

OPTIMAL CHOICE OF POLICY INSTRUMENT AND STRINGENCY UNDER UNCERTAINTY: THE CASE OF CLIMATE CHANGE

WILLIAM A. PIZER
RESOURCES FOR THE FUTURE

This version: March 3, 1997

Abstract

Considerable uncertainty surrounds both the consequences of climate change and their valuation over horizons of decades or centuries. Yet, there have been few attempts to factor such uncertainty into current policy decisions concerning stringency and instrument choice. This paper presents a framework for determining optimal climate change policy under uncertainty and compares the resulting prescriptions to those derived from a more typical analysis with best-guess parameter values. Uncertainty raises the optimal level of emission reductions and leads to a preference for taxes over rate controls. The first effect is driven primarily by uncertainty about future discount rates while the second arises because of relatively linear damages and a negative correlation between control costs and damages. Importantly, the welfare gains associated with policies computed from best-guess parameter values are significantly less than those which take uncertainty into account – on the order of 30%. This suggests that analyses which ignore uncertainty lead to inefficient policy recommendations.

1 Introduction

Uncertainty is a pervasive feature of climate change analysis. The wide range of divergent opinions on the proper policy response – ranging from a do-nothing approach to drastic emission reductions – is testament to the scale of the issue. However, there is a difference between recognizing uncertainty and using it to propose *one* policy response.

This paper presents a realistic assessment of uncertainty and optimizes over it to recommend a single course of action. This permits a useful measure of the importance of including uncertainty in the analysis. Compared to an optimization with a single set of central parameter values, does it generate different policy prescriptions? Do alternative policy instruments behave differently under uncertainty? Why? If uncertainty does lead to significantly different conclusions, it suggests that analyses which ignore uncertainty may be misleading. This is in fact the case: uncertainty leads to more stringent policy and a preference for taxes over rate controls. Both of these consequences involve significant welfare gains.

Many authors have made important contributions to the analysis of climate change policy under uncertainty but none have directly addressed how ignoring a realistic model of uncertainty can lead to significant policy errors. Much of the best known work in this area has focused on the question of learning: what happens if uncertainty is resolved now rather than later [Manne and Richels (1993,1995), Nordhaus (1994b), Manne (1996), Kelly and Kolstad (1996b), Nordhaus and Popp (1997)]. Other work has focused on sensitivity analysis: how do policy consequences and/or optimal policy choice vary when underlying parameters change [Nordhaus (1994b), Dowlatabadi and Morgan (1993), Dowlatabadi (1997) and Nordhaus and Popp (1997)]. Nordhaus (1994b), Manne (1996) and Nordhaus and Popp (1997) touch on the consequences of uncertainty for policy stringency, but do so with stylized models of uncertainty comprised of five or fewer states of nature. Use of such stylized models of uncertainty can generate misleading conclusions.

In this paper an integrated climate-economy model based on optimal intertemporal behavior is used to examine alternative policies under uncertainty. Several features distinguish

this model from other climate-economy models. First, a fast simulation technique is used to permit thousands of different states of nature to be simulated simultaneously, each capturing optimal consumer behavior. Second, the distribution of the economic states of nature is based on econometrically estimated parameter distributions. Third, these states – each of which exhibits different consumer preferences – are consistently aggregated using an explicit social welfare function.

Each of these features is important given the question to be addressed. A large number of states are necessary to explore the realistic policy consequences of uncertainty in a nonlinear model. Otherwise, important features of the uncertain parameter space may be concealed, biasing the results.¹ At the same time, intertemporal prices are likely to be wrong without a plausible model of consumer behavior. Such prices are essential for aggregating consequences over long time horizons. Econometric estimates of economic parameter distributions are important to capture well-known parameter relations and their associated uncertainty. In order to be consistent with historic interest rates, for example, time preference and intertemporal substitutability must be positively correlated. By how much and with what degree of certainty is a question best addressed by data. Finally, without a consistent method of aggregating over uncertain preferences, policy evaluation may be inadvertently skewed toward one set of preferences. Simple averaging of utility functions with different parameterizations, for example, can make the policy evaluation sensitive to the choice of units as discussed in Appendix C.

The results of this exercise indicate that uncertainty does in fact matter. Ignoring it leads to policy recommendations which are too slack, with welfare gains 30% below that of the

¹Intuitively, using only a few representative states of nature in a non-linear multidimensional model biases results when the choice of states misses important features of the parameter space. For example, consider trying to approximate the expectation of $\{y(x) : y(x) = -900, x \in [0, 0.3); y(x) = 0, x \in [0.3, 0.7); y(x) = +900, x \in [0.7, 1]\}$ for x distributed uniformly over $[0, 1]$. If x is first approximated by a random variable with two states, taken as the mid-points of $[0, 2/3]$ and $[2/3, 1]$ with probability $2/3$ and $1/3$, respectively, the expectation will be approximated as 300 rather than 0. If x is instead approximated by the midpoints of $[0, 1/3]$ and $[1/3, 1]$ with probability $1/3$ and $2/3$, respectively, this approximation becomes -300. The problem is that two states of nature cannot adequately capture the non-linearity of $y(x)$. Such problems are only exacerbated in higher dimensions.

optimal policy incorporating uncertainty. In 1995 per capita consumption terms, the policy ignoring uncertainty generates a \$58 gain versus \$73 when uncertainty is considered. Initially, the policies themselves are fairly close: the control rates are 7.9% and 8.4%, respectively, in 1995. Over time, however, these rates diverge with a difference of 20% in 30 years and over 50% after 80 years.

A related observation is that half of this observed increase in stringency can be reproduced by modeling preferences as the sole source of uncertainty (versus technology or climate parameters). This follows from the estimated distribution of risk aversion and time preference parameters which indicates that discount rates in the future may be close to zero in the presence of a slowdown in productivity growth. For an issue like climate change where the benefits from policy decisions in one period are spread out over a long horizon yet costs are incurred immediately, small changes in a discount rate near zero lead to large, asymmetrical swings in the desired level of stringency. This is analogous to small changes in the interest rate producing much larger changes in bond prices.

Finally, we observe that taxes are significantly more efficient under uncertainty than rate controls. While the optimal rate control policy generates welfare gains equivalent to a \$73 increase in current per capita consumption, the optimal tax policy generates an \$86 increase. The intuition for this result is that the damage function is relatively flat and that damages turn out to be negatively correlated with costs. Both of these features lead to a preference for price controls, as discussed by Weitzman (1974).

In order to understand the assumptions underlying these results, details of the model are first laid out in Section 2. This includes a discussion of the intertemporal model of the economy and climate, the aggregation of states into one welfare measure, and the quantification of uncertainty. Section 3 examines the question of optimal stringency under uncertainty. In particular, do models with and without uncertainty recommend the same rate of emission reductions? Section 4 explores the difference in efficiency between price and rate instruments as a mechanism for regulating carbon dioxide emissions. Section 5 concludes.

2 Model

2.1 Overview

The purpose of the model is to translate a climate change policy – either a tax or control rate – into a meaningful welfare measure of its consequences. This welfare measure, in turn, has two purposes: 1) to facilitate the calculation of optimal policies by maximizing the measure over a well-defined class of policies, and 2) to provide a metric for comparing policies drawn from different classes. Such classes might include taxes versus quantity controls, or policies which consider uncertainty versus those which ignore it.

Computing the welfare associated with a given policy can be viewed as a three-step process. First, for any given state of nature we need to determine how the economy and climate will evolve in the presence of the policy. This requires a model of consumer and producer behavior, a climate model, and links between the two. Second, a mechanism for comparing outcomes across states must be developed. When preferences vary from state to state, this is a non-trivial task even with a representative agent model within each state. Finally, uncertainty must be quantified in a reasonable way.

This paper approaches the first step by using a modified stochastic growth model to capture consumer behavior in each state. Standard stochastic growth models choose the level of aggregate consumption each period based on the optimal behavior of a representative consumer in the face of random IID shocks to productivity.² In contrast, this paper models the trend in productivity as both declining over time and endogenized to depend on climate change. However, these distinctions are ignored by the representative consumer.³ This strategy has two important consequences. Consumption remains optimally allocated across time in every state. Therefore, the intertemporal consequences of any climate change policy

²For a recent summary of the solution techniques for stochastic growth models see Taylor and Uhlig (1990).

³The productivity slowdown and economy-climate links are based on Nordhaus' (1994) DICE model. The fact that the consumer ignores these features has a negligible impact on the resulting path of consumption because the productivity slowdown and climate consequences (e.g., loss of output) occur quite gradually in the DICE model.

will be correctly valued. Equally important, for any given state the evolution of the economy and climate is easily determined: the 400 period path for each of 1000 states can be computed in less than 5 seconds on a desktop computer.

The second step develops a mechanism for aggregating the results over many states into a single welfare measure. Based on the idea of cardinal unit comparability among suitably scaled utility measures in each state, social welfare is computed as an arithmetic average. This approach requires a somewhat arbitrary choice of utility scaling and eliminates the possibility of distributional considerations. While subject to criticism on both issues, it highlights a point which has been skipped in much of the discussion of climate change: uncertainty about consumer/social preferences cannot be overcome without some aggregation scheme.

For the final step, uncertainty is quantified two ways. Economic uncertainty (describing preferences and technology) is based on an econometric analysis of historical data. Climate and long-term trend uncertainty (describing the climate consequences of economic activity, damages due to climate change, and eventual slowdowns in productivity and population growth) is based on discrete distributions developed in previous work.

2.2 State-contingent Model

2.2.1 Economic Structure

Economic behavior within each state is derived from a representative agent model where consumption must be optimally allocated across time. In a typical model with constant exogenous productivity growth, agent preferences define a steady state to which the economy converges over time. In the presence of random shocks and slowly changing trends, the economy instead converges to a distribution of states (due to the random shocks) which is itself slowly evolving (due to the slowly changing trends). For the moment we ignore these changing trends and focus on a standard stochastic growth model.

The representative consumer in this model exhibits constant relative risk aversion τ with

respect to consumption per capita. Utility is separable across time, discounted at rate ρ and weighted each period by population. With the further assumption that preferences satisfy the von Neuman-Morgenstern axioms, the consumer's optimization problem can be written as

$$\max_{\{C_t\}_{t \in \{0,1,\dots\}}} E \left[\sum_{t=0}^{\infty} (1 + \rho)^{-t} N_t \frac{(C_t/N_t)^{1-\tau}}{1-\tau} \right] \quad (1)$$

where C_t is consumption in period t and N_t is population. That is, the consumer maximizes expected discounted utility where each period's utility is population weighted. This consumption program $\{C_0, C_1, \dots\}$ is subject to the resource constraints describing production

$$Y_t = (A_t^* N_t)^{1-\theta} K_t^\theta \quad (2)$$

and capital accumulation

$$K_{t+1} = K_t(1 - \delta_k) + Y_t - C_t \quad (3)$$

where Y_t is aggregate output, $A_t^* N_t$ is effective labor input and K_t is the capital stock. A_t^* is a measure of productivity distinct from capital but not completely exogenous, as discussed later. The parameter θ summarizes the Cobb-Douglas production technology given in Equation (2) and δ_k reflects the rate of capital depreciation in the capital accumulation equation (3). Finally, there is a transversality condition for a balanced growth steady state,

$$\rho > (1 - \tau) \times (\text{asymptotic growth rate of } A_t^*) \quad (4)$$

This is always satisfied by assuming zero growth asymptotically.

Even with exogenous constant growth models for N_t and A_t^* , the dynamic optimization problem given by Equations (1–3) is difficult if not impossible to solve analytically.⁴ However, choosing

$$\Delta \ln(K_{t+1}/N_{t+1}) = \alpha_1 + \alpha_2(\ln(K_t/N_t) - \ln(A_t^*)) \quad (5)$$

⁴Long and Plosser (1983) derive an analytic solution for the case of $\delta_k = 1$ and $\lim \tau \rightarrow 1$ (log utility).

and

$$C_t = K_t(1 - \delta_k) + Y_t - K_{t+1} \quad (6)$$

– where α_1 and α_2 are functions of the parameters $(\rho, \tau, \theta, \delta_k)$ – yields a close approximation of optimal consumer behavior. This technique of approximating optimal dynamic behavior has its origins in the real business cycle literature beginning with Kydland and Prescott (1982). It is also related to the technique of feature extraction discussed by Bertsekas (1995). A simple derivation of expressions for α_1 and α_2 is given in Appendix A.

Intuitively, Equation (5) approximates behavior around a balanced growth steady state. At such a steady state, $\ln(K_t/N_t) - \ln(A_t^*)$ is constant and $\Delta \ln(K_{t+1}/N_{t+1}) = (\text{growth rate of } A_t^*) = \alpha_1 + \alpha_2(\ln(K_t/N_t) - \ln(A_t^*)) = \text{constant}$. If some unforeseen shock moves the economy away from the equilibrium value of $K_t/(A_t^*N_t)$ and α_2 is negative, e.g., the steady state is stable, then the economy will move back toward the steady state. In particular, when $K_t/(A_t^*N_t)$ is too high, capital accumulation will slow. If $K_t/(A_t^*N_t)$ is too low, capital accumulation will increase. Importantly, even if the growth rate of A_t^* is not constant, this approximation performs well as long as expected productivity growth changes gradually.

2.2.2 Long-term Growth, Climate Behavior and Damages

This section explains the remainder of the state-contingent model – specifically the evolution of A_t^* and N_t . This includes exogenous growth projections, climate behavior and damages from global warming [based primarily on the DICE model, Nordhaus (1994b)]. Exogenous labor productivity A_t is modeled as a random walk in logarithms with an exponentially decaying drift. That is,

$$\log(A_t) = \log(A_{t-1}) + \gamma_a \exp(-\delta_a t) + \sigma_a \epsilon_t \quad (7)$$

where γ_a is the initial growth rate, δ_a is the annual decline in the growth rate, σ_a is the standard deviation of the random growth shocks and ϵ_t is a standard NIID random shock. This means that productivity growth begins with a mean growth rate of γ_a (around 1.3%) in

the first period and eventually declines to zero. In addition, random and permanent shocks change the level of productivity every period. The standard error of these shocks is σ_a .

Net labor productivity A_t^* is distinguished from this exogenous measure A_t by the fact that A_t^* describes the amount of output available for consumption and investment – after output has been reduced by control costs and climate damages. To that end, A_t^* is expressed as A_t multiplied by a factors describing these two phenomena:

$$A_t^* = \left(\underbrace{\frac{1 - b_1 \mu_t^{b_2}}{1 + (D_0/9) \cdot T_t^2}}_{\substack{\text{control costs} \\ \text{damages}}} \right)^{\frac{1}{1-\theta}} A_t \quad (8)$$

μ_t is the fractional reduction in CO₂ emissions at time t (the “control rate”) versus a business as usual/no government policy baseline, while b_1 and b_2 parameterize the cost of attaining these reductions. Since b_1 and b_2 are both positive, additional rates of control involve reductions in net productivity. T_t is the average surface temperature relative to pre-industrialization in degrees Celsius and D_0 is the fractional loss in aggregate GDP from a 3° temperature increase. For temperature changes less than 10°, this is essentially a quadratic damage function.⁵ Additional details about the control cost and damage functions can be found in Nordhaus (1993) and Nordhaus (1994a), respectively.

Population is modeled in the same way as exogenous productivity but without the stochastic element:

$$\log(N_t) = \log(N_{t-1}) + \gamma_n \exp(-\delta_n t) \quad (9)$$

where γ_n is the initial growth rate and δ_n is the annual decline in the growth rate. Note that these models predict zero growth asymptotically, though this may occur centuries in the future.⁶

The remaining portion of the model explains the link between economic activity (measured as aggregate output Y_t) and warming (measured as the average surface temperature

⁵Over larger ranges, the damage function becomes S -shaped.

⁶For example, the range of parameters used in the simulations (with $\delta_{a/n} \in (0.25\%, 2.5\%)$) leads to a halving of the growth rates every 20 to 200 years.

T_t). The first step is linking output to emissions:

$$E_t = \sigma_t(1 - \mu_t)Y_t \left(\frac{A_t}{A_t^*}\right)^{1-\theta} \quad (10)$$

where E_t is emission of CO₂, σ_t is an exogenous trend in emissions/output, and μ_t is the rate of emissions reductions induced by the policymaker. The expression $Y_t \left(\frac{A_t}{A_t^*}\right)^{1-\theta}$ reflects raw output prior to the effects of climate damages and control costs. The model of σ_t is, as with labor productivity and population, based on exponentially decaying growth:

$$\log(\sigma_t) = \log(\sigma_{t-1}) + \gamma_\sigma \exp(-\delta_a t) \quad (11)$$

where γ_σ is the initial growth rate of emissions/output (a negative number) and δ_a is the annual decline in the growth rate. Note that the annual decline in the emissions/output growth rate is the same as the annual decline in labor productivity growth (δ_a).

Emissions of CO₂ accumulate in the atmosphere according to:

$$M_t - 590 = \beta E_{t-1} + (1 - \delta_m)(M_{t-1} - 590) \quad (12)$$

where M_t is the atmospheric concentration of CO₂ in billions of tons of carbon equivalent. β is a measure of the retention rate of emissions. Low values of β indicate that emissions do not, in fact, accumulate while a value of unity would mean that every ton of emitted CO₂ becomes a ton of atmospheric CO₂. The parameter δ_m plays the role of a depreciation rate: it is assumed that CO₂ in the atmosphere above the pre-industrialization level of 590 billion tons slowly decays. This decay reflects absorption of CO₂ into the oceans which are assumed to be an infinite sink.

Above average concentrations of CO₂ in the atmosphere lead to increased radiative forcings, a measure of the rate of transfer between solar energy produced by the sun and thermal energy stored in the atmosphere. This is modeled according to

$$F_t = 4.1 \times \log(M_t/590) / \log(2) + O_t \quad (13)$$

where F_t measures radiative forcings in units of watts per meter squared. The specification is such that a doubling of CO₂ concentrations leads to a roughly four fold increase in forcings (since 590 is the concentration before industrialization). O_t in this relation represents radiative forcings due to other greenhouse gases and is assumed exogenous to the model:

$$O_t = \begin{cases} 0.2604 + 0.0125t - 0.000034t^2 & \text{if } t < 150 \\ 1.42 & \text{otherwise} \end{cases} \quad (14)$$

Increased forcings lead to temperature changes according to

$$T_t = T_{t-1} + (1/R_1) [F_t - \lambda T_{t-1} - (R_2/\tau_{12})(T_{t-1} - T_{t-1}^*)] \quad (15)$$

$$T_t^* = T_{t-1}^* + (1/R_2)(R_2/\tau_{12})(T_{t-1} - T_{t-1}^*) \quad (16)$$

where T_t is the surface temperature and T_t^* is the deep ocean temperature, both expressed in changes relative to pre-industrialization levels in degrees Celsius. Note that if $M_t = 590$ and $O_t = 0$ (e.g., pre-industrialization), T_t and T_t^* will equilibrate to zero. The parameter λ describes the equilibrium change in surface temperature for a given change in radiative forcings. In particular, based on (13) and (15), a doubling of the concentration of CO₂ in the atmosphere will lead to a $4.1/\lambda$ rise in surface temperature in the long run. This parameter $4.1/\lambda$ is a measure of the temperature sensitivity of the atmosphere.⁷

The parameters R_1 , R_2 and τ_{12} describe the thermal capacity of the surface atmosphere and deep oceans and the rate of energy transfer between them, respectively.

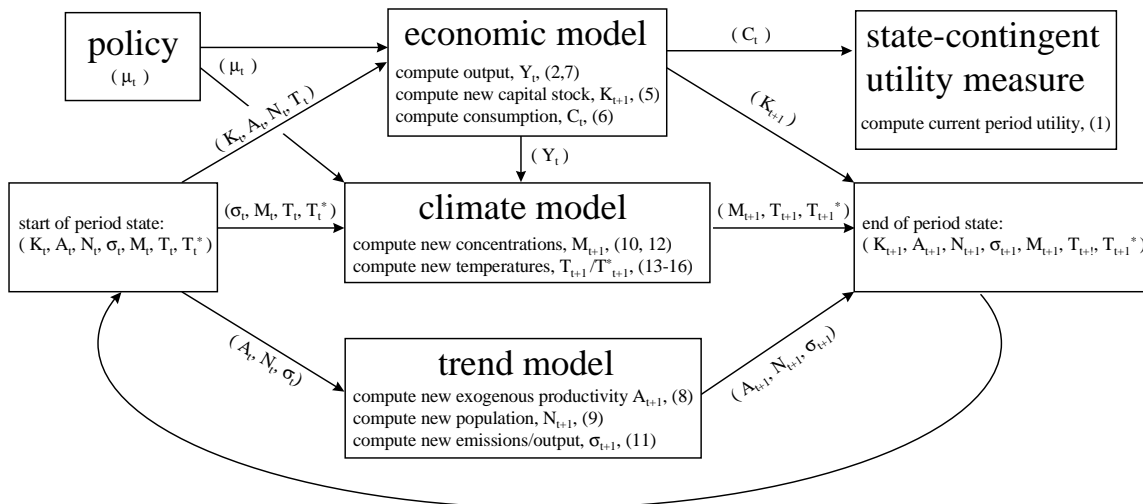
2.2.3 Summary of State-contigent Climate-Economy Model

The state-contigent model is now complete. Figure 1 shows schematically how the model evolves over time and how policy (a path of reductions $\{\mu_t\}$) is translated into a state-contigent utility measure (as measured in Equation (1)).

At any point in time, the vector $\{K_t, A_t, N_t, \sigma_t, M_t, T_t, T_t^*\}$ completely describes the state of the economy and climate. The three somewhat distinct portions of the model – economy,

⁷ λ by itself is referred to as the climate feedback parameter.

Figure 1: State-contingent model



climate and trends – are then used to evolve the current state into the next period state. The economy and climate models make use of the policy variable μ_t describing emission reductions in order to compute control costs and realized emissions, respectively. The realized path of consumption is an output of the economic model which is then used to evaluate the utility associated with implementing the policy in this state.

A key innovation relative to similar models is the use of Equation (5) in the economic model to determine the optimal value of K_{t+1} , depicted in the top center box of Figure 1, without performing a numerical optimization. This reduces by an order of magnitude the time required to evaluate policy while preserving the proper intertemporal prices.

2.3 Social Welfare

A distinguishing feature of this analysis is the use of an econometrically estimated parameter distribution describing uncertainty in the economic model. However, the consumer's objective function given by Equation (1) makes no allowance for uncertainty about the preference parameters ρ and τ which are fixed from his or her perspective. In order to encompass uncertainty about preferences, it is necessary to step back and imagine a social planner who would like to maximize the objective given in (1) but is unsure of the parameters. Since

a policy change which raises the expected utility for one set of parameters may lower the expected utility for another set, the social planner will need to specify a social welfare function to compare gains and losses. This social welfare function provides a single objective specifying how changes in utility measured with different preferences are aggregated.⁸ It is important to recognize that although parameter values in the representative agent model can be inferred from observed consumer behavior, there is no information available to estimate parameters in a social welfare function. Such information would be revealed only by observing the behavior of an actual social planner. Instead, we must rely on social choice theory and our own sense of fairness to specify the relation.

It is useful to note that the common approach in the climate change literature skirts this issue of preference aggregation by reporting a range of policy prescriptions based on a range of possible preferences and states of nature. For example, Cline (1992) presents benefit-cost analyses for 92 different cases (Tables 7.3 and 7.4). Dowlatabadi and Morgan (1993) integrate out much of the uncertainty in their analysis, but still present results for 48 scenarios. Chapter 8 of Nordhaus (1994b) gives one of the few examples where even preference uncertainty is integrated out, yielding a single welfare metric and a single policy recommendation. In a similar analysis, however, Nordhaus and Popp (1997) choose to fix preferences because of the difficulties with preference aggregation. Regardless, these authors ubiquitously observe that uncertainty about time preference has large consequences for optimal policy choice.⁹ Moreso, in fact, than uncertainty about climate sensitivity and damages. It therefore behooves us to seriously consider how to aggregate over uncertain preferences in the most reasonable way.

In this analysis social welfare is specified as an average of utility measured in each state of nature by Equation (1), *rescaled*. The rescaling serves to equate the marginal social welfare of one additional dollar of current consumption in all states of nature. While arbitrary, some adjustment is necessary to prevent the resulting policy prescription from being sensitive to

⁸E.g., providing a negative loss function across states of nature for the social planner.

⁹See discussion in Arrow, Cline, Maler, Munasinghe, and Stiglitz (1996).

the choice of units in the model (e.g., using dollars rather than thousands of dollars as the unit of account; see explanation in Appendix C). Social welfare can then be written as

$$SW(x) = I^{-1} \sum_{i=1}^I u(x, i) \quad (17)$$

where $u(x, i)$ is rescaled utility in state i with outcome x and $SW(x)$ is the social welfare associated with x . The rescaling is such that

$$u(+\$1 \text{ in initial period}, i) - u(\emptyset, i) = u(+\$1 \text{ in initial period}, j) - u(\emptyset, j) \quad \forall i, j$$

That is, a policy corresponding to an extra dollar of consumption in the initial period is assumed to have the same utility gain in every state relative to a no policy (\emptyset) baseline.

This social welfare function has its origins in the literature on social choice. Harsanyi (1977) shows that in defining social welfare over lotteries, if individual preferences satisfy the von Neumann Morgenstern axioms then social welfare must have this weighted average form. Otherwise, social preferences will fail to mimic individual preferences over lotteries involving only that individual. This functional form can also be derived from the assumption of cardinal unit comparability, as discussed by Roberts (1980). More flexible forms require additional assumptions about level or scale comparability. Our choice of welfare functions is therefore less arbitrary than it might have originally appeared: a more flexible form requires both integrating out uncertainty from the representative agent's perspective (to satisfy Harsanyi's point) and more stringent assumptions about the level of comparability (to satisfy Robert's point)¹⁰

2.4 Quantifying Uncertainty

Estimates of uncertainty in the model come from two sources: econometric analysis and subjective assessment. The model involves nineteen different parameters. Six are parameters describing observable economic activity:

¹⁰Additional levels of comparability are especially difficult with the the constant coefficient of relative aversion (CRRA) form in Equation (1) where the parameter $\tau \geq 1$. Under these assumptions, utility is alternatively bounded from above or below.

- pure time preference ρ ,
- risk aversion τ ,
- output-capital elasticity θ ,
- productivity growth γ_a ,
- variation in productivity growth σ_a^2 , and
- depreciation δ_k .

A joint distribution for these parameters is estimated with historical data. The remaining thirteen describe emissions:

- emissions rate growth γ_σ ,

climate change:

- CO₂ retention rate β ,
- temperature sensitivity $4.1/\lambda$,
- CO₂ decay rate δ_m ,
- thermal capacities and conductivities R_1 , R_2 and τ_{12} ,

control costs and damages:

- cost function parameters b_1 and b_2 ,
- fractional loss of GDP for 3° temperature rise D_0 ,

and long-term growth trends

- population growth γ_n ,
- productivity slowdown and slowdown in the growth rate of emissions/output δ_a ,
- population slowdown δ_n .

Uncertainty about these parameters is based on subjective analysis.

The econometric analysis of the six economic parameters is based on post-war U.S. data as described Appendix B. Series describing aggregate investment, capital services, output and prices are fit to the model described by Equations (2),(3) and (5). The posterior parameter distribution which arises from this analysis is summarized in Table 1.

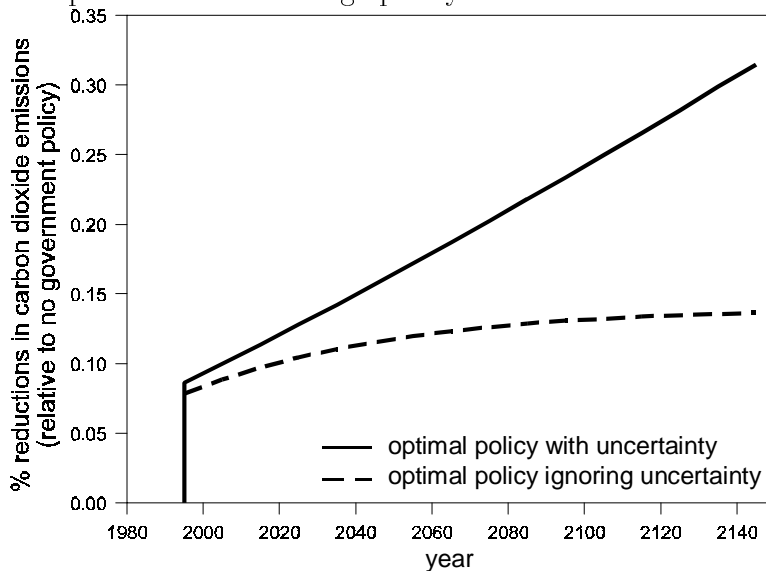
Nordhaus (1994b)¹¹ develops a distribution for the remaining parameters based on a two-step subjective analysis. The first step involves testing his model's sensitivity to each parameter being changed, one at a time, to a more extreme value. Those parameters which produce the largest variance in model output are then further scrutinized. A discrete, five-value distribution is developed for seven of these thirteen variables. The other six are fixed

¹¹Chapters 6 and 7.

Table 1: Marginal distributions of uncertain economic parameters
(narrow bars indicate values used in simulations without uncertainty)

description	symbol	equation	distribution
pure rate of time preference	ρ	(1)	
coefficient of risk aversion	τ	(1)	
output-capital elasticity	θ	(2)	
rate of capital depreciation	δ_k	(3)	
initial productivity growth rate	γ_a	(7)	
standard error of productivity shocks	σ_a	(7)	

Figure 2: Optimal climate change policy with and without uncertainty



at their best guess values. The distribution of the seven uncertain parameters is summarized in Table 2. Values of the six fixed parameters as well as initial conditions for the model are given in Table 3.

3 Does Uncertainty Matter?

To understand whether uncertainty significantly affects decision making and welfare consequences, two optimizations are performed. The first looks for the optimal CO₂ reductions based on the model developed in the previous section and the parameter distributions given in Tables 1 and 2. The second assumes there is no uncertainty. All uncertain parameters are fixed at their median values and the variance of the random exogenous growth shocks is set to zero (as indicated by the narrow bars in Tables 1 and 2). There is only one state of nature.

3.1 Policy Stringency

The resulting policies are shown in Figure 2. The initial difference between policies is small, 7.9% ignoring uncertainty versus 8.6% with uncertainty included. In thirty years these

Table 2: Discrete distributions of uncertain climate/trend parameters
(narrow bars indicate values used in simulations without uncertainty)

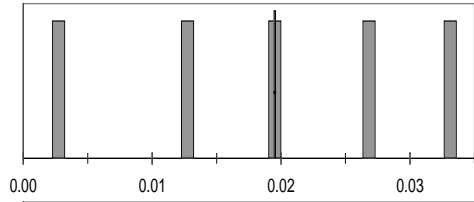
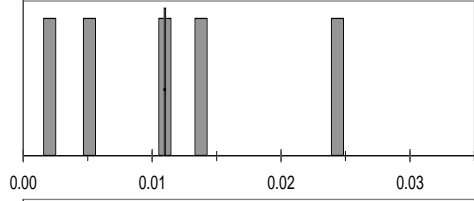
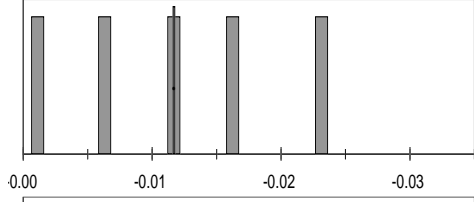
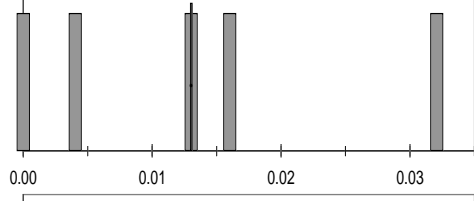
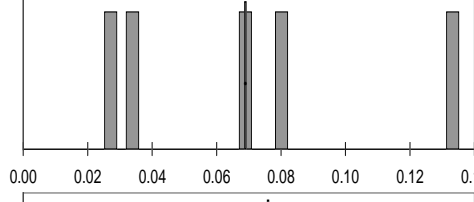
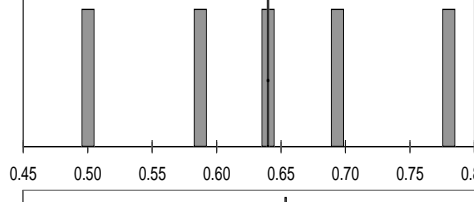
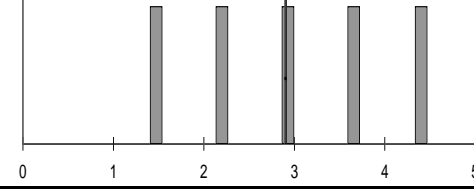
description	symbol	equation	distribution
annual decline of population growth rate	δ_n	(9)	
annual decline of productivity growth rate	δ_a	(7),(11)	
initial growth rate of CO2 per unit output	γ_σ	(10),(11)	
damage parameter (% loss of GDP for 3° temperature rise)	D_0	(8)	
cost function parameter	b_1	(8)	
retention rate for CO2 emissions	β	(12)	
temperature sensitivity to CO2 doubling (in °C)	$4.1/\lambda$	(13),(15)	

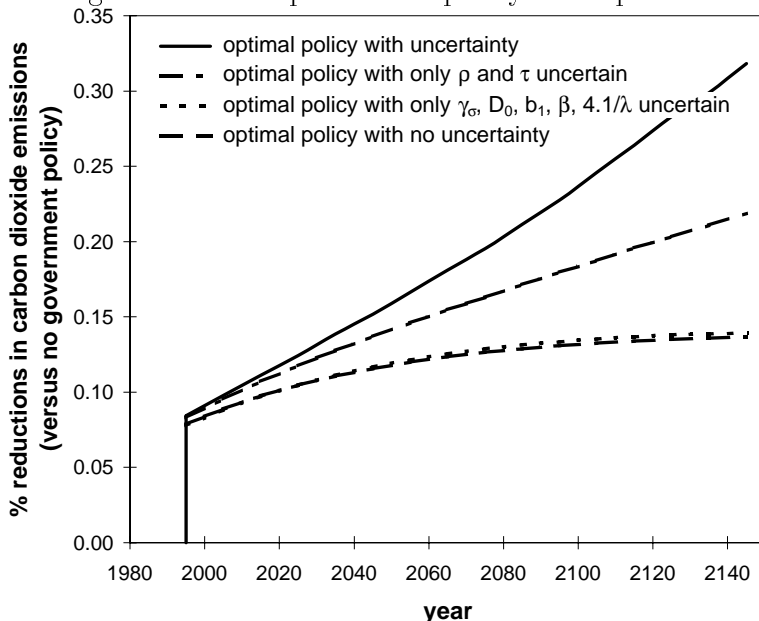
Table 3: Description of fixed parameters^a

parameter	symbol	equation	units	value
cost function parameter	b_2	(8)		2.887
decay rate of atmospheric CO ₂	δ_m	(12)		0.00833
1/thermal capacity of atmosphere	$1/R_1$	(15)	°C-meter ² /watt-year	0.048
thermal conductivity b/w atmosphere and oceans	R_2/τ_{12}	(15),(16)	watt/°C-meter ²	0.44
1/thermal capacity of deep oceans	$1/R_2$	(16)	°C-meter ² /watt-year	$\frac{0.002}{0.44}$
1995 global population	N_0	(9)	millions of people	5590
initial population growth	γ_n	(9)		0.0124
initial rate of CO ₂ emissions per unit of output	σ_0	(11)	billion tons CO ₂ per \$1989 trillions	0.385
1995 global capital stock	K_0	(2),(3)	\$1989 trillions	79.5
1995 global output	Y_0	(2),(3),(10)	\$1989 trillions	24.0
1995 atmospheric concentrations of CO ₂	M_0	(12)	billions of tons of C equivalent	763.6
1995 surface temperature ^b	T_0	(8),(15),(16)	°Celsius	0.763
1995 deep ocean temperature ^b	T_0^*	(15),(16)	°Celsius	0.117

^aAll fixed parameters are from Nordhaus (1994b). The parameters that do not depend on time are from Nordhaus' Table 2.4. Initial values for temperature, CO₂ concentrations, and output in 1995, as well as the initial annual growth rate for population, are based on the Nordhaus base case simulation. The 1995 capital stock is adjusted upward to reflect differences in the definition of capital as well as underlying parameter values. The decay rate of atmospheric CO₂ is divided by ten to convert from a decennial to annual rate. The annual thermal capacity of the ocean and atmosphere are from the second line of Nordhaus' Table 3.4b.

^bTemperatures are measured as deviations from the pre-industrialization level, circa 1900.

Figure 3: Decomposition of policy consequences



differences are higher, 10.4% versus 12.8% in 2025. By the end of the next century, however, the rate is almost doubled when uncertainty is considered. Uncertainty clearly has a dramatic effect on the future path of the optimal control rate, but with a relatively small consequence for immediate decisions.

What explains this pattern of small policy consequences now and large ones later? Surprisingly, roughly half of the observed consequences can be traced to uncertainty about future discounting, rather than climate or other economic uncertainty. This can be seen by reoptimizing the policy for a model with only pure time preference ρ and risk aversion τ uncertain, as illustrated in Figure 3. In contrast, a reoptimized policy based on modeling only the climate parameters (γ_σ , D_0 , b_1 , β , and $4.1/\lambda$) as uncertain yields essentially the same results as completely ignoring uncertainty.

Intuitively, Figure 3 shows the consequences of non-linearities in the first order conditions generating the optimal policy. If these first order conditions were linear in the uncertain parameters, replacing uncertain parameters with central estimates would have little effect. As it turns out, these conditions are non-linear functions of the discount rate determined by the preference parameters ρ and τ but fairly linear in terms of the climate parameters.

This non-linearity arises because the optimal policy each period seeks to balance marginal costs that occur in the current period with marginal benefits that occur in each period over a long horizon. Like an annuity with coupon P and interest rate i whose value is given by P/i , the value is a very non-linear function near $i = 0$. Furthermore, in these simulations the interest rate is relatively well known during the initial periods but less precisely known – and possibly near zero – in the future. Thus the policy effects of uncertainty rise in the future, as shown in Figures 2 and 3.

Why does uncertainty about the interest rate rise in the future? The historic interest rate is an observed quantity and therefore easily predicted over short horizons. More generally, it depends in a simple way on future growth, risk aversion and time preference via the Euler equation: $i_t = \rho + \tau \cdot g_{c,t}$. Here i_t is the interest rate at time t , ρ the pure rate of time preference, τ the coefficient of relative risk aversion and $g_{c,t}$ the rate of consumption at time t . Intuitively, the return to savings in period t must compensate for the fall in marginal utility between periods t and $t + 1$. This decrease occurs because future utility is discounted (ρ) and because as consumption per capita rises, marginal utility falls ($\tau \cdot g_{c,t}$).

Future uncertainty arises because we assume per capita growth $g_{c,t}$ eventually declines to zero (according to Equation (7)) leaving only pure time preference ρ to motivate non-zero interest rates. While historical data tells us that with growth of 1.3%, interest rates will be 5.5%, we have much less information about how to dissect the observed 5.5% into its component pieces ρ and $\tau \cdot 1.3\%$. Therefore, we know less about what the interest rate will be once growth declines.

An interesting interpretation of this result is that the assumption of a productivity slowdown is leading to *more* stringent policy. This runs counter to the idea that a slowdown *reduces* stringency by forecasting less growth, lower emissions and less of a climate problem (Kelly and Kolstad 1996a). Here, a productivity slowdown leads to potentially low discount rates so the present value of the damages which do occur is much higher. Uncertainty in this circumstance reverses the conclusions drawn when it is ignored.

3.2 Welfare

In spite of the differences depicted in Figure 2, the welfare consequences of the two policies remain vague. That is, it is still unclear whether the policy based on best-guess values leaves us substantially worse off compared to the policy which optimizes over uncertainty. To answer this question, the social welfare function developed in (17) can be used to compute the welfare associated with each policy, assuming uncertainty is in fact present. However, the units of social welfare have no intrinsic meaning. A better vehicle for comparison is the certainty equivalent change in first period consumption which yields same welfare consequences as the given policy. For example, implementing the policy which ignores uncertainty improves social welfare by 0.32 units based on Equation (17).¹² A similar improvement in social welfare can be achieved by exogenously raising per capita consumption by \$58 in the first period in every state of nature. This second measure has considerably more intuitive meaning.

What about the policy which optimizes over uncertainty? Here, the certainty equivalent gain is \$73 – a 25% improvement. The source of this improvement tends to be those states with low costs ($b_1 \downarrow$), high damages ($D_0 \uparrow$), high temperature sensitivity ($4.1/\lambda \uparrow$), low pure rates of time preference ($\rho \downarrow$), rapid productivity slowdown ($\delta_a \uparrow$), but high population growth ($\delta_n \downarrow$). In these states, the marginal costs to reduce emissions are low and the marginal benefits are high. Therefore, a more stringent policy improves welfare. While opposing cases (e.g., high costs, low damages) favor less stringent policy, such cases are not as extreme and the losses from a more stringent policy – though still loses – are not as significant. This is the non-linearity discussed earlier. In other words, once uncertainty is considered, bad states of the world tend to be more extreme than good states and motivate additional stringency.

¹²It turns out that computing the welfare gain is complicated by extremely fat tails in the distribution of state-contingent welfare gains. Even with 50,000 draws, the mean welfare gain had not yet converged to its asymptotic normal distribution (a sample of 38 such means rejected normality at the 10^{-5} level). The somewhat ad hoc solution used here censors all gains higher than +150 and lower than -12 utility units. This corresponds to censoring at -\$1,200 and +\$11 million in terms of certainty equivalent gains and affects 0.1% of the states. This approach was chosen so that the reported results, if anything, are understated.

3.3 A Comparison with Previous Results

The consequences of uncertainty on policy stringency have been examined previously by Nordhaus (1994b), Manne (1996) and Nordhaus and Popp (1997). Manne (1996) finds that ignoring uncertainty has negligible policy consequences. However, his experiment contrasts the case of no uncertainty with the case where both high climate sensitivity and high damages occur with only 0.25% probability. Such a low probability event – even with damages ten times higher than the baseline – is unlikely to affect policy decisions in the short term since baseline damages remain small and discounting annihilates effects far in the future. A more appropriate comparison is Nordhaus (1994b) and Nordhaus and Popp (1997) where their model of climate uncertainty (e.g., the DICE model) is identical to the one in this paper. Surprisingly, their experiments indicate significantly larger consequences: they find that ignoring uncertainty lowers the optimal first period reductions by 35%. This contrasts with the roughly 8% difference visible in Figure 2.¹³

Table 4 reveals that this difference centers on the handling of economic parameters. In Column 5, their DICE model is simulated with climate behavior and damages as the sole source of uncertainty – analogous to the dotted line in Figure 3. As in Figure 3, climate uncertainty alone has a negligible policy consequence (compare Columns 3 and 5).

How does the remaining economic behavior in the models differ? First, DICE is based on decennial simulation intervals versus annual intervals used in this paper. Since benefits do not accrue until two periods after a policy takes effect, this leads to lower control rates in DICE.¹⁴ Second, the initial 1995 discount rate in DICE is lower than the value used in our analysis – 4.8% versus 5.5%.¹⁵ This leads to higher control rates as future benefits are

¹³These percentages refer to $1 - 0.088/0.128$ and $1 - 0.079/0.086$, respectively, from the first row of Table 4.

¹⁴Emissions in period t lead to higher concentrations and forcings in period $t + 1$. Higher forcings in period $t + 1$ lead to higher temperatures in period $t + 2$, depressing output in period $t + 2$ via the damage function in Equation (8).

¹⁵Nordhaus (1994b) reports a 1995 interest rate of 5.9% in Table 5.6. However, this is based on dividing the ten-year rate by ten, $1.59/10$, rather than taking the 10th order root, $\sqrt[10]{1.59} = 1.048$ (see definition of RI on page 196).

Table 4: Comparison of Optimal Policy Results with Earlier Work
 (% Reduction in Carbon Dioxide Emissions Versus No Policy Baseline)

	Figure 3		Nordhaus (1994b)		
	no uncertainty	uncertainty	no uncertainty ^a	uncertainty ^b	only climate uncertainty ^c
1995	0.079	0.086	0.088	0.128	0.087
2005	0.089	0.100	0.096	0.157	0.096
2015	0.097	0.114	0.104	0.193	0.104
Column:	1	2	3	4	5

^aTable 5.7, Nordhaus (1994b).

^bTable 8.3, Nordhaus (1994b), last column.

^cThis column was computed by the author and is based on a policy optimization over 625 states of nature using the sampling scheme described in Nordhaus and Popp (1997), page 7, except that only parameters describing climate behavior and damages are uncertain (growth rate of emissions/output γ_σ , damages D_0 , control costs b_1 , CO₂ retention rate β and temperature sensitivity $4.1/\lambda$). The model used to generate these results is otherwise identical to the probabilistic version of DICE except: a) capital accumulation is based on the log-linear rule in Equation (5), and b) learning never occurs so the control rate is the same across all states in each period. As a consistency check, note that this model closely replicates the results in Column 4 when all parameters assume uncertain values: 0.126 in 1995, 0.149 in 2005, and 0.175 in 2015.

discounted less and costs remain the same (since costs occur in the current period). Finally, uncertainty about the discount rate, particularly the potential for low values, is larger in DICE after three periods where the values $\{0.022, 0.032, 0.042, 0.052, 0.062\}$ are each equally probable. Here, the probability of a discount rate below 0.042 is only 10%.¹⁶ This generates larger uncertainty consequences in the DICE model in earlier periods.

While exhibiting differences, these results remain consistent with the previous conclusions. In particular, ignoring uncertainty has substantial consequences. These consequences arise primarily from economic uncertainty surrounding growth and discounting. Mean-preserving uncertainty solely about climate parameters has negligible effects.

¹⁶The range for the DICE model is based on $\{0.01, 0.02, 0.03, 0.04, 0.05\}$ as equally possible values of time preference, a coefficient of relative risk aversion equal to unity, and productivity growth of 1.11% in 2020. The calculation for the data in this paper is based on probability that $\rho + \tau \cdot \gamma_a \cdot (1 - .011)^{30} < 0.04$ where the marginal parameter distributions are shown in Table 1 and $(1 - 0.011)^{30}$ represents the mean productivity slowdown from year one of the simulation.

4 Instrument Choice

In the analysis so far we have sought the path of rate controls $\{\mu_t\}$ which maximizes welfare. In any given state, this control rate translates into a marginal cost of emission reductions (\$/ton of carbon) given by $\frac{\partial Y_t}{\partial \mu_t} \div \frac{\partial E_t}{\partial \mu_t}$:

$$MC_t = \frac{b_1 \cdot b_2}{\sigma_t \cdot (1 + D_0/9 \cdot T_t^2)} \mu_t^{b_2-1} \quad (18)$$

where b_1 and b_2 are cost function parameters, D_0 is the damage from a 3° temperature rise, T_t is the temperature rise since industrialization, μ_t is the control rate, and σ_t is the rate of emissions per unit of output.

To model a tax instrument, we consider a path of taxes $\{TAX_t\}$ and in each state use Equation (18) to back out the control rate. This assumes that in response to a tax, optimizing agents choose a level of reductions that sets marginal cost equal to the tax. If the tax is particularly high, agents in some states may eliminate emissions completely ($\mu_t = 1.0$), leaving $MC_t < TAX_t$.

The important observation is that when uncertainty about parameters in Equation (18) exists, taxes and rate controls cannot be equivalent. Taxes hold marginal cost fixed and generate a distribution of control rates across states. Rate controls hold the fractional reduction in emissions μ_t fixed and generate a distribution of marginal costs. We return to this point later in this section when the value of emission rights is examined.¹⁷

4.1 Taxes versus rate controls

A tax instrument does in fact lead to a higher welfare gain compared to the rate instrument. Performing the optimization just described with $\{TAX_t\}$ as the policy variable the gain is \$86 per person. This is a 20% improvement over the rate instrument, which yields only \$73. Compared to the gain associated with the original policy ignoring uncertainty, this represents a 50% overall gain. In other words, a policymaker who ignores uncertainty, optimizes,

¹⁷Stavins (1989) discusses other important difference between policy instruments.

and implements a rate policy would improve social welfare by \$58 per person (where the measured gain is based on the assumption that uncertainty does, in fact, exist). He or she would not have any reason to consider tax instruments since they are equivalent in the absence of uncertainty. Compare this outcome to that of a second policymaker who considers uncertainty and computes both optimal tax and control rate policies. This policymaker would see that taxes perform better and, in implementing the optimal tax policy, would improve social welfare by \$86 per person. This represents a net improvement of nearly 50% over the first policymaker's outcome.

Why are taxes preferred to rate controls? Weitzman's (1974) proved that this would be true when expected marginal benefits were relatively flat. The left panel of Figure 4 illustrates this intuition: when the cost curve shifts due to random shocks (dashed lines), the optimal shadow price (intersection of costs and expected benefits) is relatively constant compared to the control rate. In terms of the model developed in Section 2, the relative flatness of marginal benefits is a consequence of the choice of damage and cost functions in (8). Over the range of realized values, the damage function $(1 + (D_0/9) \cdot T_t(\{\mu_s\}_{s \leq t})^2)^{-1}$ is more linear than the control cost function $1 - b_1 \mu_t^{b_2}$ where μ_t is the control rate and $T_t(\{\mu_s\}_{s \leq t})$ is the temperature change as a function of past control rates.¹⁸ A more optimistic (e.g., flat) view of marginal control costs and pessimistic (e.g., steep) view of marginal damages would tend to reverse this result.¹⁹

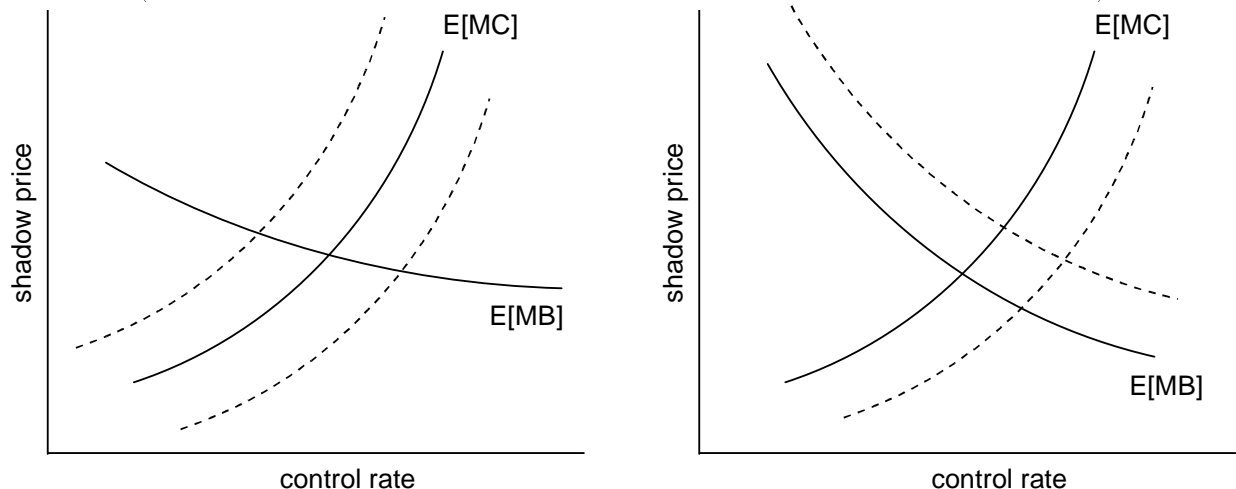
The relative curvature of the cost and damage functions is only part of the reason for preferring taxes, however. Weitzman also noted that when shocks to costs and benefits are correlated, this simple intuition breaks down.²⁰ Here, for example, high marginal benefits occur when productivity growth declines rapidly ($\delta_a \uparrow$), in turn permitting low rates of pure

¹⁸The idea that the marginal cost and damage schedules are flat or steep corresponds to the cost and damage schedules themselves being relatively linear or curved, respectively.

¹⁹Nordhaus (1994b) argues that damages are a linear function of emissions because damages depend on the stock of pollutant and the stock has a virtually linear relation to emissions (page 184). This ignores potential non-linearities in the damage function itself.

²⁰Stavins (1996) shows that under reasonable conditions such correlation can reverse the conclusions drawn from a comparison of relative elasticities.

Figure 4: Intuition for taxes being preferred to rate instruments
 (dashed lines indicate particular realizations of uncertain costs and benefits)



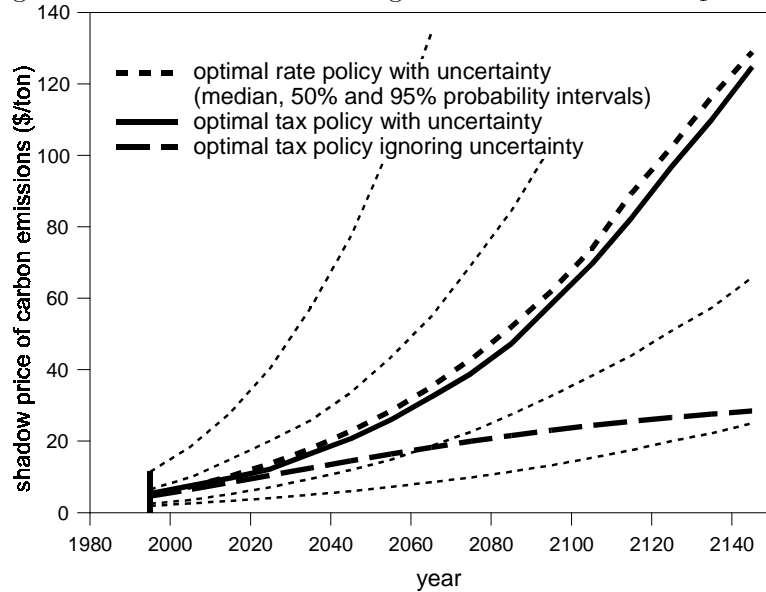
Original Weitzman (1974) result that relatively flatter (higher elasticity) marginal benefits favor taxes. Note that all three intersections correspond to similar shadow prices but different control rates. Hence the deadweight loss will be lower with a fixed shadow price rather than a fixed control rate.

Correlation of upward shifts in benefits with downward shifts in costs favors taxes. Note that both intersections lead to similar shadow prices. As in the left panel, the deadweight loss is lower with the shadow price fixed (via a tax) rather than the control rate fixed.

time preference to generate high expected benefits more quickly (once consumption stops growing, the discount rate falls to the rate of pure time preference). But a high value of δ_a also leads to a higher value of the emissions rate σ_t by (11) and, in turn, low marginal control costs by (18) (recall that growth in σ_t is initially negative so a slowdown in growth leaves it relatively high). Intuitively, a faster slowdown means the emission rates remain high and, at the margin, are cheaper. Hence upward shifts in the benefits curve are correlated with downward shifts in the cost curve. Examining the right panel of Figure 4, we see that this also leads to a preference for taxes since the optimal shadow price is again less variable than the optimal rate.

For policymakers, this result in many ways complements recent work by Goulder, Parry, and Burtraw (1996). They find that the revenue consequences of a tax or permit policy are important. In particular, the social cost of a climate change policy which does not generate revenues – such as a grandfathered permit policy where permits are given away based on past emission levels – are much higher. Therefore, the current view among many policymakers

Figure 5: Value of emission rights under alternative policies



that grandfathered permits are more appealing since they appease those directly bearing the costs is flawed in two ways. First, permits are less efficient as pointed out in this paper. Second, the lost revenue has a significant welfare consequence as pointed out by Goulder, Parry and Burtraw.

4.2 Value of emission rights

Besides focusing on efficiency, another way to view the dichotomy between taxes and rate controls is to examine the value of emission rights (e.g., marginal cost). This represents the permit price under a tradeable permit system or the tax rate under a tax system. Based on a tax system, the value of emission rights is held constant across states at the given tax rate. Under a rate or quantity control, however, emission rights will have different values in different states. Using the results of many simulations, it is possible to plot the distribution of resulting values.

Figure 5 shows the median, 50% and 95% forecast intervals for the value of emission rights under the optimal rate control, along with the fixed values under the optimal tax policy (the optimal tax policy which ignores uncertainty is presented for comparison). While

the median case under the optimal rate control closely tracks the optimal tax policy, the 95% forecast interval includes widely divergent values. For example, the median value of a right to emit one ton of carbon dioxide rises to around \$20 after 50 years under a rate control. It is entirely possible, however, that the value will be over \$100. With these magnitudes of difference between the two instruments in terms of marginal costs, the distinction in welfare consequences discussed in the previous section is not surprising. Intuitively, a relatively flat marginal benefit function would suggest that the fluctuations in marginal cost implied by the rate control are inefficient relative to a tax instrument – exactly the result we find.

5 Conclusion

This paper has sought to inject discussions of climate change policy with a dose of skepticism regarding the limitations of models which ignore uncertainty. Excluding uncertainty tends to reduce expected marginal benefits due to non-linear relations, in turn leading to policy recommendations which are too lax. Much of this effect can be related to agent preferences and their implications for future discount rates: over half the noted consequences can be replicated with preferences as the sole source of uncertainty. This raises the issue of how to aggregate over preferences, a question which is at once both controversial but necessary. Finally, uncertainty about costs introduces a dichotomy between price and quantity controls. In the case of carbon dioxide emissions, taxes perform better. This preference arises because marginal damages are relatively flat and negatively correlated with marginal costs. The welfare gain associated with the optimal tax instrument is \$86 per person whereas the gain associated with the optimal rate instrument is only \$73. An optimal rate policy computed in the absence of uncertainty has a gain of only \$58.

While the results of this analysis should be taken with the same caveats accorded other integrated assessment models (the dependence on functional form assumptions and forecasts of exogenous variables), there are important distinctions to be made. This is the first dynamic general equilibrium model to incorporate a large number of uncertain states of nature. It is

also takes advantage of econometrically quantified uncertainty. Lastly, the model explicitly addresses the issue of aggregation over uncertain preferences.

The most important consequence of this paper may be to encourage a reassessment of the current agenda for climate change research. Considerable effort has been directed towards the importance of learning; perhaps more effort should be directed at the consequences of uncertainty itself – without learning. A great deal of attention has focused on climate uncertainty; this paper suggests that important uncertainty surrounds the economy. In particular, the course of productivity growth has an important bearing on both future emissions and the valuation of future benefits via the interest rate. Finally, much of the policy discussion seems to have shifted towards quantity controls as the preferred instrument; this paper suggests that efficiency arguments may point the other way, toward taxes.

Appendix

A Notes on Solution Algorithm

This section explains the solution to the consumer optimization problem (1) given by Equation (5). In words, (5) is a log-linear approximation of the optimal decision rule based on the deterministic steady-state level of capital per efficiency unit of labor, e.g., $K_t/(A_t^*N_t)$. This approach turns out to work rather well for three reasons: First, the economy starts off relatively close to the steady state defined by the initial growth rates. Second, the decline in deterministic growth rates, described in Equations (7–9), has little effect on the approximation parameters. Third, the shocks away from the steady state – from both exogenous stochastic productivity shocks as well as endogenous climate change effects – are relatively small. While there are many ways of deriving expressions for α_1 and α_2 in terms of the underlying parameters, the simplest is based on a logarithmic approximation to the saddle path around the steady state in $C_t/(A_t^*N_t)$ and $K_t/(A_t^*N_t)$ space.

We begin with the Euler equation arising in a balanced growth equilibrium (where $\gamma_a =$

$\log(A_t^*) - \log(A_{t-1}^*)$ and $\gamma_n = \log(N_t) - \log(N_{t-1})$ are constant $\forall t$:²¹

$$\left(\frac{C_{t+1}/N_{t+1}}{C_t/N_t} \right)^\tau = (1 + \rho)^{-1} (\theta(K_{t+1}/A_{t+1}^* N_{t+1})^{\theta-1} + 1 - \delta_k) \quad (\text{A.1})$$

At the balanced growth equilibrium $\frac{C_{t+1}/N_{t+1}}{C_t/N_t} = \exp(\gamma_a)$ so (A.1) can be solved for steady state capital stock per efficiency unit of labor, $K_{t+1}/(A_{t+1}^* N_{t+1}) = \tilde{K}^\dagger = \left(\frac{(1+\rho)\exp(\gamma_a\tau) - (1-\delta_k)}{\theta} \right)^{\frac{1}{\theta-1}}$.

The capital accumulation equation (3),

$$K_{t+1} = K_t(1 - \delta_k) + (A^* N)^{1-\theta} K_t^\theta - C_t \quad (\text{A.2})$$

then yields steady state consumption per efficiency unit of labor $C_t/(A_t^* N_t) = \tilde{C}^\dagger = \tilde{K}^{\dagger\theta} + (1 - \delta_k - \exp(\gamma_a + \gamma_n))\tilde{K}^\dagger$.

The approximation (5) is made by taking total derivatives of both (A.1–A.2) with respect to $\log(\tilde{K}_{t+1})$, $\log(\tilde{K}_t)$, $\log(\tilde{C}_{t+1})$ and $\log(\tilde{C}_t)$ around the steady-state values $\tilde{K}_t = \tilde{K}_{t+1} = \tilde{K}^\dagger$ and $\tilde{C}_t = \tilde{C}_{t+1} = \tilde{C}^\dagger$. This leads to the expressions

$$\tau(dc_{t+1} - dc_t) = \kappa_{ck} dk_{t+1} \quad (\text{A.3})$$

$$\exp(\gamma_a + \gamma_n)dk_{t+1} = \kappa_{kk} dk_t + \kappa_{kc} dc_t \quad (\text{A.4})$$

where $dc_t = \log(\tilde{C}_t) - \log(\tilde{C}^\dagger)$ is the log deviation of \tilde{C}_t and $dk_t = \log(\tilde{K}_t) - \log(\tilde{K}^\dagger)$ is the log deviation of \tilde{K}_t . κ_{ck} , κ_{kk} and κ_{kc} are constants, the derivatives of the log of the right hand side of (A.1) with respect to $\log(\tilde{K}_{t+1})$ and of the log of the right hand side of (A.2) with respect to $\log(\tilde{K}_t)$ and $\log(\tilde{C}_t)$.²² Subtracting (A.4) from the same expression evaluated one period forward yields

$$\exp(\gamma_a + \gamma_n)(dk_{t+2} - dk_{t+1}) = \kappa_{kk}(dk_{t+1} - dk_t) + \kappa_{kc}(dc_{t+1} - dc_t) \quad (\text{A.5})$$

Equation (A.3) can then be used to substitute for $dc_{t+1} - dc_t$, leaving a second order difference equation in dk_t :

$$\exp(\gamma_a + \gamma_n)(dk_{t+2} - dk_{t+1}) = \kappa_{kk}(dk_{t+1} - dk_t) + \kappa_{kc} (\kappa_{ck}/\tau) dk_{t+1} \quad (\text{A.6})$$

²¹The Euler equation can be derived from Equation (1) by ignoring random productivity shocks and examining a perturbation which decreases consumption in period t , invests the extra output, then consumes the gross return (interest and principal) in period $t + 1$. If the path of consumption is indeed optimal, this should have no effect on utility and leads to the condition given in (A.1).

²²Note that $\kappa_{ck} = (1 + \rho)^{-1}(\theta - 1)\theta\tilde{K}^{\dagger\theta-1} \exp(-\gamma_a\tau)$, $\kappa_{kk} = (1 + \rho)\exp(\gamma_a\tau)$, and $\kappa_{kc} = -\tilde{C}^\dagger/\tilde{K}^\dagger$.

From (5) and the observation that when $\tilde{K}_t = \tilde{K}^\dagger$ the growth of capital stock per capita is the steady state rate of γ_a , so $\Delta k_t = \gamma_a$, we know that $\gamma_a = \alpha_1 + \alpha_2 \log(\tilde{K}^\dagger)$. This leads to

$$\alpha_1 = \gamma_a - \alpha_2 \log(\tilde{K}^\dagger) \quad (\text{A.7})$$

and suggests a guess for solution to (A.6):

$$dk_{t+1} = (\alpha_2 + 1)dk_t \quad (\text{A.8})$$

(since $dk_{t+1} - dk_t = \Delta k_{t+1} - \gamma_a = \alpha_2 dk_t$). Then (A.6) yields the characteristic equation

$$\exp(\gamma_a + \gamma_n) [(\alpha_2 + 1)^2 - (\alpha_2 + 1)] = \kappa_{kk}((\alpha_2 + 1) - 1) + \kappa_{kc}(\kappa_{ck}/\tau)(\alpha_2 + 1) \quad (\text{A.9})$$

which can be used to find α_2 in terms of κ_{kk} , κ_{kc} and κ_{ck} :

$$(\alpha_2 + 1) = \frac{e^{\gamma_a + \gamma_n} + \kappa_{kk} + \kappa_{kc}(\kappa_{ck}/\tau) - \sqrt{(e^{\gamma_a + \gamma_n} + \kappa_{kk} + \kappa_{kc}(\kappa_{ck}/\tau))^2 - 4e^{\gamma_a + \gamma_n} \kappa_{kk}}}{2e^{\gamma_a + \gamma_n}} \quad (\text{A.10})$$

For $\rho = 0.04$, $\tau = 1.2$, $\theta = 0.38$, $\delta_k = 0.05$, $\gamma_a = 0.013$ and $\gamma_n = 0.012$ we have $\tilde{K}^\dagger = 7.8$, $C^\dagger = 1.6$, $\kappa_{ck} = -0.062$, $\kappa_{kk} = 1.06$ and $\kappa_{kc} = -0.20$. The yields $\alpha_2 = -0.084$ and $\alpha_1 = 0.185$.

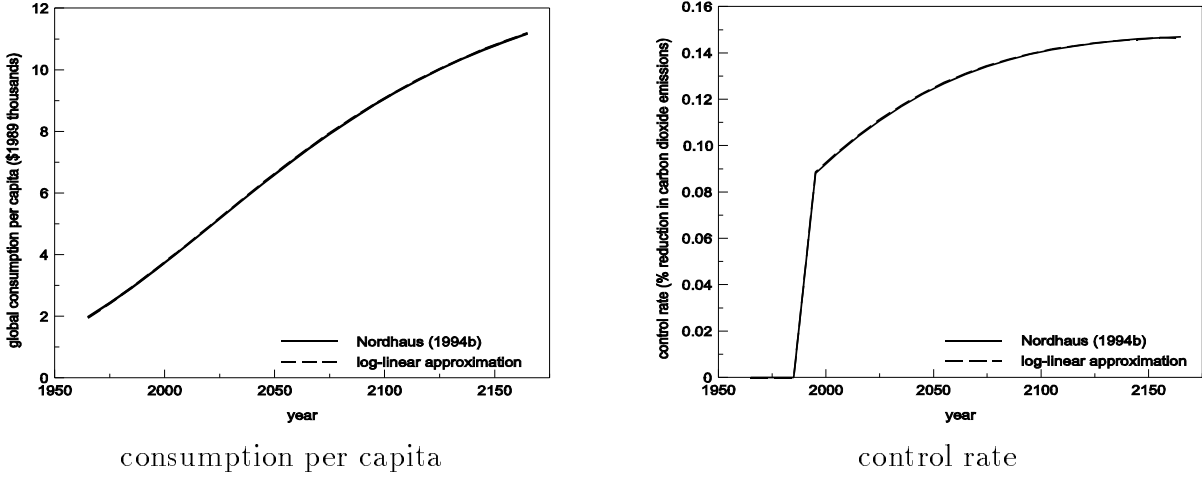
The fact that we can find a value of α_2 which satisfies (A.6) validates our initial guess. Namely, a linear rule for Δk_{t+1} in terms of $k_t - a_t^*$ as in (5) or equivalently (A.8).

To ascertain the accuracy of this approximation, the technique was applied to Nordhaus' (1994b) DICE model and compared to his original results. The results, some of which are shown in Figure 6, indicate a negligible difference.

B Notes on Econometrics

The model given by Equations (2), (3) and (5) can be coupled with the price implications of market equilibrium (Sheperd's Lemma) and a simplified model of productivity shocks

Figure 6: Comparison of Nordhaus (1994b) results with log-linear approximation



(ignoring the future slowdown and climate consequences) to yield the following econometric model:

$$\Delta a_t^* = \gamma_a + \epsilon_t \quad (\text{B.11})$$

$$\Delta k_t = (\alpha_1 - \alpha_2 a_0^*) + \alpha_2(k_{t-1} - (a_{t-1}^* - a_0^*)) + \eta_{2t} \quad (\text{B.12})$$

$$y_t = (1 - \theta)a_0^* + \theta k_t + (1 - \theta)(a_t^* - a_0^*) + \eta_{1t} \quad (\text{B.13})$$

$$\frac{K_{t+1} - (Y_t - C_t)}{K_t} = 1 - \delta_k + \eta_{3t} \quad (\text{B.14})$$

$$\frac{P_{K,t} K_t}{P_{Y,t} Y_t} = \theta + \nu_t \quad (\text{B.15})$$

$$\nu_t = r_v \nu_{t-1} + \eta_{4t} \quad (\text{B.16})$$

Here $a_t^* = \log(A_t^*)$, $k_t = \log(K_t/N_t)$, $y_t = \log(Y_t/N_t)$, $P_{K,t}$ is the price of capital services and $P_{Y,t}$ is the price of output; the remaining variables are as defined in Section 2.²³ The disturbances η_t and ϵ_t are assumed to be jointly NIID with zero covariance (note that the capital share is defined with an autoregressive error). This specification defines a likelihood function describing the distribution of data and unobserved productivity shocks conditional

²³The data, for the years 1952-1992, is from the U.S. Worksheets [Jorgenson (1980); see Ho (1989) and Wilcoxon (1988) for further details]. Output is a divisia index of output of consumption goods (c790+c781), investment goods (c789) and government services (c943). Capital is aggregate capital services (c667). Changes in relative prices among output goods (c792, c761, c791, c395) are subsumed into quantity changes for simplicity. Population is 16+ (c941). Numbers indicate variables from the U.S. Worksheets

on the model parameters.

This likelihood function coupled with a prior distribution over the parameters can be used to obtain a function describing the posterior density of the parameters and productivity shocks conditional on the observed data.²⁴ This is Bayes' rule. Obtaining a sample of draws from the posterior density is difficult, however, as the distribution is non-standard. To generate the distributions shown in Table 1, the Gibbs sampler was used.²⁵ Rather than drawing combinations of all the parameters at once, the Gibbs sampler draws each unobserved parameter sequentially based on its distribution conditional on the previous draw of all other parameters. A chain of draws is formed which converges to the joint posterior distribution. The distribution used in this paper is based on ten chains of 1500 draws with the first 500 draws dropped to reduce start-up effects. The sample was further reduced by about 10% by requiring that the reduced form parameters α_1 and α_2 map back into meaningful structural parameters ρ and τ .

The marginal parameter distributions shown in Table 1 are broadly consistent with previous estimates of the preference parameters ρ and τ (Hansen and Singleton 1983; Hansen and Singleton 1982), productivity growth γ_a (Jorgenson, Gollop, and Fraumeni 1987), depreciation δ_k (Hulten and Wykoff 1981) and capital share θ (National Income and Product Accounts²⁶). Further details about the econometric methodology are given in Pizer (1996).

C Notes on Utility Rescaling

Policy consequences in a given state are always expressible in terms of a first period consumption equivalent. Letting $C_{i,1}(x)$ be this consumption equivalent for state i and policy

²⁴The Jeffrey's prior (see Section 2.8 of Gelman, Carlin, Stern, and Rubin (1995)) is used for each parameter, restricted to the economically relevant parameter space (e.g., $0 < \theta < 1$, $\delta_k > 0$, $\alpha_2 < 0$).

²⁵Geman and Geman (1984), Tanner and Wong (1987), Jacquier, Polson, and Rossi (1994).

²⁶Discrepancies exist because the U.S. Worksheets (versus the NIPA) treat consumer durables, institutional producer durables, institutional producer real estate and owner-occupied housing as capital stock.

x , a social welfare function which simply averages utility yields

$$SW(x) = I^{-1} \sum_i \frac{(C_{i,1}(x)/N_1)^{1-\tau_i}}{1-\tau_i}$$

where N_1 is the first period population (known with certainty) and τ_i is the coefficient of relative risk aversion in state i . But scaling the units of consumption by a factor of ψ changes social the social welfare associated with policy x to:

$$SW(x) = I^{-1} \sum_i \psi^{1-\tau_i} \frac{(C_{i,1}(x)/N_1)^{1-\tau_i}}{1-\tau_i}$$

Unless the consequences are the same for each state, this simple change in units – which leads to an unintentional reweighting among states – will change the ordering among policies. The proposed solution takes the form

$$SW(x) = I^{-1} \sum_i (C_{i,1}(\emptyset)/N_1)^{\tau_i} \frac{(C_{i,1}(x)/N_1)^{1-\tau_i}}{1-\tau_i}$$

where $C_{i,1}(\emptyset)$ is the consumption level in the absense of policy and

$$u(x, i) = (C_{i,1}(\emptyset)/N_1)^{\tau_i} \frac{(C_{i,1}(x)/N_1)^{1-\tau_i}}{1-\tau_i}$$

is the rescaled utility measure in (17). Note that renormalizing the units of consumption no longer affects the policy ordering.

References

- Arrow, K. J., W. R. Cline, K.-G. Maler, M. Munasinghe, and J. E. Stiglitz (1996). Intertemporal equity and discounting. In J. Bruce, H. Lee, and E. Haites (Eds.), *Climate Change 1995: Economic and Social Dimensions of Climate Change*. Cambridge: Cambridge University Press.
- Bertsekas, D. P. (1995). *Dynamic Programming and Optimal Control*, Volume 1. Belmont, MA: Athena Scientific.
- Cline, W. (1992). *The Economics of Global Warming*. Washington: Institute of International Economics.
- Dowlatabadi, H. (1997). Climate change mitigation: Sensitivity to assumptions about convention frameworks and technical change. Seminar/mimeo presented at Resources for the Future, February 5.

- Dowlatabadi, H. and M. Morgan (1993). A model framework for integrated studies of the climate problem. *Energy Policy* 21(3), 209–221.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (1995). *Bayesian Data Analysis*. New York: Chapman & Hall.
- Geman, S. and D. Geman (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721–741.
- Goulder, L., I. W. Parry, and D. Burtraw (1996). Revenue-raising vs. other approaches to environmental protection: The critical significance of pre-existing tax distortions. NBER Working Paper No. 5641.
- Hansen, L. P. and K. J. Singleton (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica* 50(5), 1269–1286.
- Hansen, L. P. and K. J. Singleton (1983). Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy* 91(2), 249–265.
- Harsanyi, J. C. (1977). *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. New York: Cambridge University Press.
- Ho, M. S. (1989). *The effects of external linkages on U.S. economic growth: A dynamic general equilibrium analysis*. Ph. D. thesis, Harvard University.
- Hulten, C. R. and F. C. Wykoff (1981). The measurement of economic depreciation. In C. R. Hulten (Ed.), *Depreciation, Inflation, and the Taxation of Income from Capital*. Washington: Urban Institute Press.
- Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics* 12(4), 371–389.
- Jorgenson, D., F. Gollop, and B. Fraumeni (1987). *Productivity and U.S. Economic Growth*. Cambridge: Harvard University Press.
- Jorgenson, D. W. (1980). Accounting for capital. In G. M. von Furstenberg (Ed.), *Capital, Efficiency, and Growth*, pp. 251–320. Cambridge: Ballinger.
- Kelly, D. L. and C. D. Kolstad (1996a). Malthus and climate change: Betting on a stable population. Working Paper in Economics #9-96R, University of California, Santa Barbara.
- Kelly, D. L. and C. D. Kolstad (1996b). Tracking the climate change footprint: Stochastic learning about climate change. Working Paper in Economics #3-96R, University of California, Santa Barbara.
- Kydland, F. E. and E. C. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica* 50(6), 1345–1370.
- Long, J. B. and C. I. Plosser (1983). Real business cycles. *Journal of Political Economy* 91, 39–69.
- Manne, A. and R. Richels (1995). The greenhouse debate: Economic efficiency, burden sharing and hedging strategies. *The Energy Journal* 16(4), 1–37.

- Manne, A. S. (1996). Hedging strategies for global carbon dioxide abatement: A summary of poll results. Energy Modeling Forum 14 Subgroup – Analysis for Decisions Under Uncertainty.
- Manne, A. S. and R. G. Richels (1992). *Buying Greenhouse Insurance*. Cambridge: MIT Press.
- Nordhaus, W. D. (1993). The cost of slowing climate change. *Energy Journal* 12(1), 37–65.
- Nordhaus, W. D. (1994a). Expert opinion on climate change. *American Scientist* 82(1), 45–51.
- Nordhaus, W. D. (1994b). *Managing the Global Commons*. Cambridge: MIT Press.
- Nordhaus, W. D. and D. Popp (1997). What is the value of scientific knowledge? an application to global warming using the PRICE model. *Energy Journal* 18(1), 1–45.
- Pizer, W. A. (1996). *Modeling Long-Term Policy Under Uncertainty*. Ph. D. thesis, Harvard University.
- Roberts, K. (1980). Interpersonal comparability and social choice theory. *Review of Economic Studies* 47, 421–439.
- Stavins, R. N. (1989). Harnessing market forces to protect the environment. *Environment* 31(1), 5–7, 28–35.
- Stavins, R. N. (1996). Correlated uncertainty and policy instrument choice. *Journal of Environmental Economics and Management* 30(2), 218–232.
- Tanner, M. and W. Wong (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* 82, 528–550.
- Taylor, J. B. and H. Uhlig (1990). Solving nonlinear stochastic growth models: A comparison of alternative solution methods. *Journal of Business and Economic Statistics* 8(1), 1–17.
- Weitzman, M. L. (1974). Prices vs. quantities. *Review of Economic Studies* 41(4), 477–491.
- Wilcoxon, P. J. (1988). *The effects of environmental regulation and energy prices on U.S. economic performance*. Ph. D. thesis, Harvard University.