterministic.
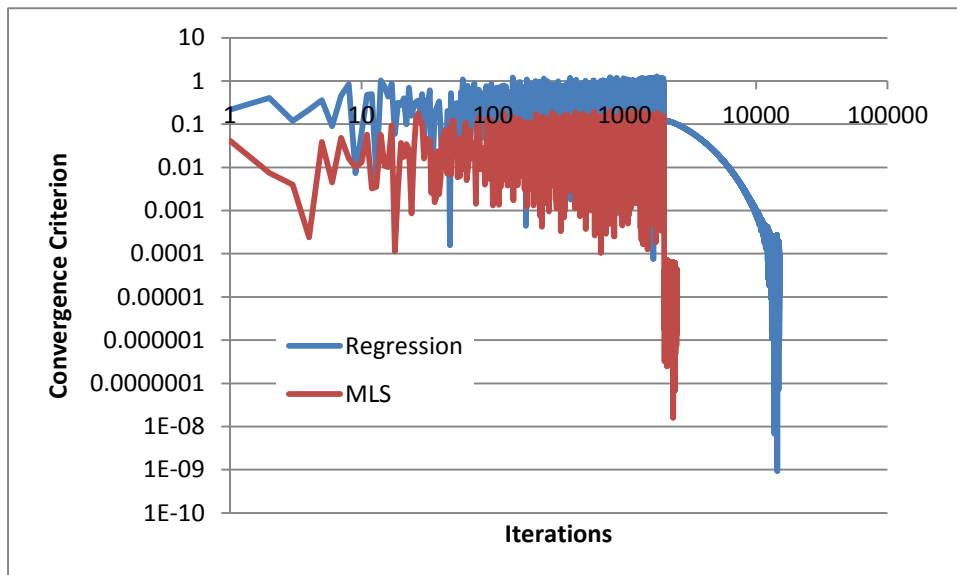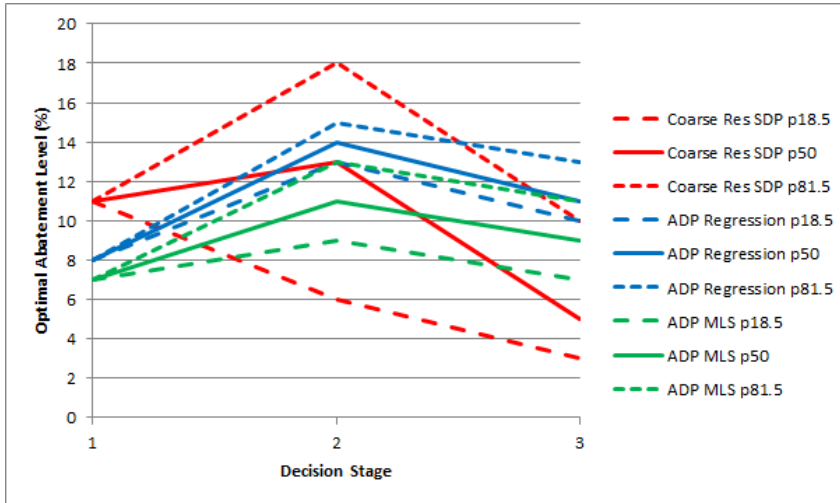


**Fig. 5** Convergence criterion by iteration for global regression and moving least squares algorithms.

**Fig. 6** Comparison of 3-Stage Stochastic Dynamic Programming Solution with ADP solution. Percentiles shown correspond approximately to the three-point Pearson-Tukey discretization used in the SDP.

induction algorithm. Given the need to iterate over discrete actions $\mu(t)$, cost growth shocks $\theta(t)$, and the DICE model's seven state variables, we choose a three-stage model as the maximum-dimension model that can be solved using traditional numerical techniques within a reasonable timeframe. Experiments with simplifying techniques to run a longer model failed to produce reasonable results. For example, a technique to reduce the form of the (deterministic) model by exploiting the physical correlations between various state variables yields accurate early-stage optimal actions, but later-stage results that match neither the ADP or NLP model.

The three-stage models choose optimal abatement levels at an initial stage, at 100 years, and at 200 years. To construct the SDP, the action space $\mu$ is discretized in steps of one-percent from zero abatement to twenty-five percent abatement. The original growth cost shocks, $\theta(t) \sim N(1, 0.4)$ are discretized using a Pearson-Tukey three-point approximation [12], at 0.34, 1.00, and 1.66 with probabilities of 0.185, 0.630, and 0.185, respectively. Finally, each of the state variables are divided into six discrete equal intervals. For the purposes of comparison, we keep all other parameters between the ADP and SDP models the same. Even with the extremely unsatisfying coarse resolution of the SDP, the first-period optimal $\mu$ is roughly consistent with the ADP solutions. The SDP result also exhibits a similar behavior and range for later stage decisions as the ADP implementations, as shown in Figure 6.
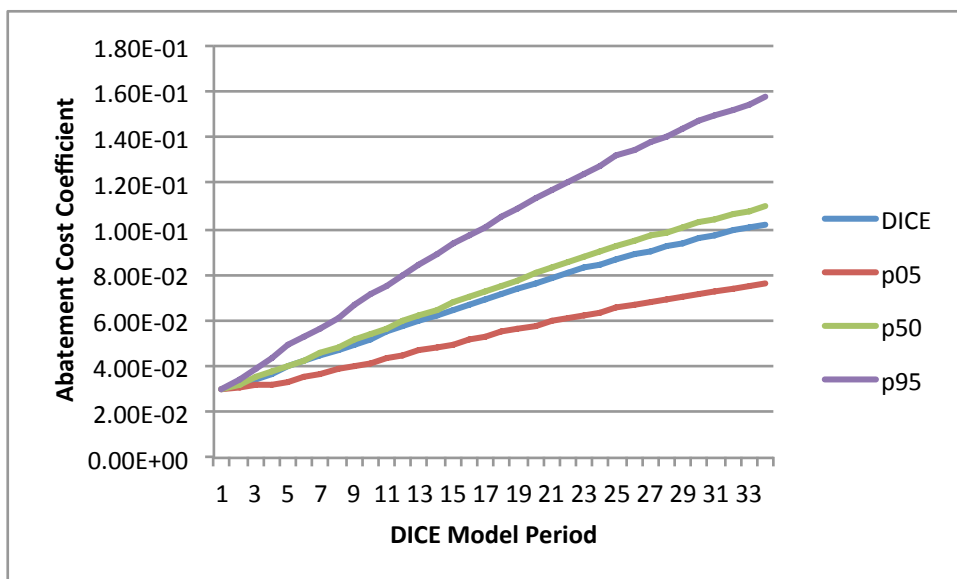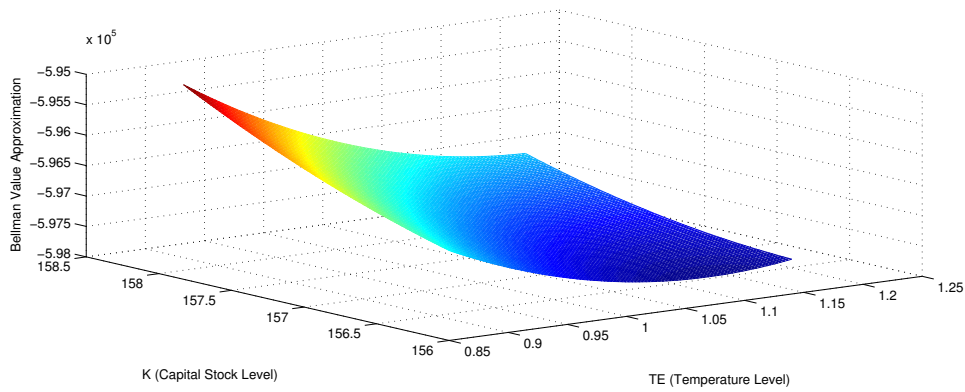
**Fig. 7** Uncertainty in abatement cost coefficient based on 20000 iterations and normally distributed multiplicative shock with mean 1 and standard deviation 0.4.

## 4 Results of Numerical Experiments

4.1 Results with Uncertainty in Abatement Cost

We now introduce uncertainty in the growth of the abatement cost coefficient as described above. We assume a reference value for the standard deviation in cost shocks of 40% based on [31]. Figure 7 shows the range of uncertainty in the cost coefficient that results from 20000 samples in each stage of the normally distributed multiplicative shock applied to the growth rate. Using the ADP algorithm with global regression, we solve a seven-stage stochastic dynamic program with uncertain growth in abatement cost in each stage. The second-stage total value function for this program is shown in Figure 8; we present only the second-stage value function for brevity. The median and upper and lower 90% range of the resulting optimal control rate in each stage are shown in Figure 9. Although this figure shows the unconditional range of control rates in each stage, the actual range in each stage $t > 1$ will depend on the previous shocks and decisions. For example, conditional on a stage 1 cost shock of 1.4, the 90% range on optimal stage 2 control is 7-9%, while conditional on a stage 1 cost shock of 0.6, the 90% range on stage 2 optimal control is 2-4%. In this version of the model, the initial optimal control rate is slightly decreasing in the variance of the cost uncertainty (Figure 10).

**Fig. 8** Second-stage Bellman Value function by state variables, Capital Stock (K) and Temperature (TE).
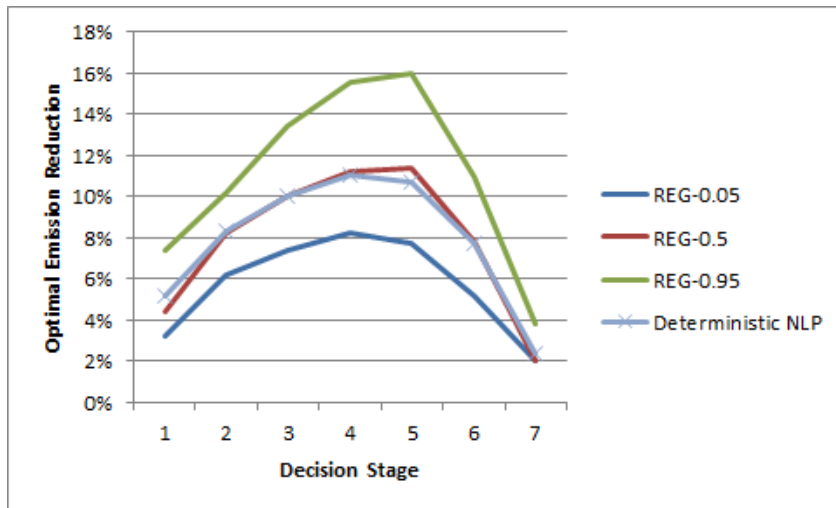


**Fig. 9** Optimal control rate over decision stages with uncertain abatement costs.

4.2 Results with Decision-Dependent Uncertainty in Abatement Cost

In this section, we show the results of adding endogenous decision-dependency to the model. As described in Section 2.2, the mean of the distribution of abatement costs in each stage is now a function of previous abatement decisions. This stylized model is intended to capture the generic effects of assuming that emissions limits induce technological change, which in turn lowers the costs of future abatement. Although not resolved in this model, this phenomenon could represent either learning by experience in technological change [1, 35, 34] or induced R&D spending leading to technological change [24, 11].
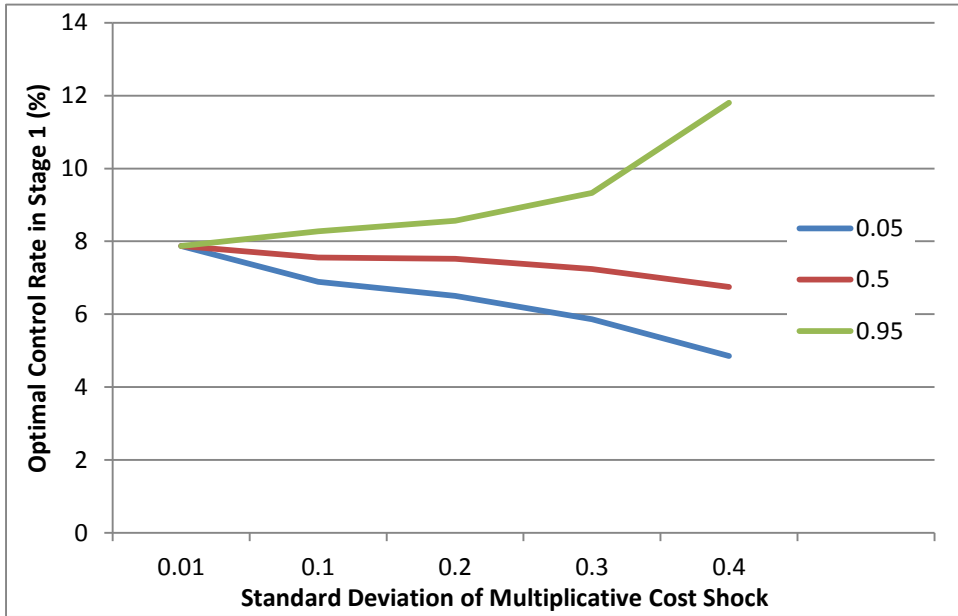
**Fig. 10** Optimal control rate (%) in first decision stage, shown as median and upper/lower 90% bounds, as a function of the standard deviation of the cost growth rate shock.

Figure 11 shows the median and 90% range of the optimal control rate in each decision stage with uncertain decision-dependent abatement costs ($\alpha = 0.2$ in equation 7). Also shown is the range of optimal controls without decision-dependency ($\alpha = 0$) from the previous section. Note that the effect of endogenous decision-dependency is to increase the optimal control rate in all stages. This is not surprising, since the approach here results in a distribution of abatement costs with a lower mean. In terms of the Bellman equation, there is additional marginal value in the next stage ($t+1$) to each unit of abatement in period $t$. This source of additional value in terms of technological change is often a primary motivation for near-term emissions reductions. However, it is typically not represented within integrated assessment models except for specialized versions focused on developing representations for technical change. Although this stylized model is not the appropriate tool for determining the empirical strength of the decision-dependent effect, a sensitivity analysis over a range of values (Figure 12) suggests that the effect need not be very significant to alter optimal control rate decisions.

## 5 Discussion

Decision problems about how to respond to climate change will occur gradually over very long time-scales, under a great deal of uncertainty, and with learning along the way about these uncertainties. Further, sociotechnical systems,
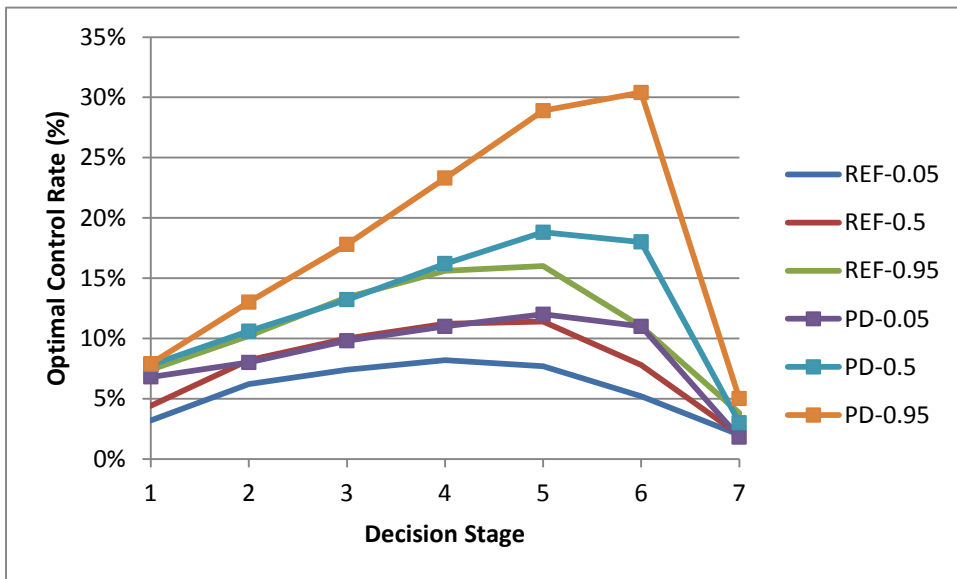
**Fig. 11** Optimal control rate over decision stages with uncertain abatement costs with and without decision-dependency.
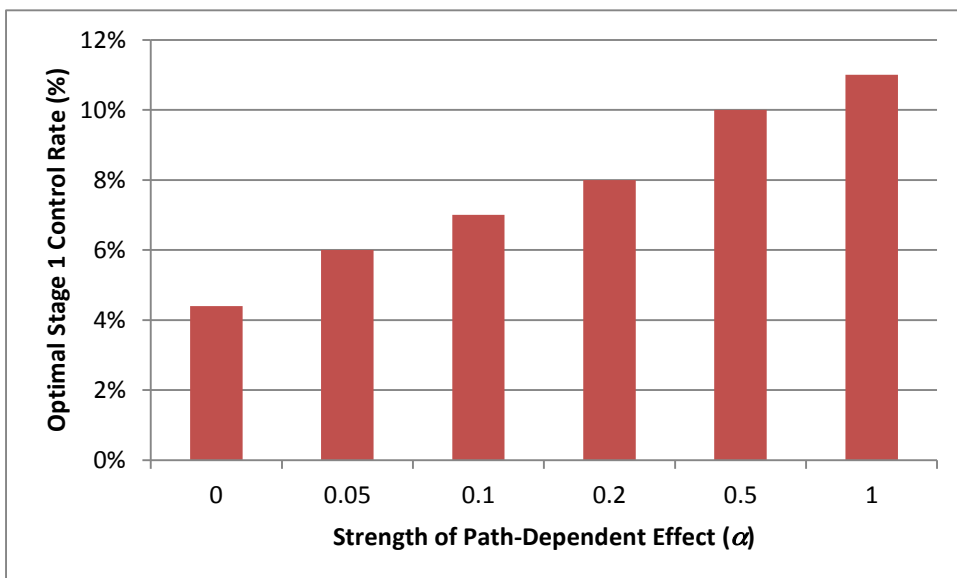


**Fig. 12** Optimal Stage 1 control rate for different strengths of the decision-dependency.

such as the global economic-energy-emissions nexus, often exhibit feedbacks between decisions and uncertainties over time, in contrast to assumptions of exogeneity conventionally used to keep computations tractable. Use of optimization and simulation models of such complex systems typically neglect uncertainty or simplify the model to two decision stages, where uncertainty is resolved in the second period. To explore more realistic decision problems, we require algorithms that can tractably solve multi-period stochastic optimization with multi-dimensional state variables and feedbacks between decisions and uncertainties.

We have demonstrated two variants on such an algorithm here, and applied it to the problem of optimal emissions abatement under abatement cost uncertainty to respond to climate change. These algorithms converge quickly without having sacrificed resolution in state space, action space, or number of decision stages.

The general insights for climate policy are that 1) in the presence of abatement cost uncertainty, near-term emissions reductions should be lower the greater the uncertainty, and 2) if future (uncertain) abatement costs depend on current abatement decisions, then optimal near-term emissions reductions will be greater. We have compared two alternative methods for value function approximation, one parametric and one non-parametric. The non-parametric approach using moving least squares converges in fewer iterations, but each iteration is more computationally expensive. This particular application happens to have smooth value function surfaces that are well approximated by global second-order polynomials. However, other questions may involve highly nonlinear or discontinuous surfaces, such as considering threshold effects or tipping points in the climate system. We expect that the non-parametric approach will have significant advantages for such applications, and constitutes a basis of our future work with these algorithms.

# References

1. Arrow, K.: The economic implications of learning by doing. Review of Economic Studies **29**, 155–173 (1962)
2. Baker, E., Solak, S.: Optimal Climate Change Policy: R&D Investments and Abatement under Uncertainty (2011). Under Review
3. Bellman, R.: Dynamic programming. Dover Pubns (2003)
4. Bertsekas, D.: Dynamic programming and optimal control. Athena Scientific Belmont, MA (2007)
5. Bertsekas, D., Tsitsiklis, J.: Neuro dynamic programming. Athena Scientific Belmont, MA (1996)
6. Crost, B., Traeger, C.: Risk and aversion in the integrated assessment of climate change (2010). CUDARE Working Paper No. 1104
7. Fasshauer, G.: Meshfree approximation methods with Matlab. World Scientific Pub Co Inc (2007)

8. Gerst, M., Howarth, R., Borsuk, M.: Accounting for the risk of extreme outcomes in an integrated assessment of climate change. Energy Policy **38**(8), 4540–4548 (2010)
9. Goel, V., Grossmann, I.: A class of stochastic programs with decision dependent uncertainty. Mathematical programming **108**(2), 355–394 (2006)
10. Hammitt, J.K., Lempert, R.A., Schlesinger, M.E.: A sequential-decision strategy for abating climate change. Nature **357**, 315–318 (1992)
11. Jaffe, A., Newell, R., Stevins, R.: Handbook of Environmental Economics, Vol 1., chap. Technological change and the environment, pp. 461–516. North-Holland/Elsevier (2003)
12. Keefer, D., Bodily, S.: Three-point approximations for continuous random variables. Management Science **29**(5), 595–609 (1983)
13. Kelly, D., Kolstad, C.: Bayesian learning, growth, and pollution. Journal of Economic Dynamics and Control **23**, 491–518 (1999)
14. Kolstad, C.: Learning and stock effects in environmental regulation: The case of greenhouse gas emissions. Journal of Environmental Economics and Management **31**, 1–18 (1996)
15. Leach, A.: The climate change learning curve. Journal of Economic Dynamics and Control **31**, 1728–1752 (2007)
16. Lemoine, D., Traeger, C.: Tipping points and ambiguity in the integrated assessment of climate change (2011). NBER Environmental and Energy Economics Summer Institute 2011 Paper
17. Manne, A., Richels, R.: The costs of stabilizing global co2 emissions: A probabilistic analysis based on expert judgment. The Energy Journal **15**(1), 31–56 (1994)
18. Martens, P., Rotmans, J.: Climate change: An integrated perspective. In: P. Martens, J. Rotmans, D. Jansen, K. Vrieze (eds.) Climate Change: An Integrated Perspective, *Advances in Global Change Research*, vol. 1, pp. 1–10. Springer Netherlands (2003)
19. McKay, M.D., Beckman, R.J., Conover, W.J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics **21**(2), 239–245 (1979)
20. Nordhaus, W.: The challenge of global warming: Economic models and environmental policy. Available at: http://nordhaus.econ.yale.edu/ (2007). NBER Working Paper 14832
21. Nordhaus, W., Boyer, J.: Warming the world: economic modeling of global warming. Massachusetts Institute of Technology Press, Cambridge, MA, USA (2000)
22. Nordhaus, W., Popp, D.: What is the value of scientific knowledge? an application to global warming using the price model. The Energy Journal **18**(1), 1–45 (1997)
23. Parpas, P., Webster, M.: A Stochastic Minimum Principle and a mesh-free method for Stochastic Optimal Control (2011). Submitted
24. Popp, D., Newell, R., Jaffe, A.: Energy, the environment, and technological change. howpublished (2009). NBER Working Paper 14832
25. Powell, W.: Approximate Dynamic Programming: Solving the curses of dimensionality, vol. 703. Wiley-Blackwell (2007)
26. Reilly, J., Edmonds, J., Gardner, R., Brenkert, A.: Monte carlo analysis of the iea/orau energy/carbon emissions model. The Energy Journal **8**(3), 1–29 (1987)
27. Scott, M., Sands, R., Edmonds, J., Liebetrau, A., Engel, D.: Uncertainty in integrated assessment models: Modeling with minicam 1.0. Energy Policy **27**(14), 597–603 (1999)
28. Webster, M.: The curious role of learning: Should we wait for more data? The Energy Journal **23**, 97–119 (2002)
29. Webster, M.: Incorporating path-dependency into decision analytic methods: An application to global climate change policy. Decision Analysis **5**(2), 60–75 (2008)
30. Webster, M., Jakobovits, L., Norton, J.: A class of stochastic programs with decision dependent uncertainty. Climatic Change **89**(1-2), 67–85 (2008)
31. Webster, M., Paltsev, S., Parsons, J., Reilly, J., Jacoby, H.: Uncertainty in greenhouse emissions and costs of atmospheric stabilization. Tech. rep., MIT JPSPGC (2008). Report No. 165
32. Webster, M., Sokolov, A., Reilly., J., Forest, C., Paltsev, S., Schlosser, A., Wang, C., Kicklighter, D., Sarofim, M., Melillo, J., Prinn, R., Jacoby, H.: Analysis of climate policy targets under uncertainty. Tech. rep., MIT JPSPGC (2009). Report No. 180

33. Weyant, J., Davidson, O., Dowlabathi, H., Edmonds, J., Grubb, M., Parson, E., Richels, R., Rotmans, J., Shukla, P., Tol, R., Cline, W., Fankhauser, S.: Integrated assessment of climate change: an overview and comparison of approaches and results. In: Economic and Social Dimensions of Climate Change, Equity and Social Considerations. Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change, chap. Integrated assessment of climate change: an overview and comparison of approaches and results, pp. 367–396. Cambridge University Press (1996)
34. Wright, T.: Factors affecting the cost of airplanes. Journal of Aeronautical Sciences **3**, 122–128 (1936)
35. Wright, T.: Inside the Black Box: Technology and Economics. Cambridge University Press (1982)
36. Yohe, G., Andronova, N., Schlesinger, M.: To hedge or not against an uncertain climate future? Science **306**, 416–417 (2004)