Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Robust Dynamic Optimal Taxation and Environmental Externalities

Xin Li, Borghan N. Narajabad, and Ted Temzelides

2014-75

Robust Dynamic Optimal Taxation and Environmental Externalities *

Xin Li	Borghan Narajabad	Ted Temzelides
IMF	Federal Reserve Board	Rice University

First Version: 11/26/2012
Current Version: 05/29/2014

Abstract

We study a dynamic stochastic general equilibrium model in which agents are concerned about model uncertainty regarding climate change. An externality from greenhouse gas emissions damages the economy's capital stock. We assume that the mapping from climate change to damages is subject to *uncertainty*, and we use robust control theory techniques to study efficiency and optimal policy. We obtain a sharp analytical solution for the implied environmental externality and characterize dynamic optimal taxation. A small increase in the concern about model uncertainty can cause a significant drop in optimal fossil fuel use. The optimal tax that restores the socially optimal allocation is Pigouvian. Under more general assumptions, we develop a recursive method and solve the model computationally. We find that the introduction of uncertainty matters qualitatively and quantitatively. We study optimal output growth in the presence and in the absence of concerns about uncertainty and find that these concerns can lead to substantially different conclusions.

*We thank Lars Peter Hansen, participants at the 2013 Midwest Macro Conference and at the Econometric Society Summer 2013 meetings for comments and suggestions. The opinions expressed do not necessarily reflect those of the Federal Reserve System.

1 Introduction

We study optimal taxation in a dynamic stochastic general equilibrium model in which agents are concerned about model uncertainty. We assume that an externality through global temperature changes resulting from greenhouse gas emissions *(GHG)* adversely affects the economy's capital stock and, thus, its output. The precise effects of this externality, however, are subject to uncertainty. Most existing approaches, however, only incorporate the uncertainty associated with climate change in a limited way (Stern, 2013). To fill the gap, we focus on the implications of this uncertainty. In order to model the effect of the emissions created by economic activity on the environment, we employ the framework used in Golosov, Hassler, Krusell, and Tsyvinski (2013, GHKT hereafter).[1] While they assume that the mapping from climate change to damages is subject to *risk*, in our model this mapping is subject to *Knightian uncertainty.* We study the implications of this assumption using a robust control approach. We believe that this is an appropriate application of uncertainty in economic modeling. After all, man-made climate change is unprecedented, and there is an ongoing heated debate about its potential effects. Although our model does not include the risks of large-scale human migration or conflict resulting from climate change, it proposes a robust control approach as an alternative to standard probability distribution-based modeling. More specifically, concerned about model uncertainty, a social planner in our model maximizes social welfare under a "worst-case scenario."

In addition to taking model uncertainty into consideration, there are two additional differences between our assumptions and those in GHKT. First, we find it convenient to assume that the environmental externality indirectly affects output through the capital stock. As a result, the theoretical analysis in our model brings different results, although the two assumptions lead to identical results if we assume 100 percent capital depreciation (as we do in the computational part). A second difference is that our estimates of total fossil fuel supplies are significantly larger than theirs. This is partly due to the addition of the supply of unconventional oil and gas, but mainly due to our consideration of estimated methane hydrate resources.[2]

Under additional assumptions, we obtain a sharp analytical solution for the implied pollution externality, and we characterize dynamic optimal taxation. A small increase in the concern about model uncertainty can cause a significant drop in optimal energy extraction. The optimal tax, which restores the social optimal allocation, is Pigouvian. Under more general assumptions, we develop a simple recursive method that allows us to solve the model computationally. We find that the introduction of uncertainty matters in the sense that our model produces results that are qualitatively different - for example, in terms of oil consumption - from those of GHKT. At the same time, concerns about uncertainty do not affect renewable energy adoption. The reason is that, rather than being driven by renewable

[1] Acemoglu, Aghion, Bursztyn, and Hemous (2012) study related issues. See Nordhaus and Boyer (2000) and Stern (2007) for earlier work that also points to the importance of uncertainty.

[2] See Boswell and Collett (2011), Hartley, Medlock, Temzelides, and Zhang (2012), and references therein for a more detailed discussion on total estimated fossil fuel resources.

energy use, the margin that determines short-term decisions regarding energy sources is driven by two factors: the trade-off between higher versus lower total energy consumption, and the choice of coal versus gas/oil. We find that oil use in our model can be flat for some parametrizations. We study optimal output growth in the presence and in the absence of concerns about uncertainty and find that the results can be very different. In the worst-case scenario, optimality implies that a small sacrifice in yearly output can prevent a large future welfare loss.

As the green energy sector does not create emissions in our model, we find that the optimal path for the use of green energy does not directly depend on the level of concern about model uncertainty. However, since green energy, coal, and oil are substitutes, model uncertainty does indirectly affect the use of green energy, through its impact on coal and oil. We also find that an increase in the concern about model uncertainty causes a significant decline in the use of coal, while the use of oil is slightly delayed. Holding other parameters fixed, the optimal path of oil consumption is jointly determined by the resource scarcity effect and the model uncertainty effect. Naturally, we do not find a significant difference in oil consumption when the scarcity effect dominates. However, when we consider a higher level of initial resources of fossil fuel, the concern about model uncertainty substantially discourages the use of oil.

Kolstad (1996) discusses uncertainty in integrated assessment models but does not employ techniques from robust control. Existing work that employs robust control or related techniques in order to address issues related to model uncertainty includes Hennlock (2008, 2009), Funke and Paetz (2010), Sterner and Hennlock (2011), and Lemoine and Traeger (2011). These papers employ a version of Nordhaus's DICE model and we build our analysis closely on GHKT (2013), which is consistent with the DICE model. Using GHKT allows to derive analytical results under a set of additional assumptions. In related recent work, Weitzman (2014) considers the social costs of carbon when catastrophic climate-related events follow a fat-tailed distribution.[3] In addition to building on GHKT, our paper relies on existing work in robust control theory from both economics and engineering. In the traditional stochastic control literature, uncertainties in the system are modeled using probability distributions. The goal there is to derive a policy that works best "on average." In contrast, given a bound on uncertainty, robust control is concerned with optimizing performance under a so-called worst-case scenario.[4] Hansen and Sargent (2001) introduce techniques from robust control theory to dynamic economic decision making problems.[5] They point out the connection between the max-min expected utility theory of Gilboa and Schmeidler (1989) and the applications of robust control theory proposed by Anderson et al. (2000) and Dupuis et al. (1998). Hansen, Sargent, Turmuhambetova and Williams (2006) give a thorough in-

[3]See also Barrage (2013). Other related work includes Hotelling (1931), Dasgupta and Heal (1974), Nordhaus (2000, 2008), Hoel (1978), Stern (2007), Sinn (2008), Gars, Golosov, and Tsyvinski (2009), Krusell and Smith (2009), and Ploeg and Withagen (2012, 2012). GHKT (2013) provide an excellent review of this literature.

[4]See, for example, Lewis (1986) and Chandrasekharan (1996).

[5]See Knight (1921), Savage (1954), Ellsberg (1961), Gilboa and Schmeidler (1989), Hansen and Sargent (2001 and 2010) for related research.

troduction to the robust control approach. They discuss applications to a wide range of problems within the linear-quadratic-Gaussian framework.[6]

As is standard in the robust control literature, our paper postulates the problem of optimal fossil fuel extraction as a two-person zero-sum dynamic game: in each stage, a social planner (a representative household in the decentralized version) maximizes social welfare (lifetime utility) by choosing the level of energy extraction, consumption, labor and capital investment. Subsequently, a malevolent player chooses alternative distributions in order to minimize the respective payoff. Our work contributes to the existing literature of applications of robust control in economics in two ways. First, it explores a class of models under a non-quadratic objective and non-linear constraints. In that regard, we demonstrate that models of the type used in GHKT (2013) can be restated in a robust control framework. We then derive some sharp analytical results, and compute the resulting model numerically. Second, we employ the exponential distribution as the approximating distribution. While existing studies usually employ the linear-quadratic model combined with Gaussian distributions in order to produce analytical solutions, our work shows that the approximating distribution for models with log-utility and full depreciation of capital can be drawn from either the normal or the exponential family.

The paper proceeds as follows. Section 2 presents the basic model. Section 3 studies the model analytically, while Section 4 presents our numerical and quantitative findings. A brief conclusion follows. Technical material appears in the appendices.

2 The Model

In order to characterize the optimal policy for the case where there is a concern about climate change and model uncertainty, we first formulate a general framework for the "robust planner's problem," a benchmark that we will subsequently compare to decentralized market solutions.

Time, t, is discrete and the horizon is infinite. The world economy is populated by a $[0,1]$-continuum of infinite-lived representative agents with utility

$$E_0 \sum_{t=0}^{\infty} \beta^t u(C_t). \tag{1}$$

The function u is a standard concave period utility function, C_t represents final-good consumption in period t, and $\beta \in (0,1)$ is the discount factor. The final goods sector uses energy, E, capital, K, and labor, N, to produce output. Labor supply is inelastic. The economy's capital stock depreciates at rate $\delta \in (0,1)$. Henceforth, $\tilde{K}$ represents the end-of-period capital (before interacting with the climate factor through the process described below). The feasibility constraint in the final goods sector is given by

[6] Related work includes Hansen, Sargent and Tallarini (1999), Hansen and Sargent (2003), Colgey, Colacito, and Hansen and Sargent (2008). See Williams (2008) for a review. In a recent paper, Bidder and Smith (2012) use robust control theory to study the implications of model uncertainty for business cycles generated through "animal spirits."

$$C_t + \tilde{K}_{t+1} = Y_t + (1-\delta)K_t. \tag{2}$$

There are four production sectors. The final-goods sector, indexed by $i = 0$, produces the consumption good. The corresponding production function is given by $Y = F(K, N_0, E)$. Thus, in addition to capital and labor, production of the final good requires the use of energy, E. The three energy-producing sectors for oil, coal, and green energy (labelled by $i = 1, 2, 3$, respectively) produce energy amounts E_1, E_2 and E_3 (measured in carbon equivalents). The oil sector is assumed to produce oil at zero cost. We denote by R the total oil energy stock, and we impose the resource constraint, $R_t \geq 0$, for all t. Both the coal and the green energy sectors use linear technologies

$$E_i = A_i N_i,\ i = 2, 3. \tag{3}$$

We follow GHKT in modeling a simplified carbon cycle as follows. The variable S (measured in units of carbon content) represents the GHG concentration in the atmosphere in excess of the pre-industrial level. We denote by P and T the permanent and temporary components of S, respectively. These evolve according to the following equations.

$$\begin{aligned} P' &= P + \phi_L(E_1 + E_2), & (4)\\ T' &= (1-\phi)T + (1-\phi_L)\phi_0(E_1 + E_2), & (5)\\ S' &= P' + T'. & (6) \end{aligned}$$

We introduce model uncertainty regarding climate change through a stochastic variable, γ, which reduces the end-of-period capital stock $\tilde{K}'$ by a factor of $h(S', \gamma)$ to K'. That is, $K' = h(S', \gamma)\tilde{K}'$. While γ directly affects output in GHKT, we find it convenient to assume that γ adversely affects the economy's capital stock. The two assumptions are identical under a Cobb-Douglas production function and an exponential damage function (which we assume throughout this paper).[7] We use $\pi(\gamma)$ to denote the approximating distribution of γ, while $\hat{\pi}(\gamma)$ denotes the welfare-minimizing distribution and $m(\gamma) = \frac{\hat{\pi}(\gamma)}{\pi(\gamma)}$ is the likelihood ratio. The distance, ρ, between $\hat{\pi}(\gamma)$ and $\pi(\gamma)$ is measured by relative entropy:

$$\rho(\hat{\pi}(\gamma), \pi(\gamma)) \equiv E[m(\gamma)\log m(\gamma)] \equiv \hat{E}[\log m(\gamma)] \equiv \int [m(\gamma)\log m(\gamma)]\pi(\gamma)\mathrm{d}\gamma. \tag{7}$$

As is standard in robust control, the concern about model uncertainty is represented by a two-person zero-sum dynamic game in which, after observing the choice of a social planner, a malevolent player chooses the worst specification of the model in each period. This game

[7] Our specification allows us to assume that, in each period, the social planner moves before nature. The resulting max-min game is easier to analyze. To see the equivalence with GHKT, assume that the economy enters the current period with capital k and carbon concentration S. In GHKT, the final good production is given by $A_0 e^{-\gamma S} K^\theta N_0^{1-\theta-\nu} E^\nu$, while in our model the final good production is given by $A_0(e^{-\gamma S}K)^\theta N_0^{1-\theta-\nu} E^\nu = A_0 e^{-\theta\gamma S} K^\theta N_0^{1-\theta-\nu} E^\nu$. The two production technologies are identical if the damage parameter, γ, in our model is scaled up by a factor of $\frac{1}{\theta}$.

proceeds as follows. Our attention will be restricted to a particular type of equilibrium, the so-called Markov perfect (or *feedback*) equilibrium. This equilibrium is *strongly time-consistent*. At the beginning of a period, the state value of (K, N, P, T, R) is revealed. Then, the planner chooses $(C, E_i, N_i, \tilde{K}', P', T', S', R')$ in order to maximize social welfare. After observing the planner's choice, nature (the "malevolent player") chooses an alternative distribution $\hat{\pi}(\gamma)$ or, equivalently, $m(\gamma)$, to minimize welfare. Note that any deviation from the approximating distribution will be penalized by adding $\alpha\rho(\hat{\pi}(\gamma), \pi(\gamma))$ to the objective function. Here, α represents the magnitude of the "punishment." A greater α means a greater penalty associated with the deviation of γ from its approximating distribution and thus a lower concern about robustness.

This leads to the following social planner's problem:

$$
\begin{aligned}
& V(K, N, P, T, R) = \max_{\{C, E_i, N_i, \tilde{K}', P', T', S', R'\}} \min_{m(\gamma)} \\
& \left\{ u(C) + \beta \int \left[m(\gamma) V(K', N', P', T', R') + \alpha m(\gamma) \log m(\gamma) \right] \pi(\gamma) \mathrm{d}\gamma \right\} \\
s.t. \quad & \\
& E_i = A_i N_i; \; i = 2, 3 \\
& E = (\kappa_1 E_1^\rho + \kappa_2 E_2^\rho + \kappa_3 E_3^\rho)^{1/\rho} \\
& N = N_0 + N_2 + N_3 \\
& \tilde{K}' = F(K, N_0, E) + (1 - \delta) K - C \\
& K' = h(S', \gamma) \tilde{K}' \\
& R' = R - E_1 \geq 0 \\
& N' = A_N N \\
& P' = P + \phi_L (E_1 + E_2) \\
& T' = (1 - \phi) T + (1 - \phi_L) \phi_0 (E_1 + E_2) \\
& S' = P' + T' \\
& 1 = \int m(\gamma) \pi(\gamma) \mathrm{d}\gamma
\end{aligned}
$$

The social planner's problem can be solved analytically under a set of additional assumptions, and we will focus on the analytical solution first. We will discuss the decentralized problem and show that the socially optimal allocation can be restored by imposing appropriate fossil fuel taxes on the energy-producing sector.

3 The Analytical Solution

In this section, we will make the following additional assumptions. While these assumptions are admittedly strong, they allow us to fully solve the model analytically. As we shall see, certain aspects of the solution remain instructive in the next section, when the restrictive assumptions are dropped and the model is solved numerically.

(A1) The period utility function is given by $u(C) = \log(C)$.

(A2) Capital depreciates fully; i.e., $\delta = 1$.

(A3) The production function is given by $F(K, N_0, E) = A_0 K^{\theta} N_0^{1-\theta-\nu} E^{\nu}$.

(A4) The damage function is given by $h(S', \gamma) = e^{-S'\gamma}$.[8]

(A5) The approximating distribution for γ is exponential with mean λ^{-1} and variance λ^{-2}; i.e., $\pi(\gamma) = \lambda e^{-\lambda\gamma}$.[9]

(A6.1) $\phi_L = 0$.[10]

(A6.2) $\phi = 0$.

(A7) There is a single fossil energy sector producing oil at zero cost. Production is subject to a resource feasibility constraint: $R' \geq 0$. As a result, $N_1 = 0$ and $N_0 = N$.

(A8) There is no population growth, and the aggregate labor supply is normalized to 1. That is, $A_N = 1$ and $N = 1$ in all periods.

(A9) There is no technology improvement. That is, A_0 is constant over time. We normalize $A_0 = 1$.

(A10) The resource feasibility constraint is not binding.

We will first solve the social planner's problem. We will then discuss the decentralized problem and show that the socially optimal allocation can be restored by implementing fossil fuel taxes on the energy-producing sector.

Under **A1-A10**, the social planner's problem can be rewritten as:

$$
\begin{aligned}
V(K,S) \quad &= \quad \max_{\{C,E,\tilde{K}',S'\}} \min_{m(\gamma)} \{u(C) + \beta \int \left[m(\gamma)V(K',S') + \alpha m(\gamma)\log m(\gamma)\right] \pi(\gamma)\mathrm{d}\gamma\} \quad (8) \\
&s.t. \\
&\quad \tilde{K}' = F(K,E) - C \\
&\quad K' = h(S',\gamma)\tilde{K}' \\
&\quad S' = S + \phi_0 E \\
&\quad 1 = \int m(\gamma)\pi(\gamma)\mathrm{d}\gamma \quad (9)
\end{aligned}
$$

where $h(S', \gamma) = e^{-S'\gamma}$ and $F(K, E) = K^{\theta}E^{\nu}$. To solve this problem, we first guess that $V(\cdot)$

[8] There exists a constant, Δ, such that if the GHG concentration, S, is greater than $\frac{1}{\Delta}$, the system cannot be "robustified," in the sense that the value of the game goes to negative infinity. However, if the economy starts with an initial $S_0 < \frac{1}{\Delta}$, then S_t will converge to $\frac{1}{\Delta}$ as $t \to +\infty$.

[9] The exponential distribution with mean λ^{-1} is the maximum-entropy distribution among all continuous distributions supported in $[0, \infty]$ that have mean λ^{-1}. The worst-case distribution for γ is also exponential with mean $(\lambda^*)^{-1}$ and variance $(\lambda^*)^{-2}$, where $\lambda^* = \lambda(1 - \Delta S'^*) = \lambda(1 - \Delta\phi_0 c_E)(1 - \Delta S)$. That is, $\pi^*(\gamma) = \lambda^* e^{-\lambda^*\gamma}$. Since $\lambda^* = \lambda(1 - \Delta S'^*) < \lambda$, the worst-case mean of γ, $(\lambda^*)^{-1}$, is strictly greater than the approximating mean, λ^{-1}.

[10] If $\phi_L > 0$, we need to depict the dynamics of P and T separately before we sum them in order to obtain the dynamics of S. Assuming that $\phi_L = 0$ allows us to express the dynamics of S without the need to consider P and T separately. That is, $S' = (1 - \phi)S + \phi_0 E$. Moreover, **(A6.1)** and **(A6.2)** imply that $S' = S + \phi_0 E$, which is necessary for an analytical solution.

takes the form

$$V(K', S') = f(S') + \bar{A}\log(K') + \bar{D} = f(S') + \bar{A}\log(h(S', \gamma)\tilde{K}') + \bar{D} \tag{10}$$

where $\bar{A}$ and $\bar{D}$ are undetermined coefficients. The functional form for $f(\cdot)$ will be derived when we solve the minimizing player's problem.

First, we define the robustness problem (the *inner minimization problem*) by

$$\begin{aligned}
\mathcal{R}(V)(\tilde{K}', S') &= \min_{m(\gamma)} \int [m(\gamma)V(K', S') + \alpha m(\gamma)\log m(\gamma)]\,\pi(\gamma)\mathrm{d}\gamma \\
s.t. & \\
& K' = e^{-S'\gamma}\tilde{K}' \\
& 1 = \int m(\gamma)\pi(\gamma)\mathrm{d}\gamma
\end{aligned}$$

The first-order condition for $m(\gamma)$ implies that

$$m^*(\gamma) = \frac{\exp(-\frac{V(K',S')}{\alpha})}{\int \exp(-\frac{V(K',S')}{\alpha})\pi(\gamma)\mathrm{d}\gamma} = (1 - \Delta S')e^{\Delta S'\lambda\gamma}$$

or, equivalently,

$$\hat{\pi}^*(\gamma) = m^*(\gamma)\pi(\gamma) = \lambda^* e^{-\lambda^*\gamma},$$

where we define $\Delta = \frac{\bar{A}}{\alpha\lambda}$ and $\lambda^* = \lambda(1 - \Delta S')$. The worst-case distribution of γ remains exponential with a distorted mean $(\lambda^*)^{-1}$ and variance $(\lambda^*)^{-2}$. Therefore,

$$\begin{aligned}
\mathcal{R}(V)(\tilde{K}', S') &= \int [m^*(\gamma)V(K', S') + \alpha m^*(\gamma)\log m^*(\gamma)]\,\pi(\gamma)\mathrm{d}\gamma \\
&= -\alpha\log[\int \exp(-\frac{V(K', S')}{\alpha})\pi(\gamma)\mathrm{d}\gamma]. \qquad (11)
\end{aligned}$$

Substituting equation(10) into equation(11), we obtain

$$\mathcal{R}(V)(\tilde{K}', S') = f(S') + \bar{A}\log(\tilde{K}') + \bar{D} + H(S'; \alpha, \bar{A}),$$

where $H(S'; \alpha, \bar{A})$, the robust version of the externality from carbon emissions, is given by

$$H(S'; \alpha, \bar{A}) = -\alpha\log[\int h^{-\frac{\bar{A}}{\alpha}}(S', \gamma)\pi(\gamma)\mathrm{d}\gamma]$$

It follows from **(A4)-(A5)** that

$$H(S'; \alpha, \bar{A}) = \alpha\log(1 - \Delta S').$$

Next, we define the optimal choice problem (the *outer maximization problem*). Using the analysis above, this problem can be written as

$$V(K,S) = \max_{\{C,E,\tilde{K}',S'\}} \{\log(C) + \beta \mathcal{R}(V)(\tilde{K}', S')\}$$

or equivalently,

$$\begin{aligned} f(S) + \bar{A}\log(K) + \bar{D} &= \max_{C,E}\{\log(C) + \beta[f(S') + \bar{A}\log(\tilde{K}') + \bar{D} + H(S';\alpha,\bar{A})]\} \\ s.t. & \\ & \tilde{K}' = F(K,E) - C \\ & S' = S + \phi_0 E \\ & H(S';\alpha,\bar{A}) = \alpha\log(1-\Delta S'). \end{aligned}$$

The first-order conditions imply

$$C = \frac{F(K,E)}{1+\beta\bar{A}} \tag{12}$$

$$-\phi_0\left[\frac{\partial f(S')}{\partial S'} + \frac{\partial H(S';\alpha,\bar{A})}{\partial S'}\right] = \frac{1+\beta\bar{A}}{\beta}\frac{\frac{\partial F(K,E)}{\partial E}}{F(K,E)}. \tag{13}$$

Noting that $H(S;\alpha,\bar{A})$ is a logarithmic function of S, we guess that $f(S) = \bar{B}\log(1-\Delta S)$, where $\bar{B}$ is an undetermined coefficient. As a result, the above F.O.N.C. can be simplified to

$$\begin{aligned} C &= \frac{K^\theta E^\nu}{1+\beta\bar{A}} \\ \frac{\beta\phi_0\Delta(\alpha+\bar{B})}{1-\Delta S'} &= \frac{\nu(\beta\bar{A}+1)}{E} \end{aligned}$$

After some derivations, we obtain

$$\begin{aligned} \bar{A} &= \frac{\theta}{1-\beta\theta} \\ \bar{B} &= \frac{1}{1-\beta}[\alpha\beta + \frac{\nu}{1-\beta\theta}] \end{aligned}$$

The expression for $\bar{D}$ is more complicated and less intuitive. Substituting $\bar{A} = \frac{\theta}{1-\beta\theta}$ into the first-order conditions, we obtain the optimal allocation. We summarize the above discussion in the following.

Proposition 1. *Assume that (A1)-(A10) hold. The two-person zero-sum dynamic game described by equation (8)-equation (9) admits a feedback (Markov perfect) equilibrium. The* *equilibrium strategies are given by:*

$$\begin{aligned} C^* &= (1-\beta\theta)K^\theta E^{*\nu} = (1-\beta\theta)K^\theta[c_E(1-\Delta S)]^\nu \\ E^* &= c_E(1-\Delta S) \\ S'^* &= S + \phi_0 c_E(1-\Delta S) \\ \hat{\pi}^*(\gamma) &= \lambda^* e^{-\lambda^*\gamma} \end{aligned}$$

where $c_E = \frac{\nu(1-\beta)}{[\beta\alpha(1-\beta\theta)+\nu]\phi_0\Delta}$ *and* $\lambda^* = \lambda(1-\Delta S'^*)$.

A few technical remarks are in order. First, the function $V(K,S)$ is increasing in K, decreasing in S, and jointly concave in K and S. The value of $\bar{A}$ is the same as in the model without concern about model uncertainty. Both E^* and S'^* are affine functions of S. In addition, it can be shown that, given S, both E^* and S'^* are increasing functions of α. This is intuitive since a greater α implies a larger resulting penalty from a deviation of γ from its approximating distribution, and thus a lower concern about model-uncertainty. Note that C^* is affected by S only through E^*. This is due to logarithmic utility. As a result, a greater concern about model-uncertainty will lower both E^* and C^*. The value of the externality from one unit of emissions evaluated at E^* is given by

$$\lambda^s = -\beta\frac{\partial V(K',S')}{\partial E}|_{K'^*,S'^*} = \frac{\beta\phi_0\Delta(\bar{B}+\alpha)}{1-\Delta S'^*} = \frac{\nu}{c_E(1-\beta\theta)(1-\Delta S)} = \frac{\nu}{(1-\beta\theta)E^*}$$

Our model so far is similar to the oil regime in GHKT, except that we assume that the resource constraint is not binding. Since $S_{t+1} = S_t + \phi_0 E_t$, we arrive at the following expression for the aggregate oil extraction

$$\sum_{t=0}^{+\infty} E_t = \lim_{t\to+\infty}\phi_0^{-1}(S_t - S_0) = \phi_0^{-1}(\frac{1}{\Delta} - S_0)$$

Thus, the resource constraint is not binding if and only if the aggregate oil reserves are greater than $\phi_0^{-1}(\frac{1}{\Delta} - S_0)$. Figures 1, 2, and 3 below illustrate how E^* responds to a concern about model uncertainty. Figures 1 and 2 show how E^* reacts to a change in the penalty parameter, α.

Another natural measurement for model uncertainty is the distance between $\hat{\pi}^*(\gamma)$ and $\pi(\gamma)$, δ, given by the relative entropy

$$\delta \equiv \rho(\hat{\pi}^*(\gamma),\pi(\gamma)) = \log(1-\Delta S'^*) + \frac{\Delta S'^*}{1-\Delta S'^*},$$

where $\rho(\hat{\pi}^*(\gamma),\pi(\gamma))$ can be viewed as the maximum deviation allowed from the approximating model, $\pi(\gamma)$, given any penalty parameter, α. It is straightforward to verify that $\rho(\hat{\pi}^*(\gamma),\pi(\gamma))$ is decreasing as α increases. Figure 3 shows how E^* changes as we relax δ, allowing for more uncertainty about the approximating model. In the appendix we show that $\frac{\partial E^*}{\partial\delta}|_{\delta=0} = -\infty$. That is, even an infinitesimal concern about model uncertainty can cause a significant drop in the optimal energy extraction.

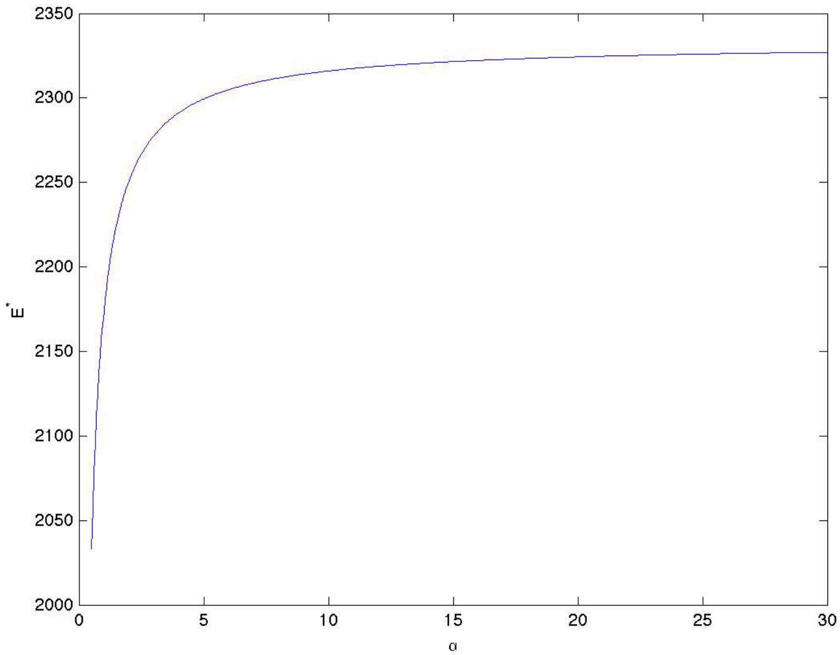

Figure 1: The Effect of Penalty Parameter α on Optimal Carbon Emissions, E

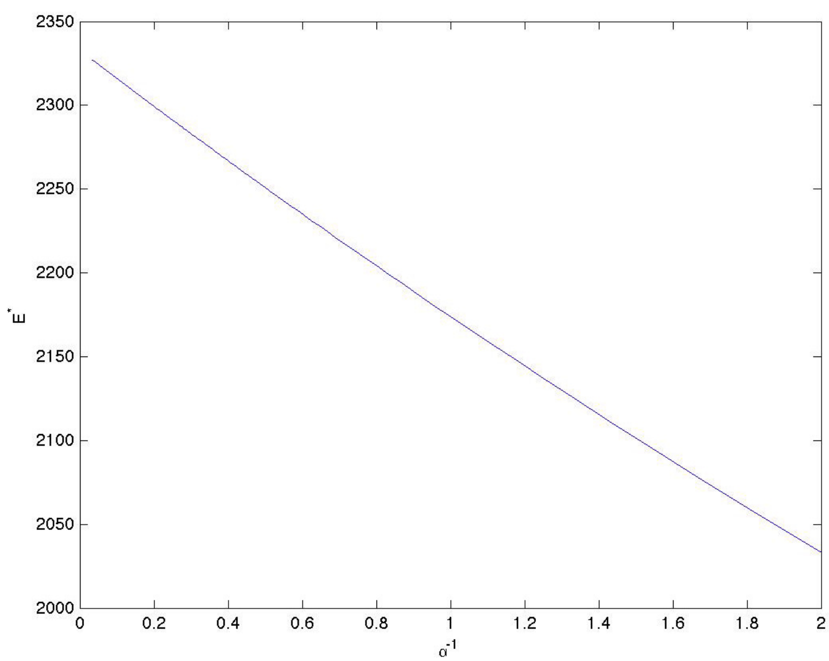

Figure 2: The Effect of α^{-1} on E

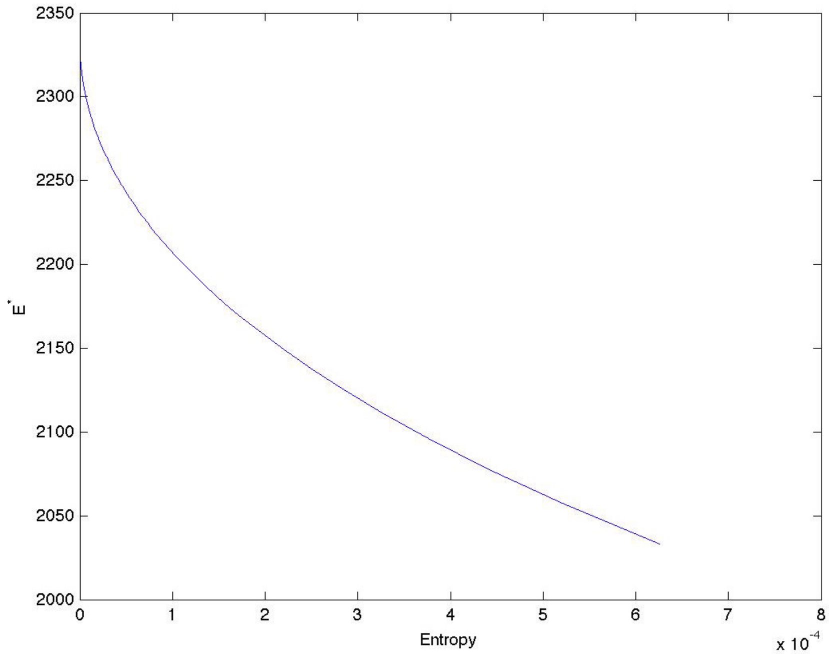

Figure 3: The Effect of Model Deviation as Measured by Entropy, δ, on E

Robust control modeling can be introduced in a variety of ways. So far, we have used a closed-loop zero-sum dynamic game in which the social planner moves first in each period. Alternatively, we can construct a game with the same information structure by interchanging the order of max and min in equation (8). The two games differ only in terms of the timing protocol. However, both lead to the same (unique) feedback saddle-point equilibrium if certain conditions are satisfied. More precisely, if (A1)-(A10) hold, then the objective in (8) is strictly concave in C and E, and strictly convex in $m(\gamma)$. Consequently, the two closed-loop zero-sum dynamic games admit the same unique pure strategy saddle-point Nash equilibrium, which is the one described in Proposition 1.

Let us now turn to the decentralized problem. Suppose a percentage tax, τ_t, is imposed on emissions, E_t. Since the extraction cost of energy (the cost of creating emissions) is zero, it must be true that

$$\tau_t = p_t = \frac{\partial F(K_t, E_t)}{\partial E_t} = \nu K_t^{\theta} E_t^{\nu-1}$$

The above equation captures the one-to-one relationship between E_t and τ_t. Therefore, to achieve the optimal emissions level, $E_t = c_E(1 - \Delta S)$ in equation (14), we must impose $\tau_t = \nu c_E^{\nu-1}(1 - \Delta S_t)^{\nu-1} K_t^{\theta}$. It is straightforward to show that $\tau_t = \frac{\lambda^s}{u'(C_t^*)}$, where C_t^* is the optimal consumption, given by equation (14). That is, the optimal tax on emissions is equal to the corresponding GHG externality measured in units of the consumption good. It remains to show that C_t^* can be recovered under the optimal tax. This can be shown using the representative household's problem as follows. Since we have established a one-to-one relationship between E_t and τ_t, we may assume without loss of generality that the planner chooses E_t. Further, assume that E_t is chosen as a function of S_t only. This is without loss of generality, since our goal is to recover the optimal emissions in equation (14), which depends only on S_t. Given $E = E(S)$, k, K, and S, a representative household solves:

$$\begin{aligned} V(k,K,S) &= \max_{c,\tilde{k}'} \min_{\hat{\pi}(\gamma)} \left\{ u(c) + \beta \hat{E}_\gamma \left[V(k',K',S') + \alpha \log\left(\frac{\hat{\pi}(\gamma)}{\pi(\gamma)}\right)\right]\right\} \\ & s.t. \\ c + \tilde{k}' &= r(K,S)k + \tau(K,S)E(S) + \pi^{profit} \\ \tilde{K}' &= G(K,S) \\ k' &= e^{-\gamma S'}\tilde{k}' \\ K' &= e^{-\gamma S'}\tilde{K}' \\ S' &= S + \phi_0 E(S) \end{aligned}$$

where $u(c) = \log(c)$, $r(K,S) = \theta K^{\theta-1}[E(S)]^\nu$, $\tau(K,S) = \nu K^\theta [E(S)]^{\nu-1}$, π^{profit} is the firm's profit, and $\tilde{K}' = G(K,S)$ is the equilibrium transition law for the aggregate capital stock. Here, (k,K,S) stands for the beginning-of-period and $(\tilde{k}', \tilde{K}', S')$ for the end-of-period state. Notice that $(\tilde{k}', \tilde{K}')$ is not equal to the beginning-of-next-period state, (k', K'), due to capital deterioration by a factor $e^{-\gamma S'}$. In addition, $\hat{E}_\gamma$ is calculated with respect to the worst-case distribution for γ, $\hat{\pi}(\gamma)$, as chosen by the minimizing player. Since the minimizing player moves after the maximizing player, the worst distribution is, in general, conditional on the end-of-period state, $(\tilde{k}', \tilde{K}', S')$. It can be shown that the optimal consumption sequence satisfies the following Euler equation:

$$u'(c^*) = \beta \frac{\int e^{-\gamma S'} r(K',S') u'(c'^*) e^{-\frac{V(k',K',S')}{\alpha}} \pi(\gamma) d\gamma}{\int e^{-\frac{V(k',K',S')}{\alpha}} \pi(\gamma) d\gamma}$$

This yields the following proposition.

Proposition 2. *Assume that (A1) - (A10) hold. The optimal energy consumption is* $E = c_E(1 - \Delta S)$. *The optimal tax is* $\tau_t = \frac{\lambda^s}{u'(C^*)}$, *with tax proceeds rebated lump-sum to the representative consumer. The resulting competitive equilibrium allocation coincides with the solution to the planner's problem. That is,* $c^* = C^* = (1 - \beta\theta)K^\theta[c_E(1 - \Delta S)]^\nu$.

4 The Computational Solution and Calibration

In this section we first extend the analytical model by relaxing assumptions **(A6.1)** and **(A6.2)**. For our baseline model, we will assume that $\pi(\gamma)$, the approximating distribution of γ, is exponential. As we now allow for $\phi_L > 0$, we need to introduce two additional state variables (P and T), since keeping track of the sum $S = P + T$ will no longer suffice. We will also relax **(A7)** by incorporating a "coal" and a "green" sector into the model. Furthermore, we will relax **(A8)** and **(A9)** by allowing A_2N_2 and A_3N_3 to grow at a rate of 2 percent per year. Last, we will drop **(A10)**.

The social planner's problem becomes:

$$
\begin{aligned}
V(K,N,P,T,R) = & \max_{\{C,E_1,E_2,E_3,E,\tilde{K}',P',T',S',R'\}} \min_{m(\gamma)} \\
& \{u(C)+\beta\int\left[m(\gamma)V(K',N',P',T',R')+\alpha m(\gamma)\log m(\gamma)\right]\pi(\gamma)\mathrm{d}\gamma\} \\
s.t. & \\
E &= (\kappa_1 E_1^{\rho}+\kappa_2 E_2^{\rho}+\kappa_3 E_3^{\rho})^{1/\rho} \\
\tilde{K}' &= F\left(K, N(1-\frac{E_2}{A_2 N}-\frac{E_3}{A_3 N}), E\right)-C \\
K' &= h(S',\gamma)\tilde{K}' \\
A_2'N' &= (1+g)A_2 N \\
A_3'N' &= (1+g)A_3 N \\
R' &= R-E_1 \geq 0 \\
P' &= P+\phi_L(E_1+E_2) \\
T' &= (1-\phi)T+(1-\phi_L)\phi_0(E_1+E_2) \\
S' &= P'+T' \\
1 &= \int m(\gamma)\pi(\gamma)\mathrm{d}\gamma
\end{aligned}
$$

To solve this problem we first argue that most of the analysis conducted in Section 3 carries over. The only difference is that the function $f(\cdot)$ no longer has a closed form expression. We will again apply the outer-inner loop method used in Section 3. The inner loop minimization problem is unchanged, while the outer loop maximization problem will be solved in parts. In that regard, it is important to note that solving the optimization problem for E_i, P', T', and R' can be carried out separately from solving for C and $\tilde{K}'$. Furthermore, the solution to the second optimization problem remains the same as in Section 3; i.e., $C^* = (1-\beta\theta)Y^*$ and $\tilde{K}'^* = \beta\theta Y^*$, where Y^* denotes the optimal output level. After substituting for C^*, the optimization problem for E_i, P', T', and R' can be simplified, leading to the dynamic programming problem below:

$$f(N,P,T,R) \quad = \quad \max_{E_1,E_2,E_3,E,P',T',S',R'} \left\{ \begin{array}{l} \frac{1}{1-\beta\theta}\log\left[(1-\frac{E_2}{A_2N}-\frac{E_3}{A_3N})^{1-\theta-\nu}E^{\nu}\right] \\ \\ +\beta\left[f(N',P',T',R')+\alpha\log(1-\Delta S')\right] \end{array} \right\} \tag{14}$$

$s.t.$

$$\begin{aligned} &E=(\kappa_1E_1^{\rho}+\kappa_2E_2^{\rho}+\kappa_3E_3^{\rho})^{1/\rho}\\ &N'=(1+g)N\\ &R'=R-E_1\geq 0\\ &P'=P+\phi_L(E_1+E_2)\\ &T'=(1-\phi)T+(1-\phi_L)\phi_0(E_1+E_2)\\ &S'=P'+T' \end{aligned}$$

Next, we characterize the optimality conditions for E_3, E_2, and E_1, respectively. The first-order condition for E_3 implies

$$\frac{\nu\kappa_3}{E_3^{1-\rho}E^{\rho}}=\frac{1-\theta-\nu}{A_3N_0}.$$

The first-order condition for E_2 gives

$$\begin{aligned} &\frac{1-\theta-\nu}{A_2N_0}\\ =\;&\frac{\nu\kappa_2}{E_2^{1-\rho}E^{\rho}}+(1-\beta\theta)\beta\left[\phi_L\left(\frac{\partial f}{\partial P'}-\frac{\alpha\Delta}{1-\Delta S'}\right)+(1-\phi_L)\phi_0\left(\frac{\partial f}{\partial T'}-\frac{\alpha\Delta}{1-\Delta S'}\right)\right]. \end{aligned}$$

Applying the envelope theorem to P and T gives

$$\frac{\partial f}{\partial P} \quad = \quad \beta\left(\frac{\partial f}{\partial P'}-\frac{\alpha\Delta}{1-\Delta S'}\right) \tag{15}$$

$$\frac{\partial f}{\partial T} \quad = \quad \beta(1-\phi)\left(\frac{\partial f}{\partial T'}-\frac{\alpha\Delta}{1-\Delta S'}\right). \tag{16}$$

Defining $\hat{\Lambda}^P=-(1-\beta\theta)\frac{\partial f}{\partial P}$ and $\hat{\Lambda}^T=-(1-\beta\theta)\frac{\partial f}{\partial T}$ to be the marginal values of the externality caused by P and T, respectively, the first-order condition for E_2 becomes

$$\frac{1-\theta-\nu}{A_2N_0}=\frac{\nu\kappa_2}{E_2^{1-\rho}E^{\rho}}-\left[\phi_L\hat{\Lambda}^P+\frac{(1-\phi_L)\phi_0}{1-\phi}\hat{\Lambda}^T\right]$$

The marginal externality of S, $\hat{\Lambda}^S$, can be calculated as the consequence of a unitary increase in E_1+E_2. Note that increasing E_1+E_2 by one unit is equivalent to simultaneously increasing P by ϕ_L units and T by $\frac{(1-\phi_L)\phi_0}{1-\phi}$ units. Therefore, $\hat{\Lambda}^S$ is given by

$$\hat{\Lambda}^S=\phi_L\hat{\Lambda}^P+\frac{(1-\phi_L)\phi_0}{1-\phi}\hat{\Lambda}^T.$$

Thus, we obtain

$$\frac{\nu\kappa_2}{E_2^{1-\rho}E^\rho} - \hat{\Lambda}^S = \frac{1-\theta-\nu}{A_2 N_0}.$$

This equation has the same form as the corresponding equation in GHKT, but under a different interpretation for $\hat{\Lambda}^S$. To see the difference, it is convenient to restore the time index, t. From equation (15) and equation (16) we have

$$\begin{aligned}
\hat{\Lambda}_t^P &= (1-\beta\theta)\alpha\Delta \sum_{j=1}^{+\infty} \frac{\beta^j}{1-\Delta S_{t+j}} = \theta\bar{\gamma}\sum_{j=1}^{+\infty} \frac{\beta^j}{1-\Delta S_{t+j}} \\
\hat{\Lambda}_t^T &= (1-\beta\theta)\alpha\Delta \sum_{j=1}^{+\infty} \frac{[\beta(1-\phi)]^j}{1-\Delta S_{t+j}} = \theta\bar{\gamma}\sum_{j=1}^{+\infty} \frac{[\beta(1-\phi)]^j}{1-\Delta S_{t+j}}.
\end{aligned}$$

The second equality in either equation is obtained by using $(1-\beta\theta)\alpha\Delta = (1-\beta\theta)\alpha\frac{\bar{A}}{\alpha\lambda} = \theta\lambda^{-1} = \theta\bar{\gamma}$, where $\lambda^{-1} = \bar{\gamma}$ is the mean of γ under the approximating model. It follows immediately that $\hat{\Lambda}_t^S$ can be expressed as

$$\hat{\Lambda}_t^S = \theta\bar{\gamma}\sum_{j=1}^{+\infty}\left[\phi_L \frac{\beta^j}{1-\Delta S_{t+j}} + \frac{(1-\phi_L)\phi_0}{1-\phi}\frac{[\beta(1-\phi)]^j}{1-\Delta S_{t+j}}\right].$$

It is instructive to consider the case when $\alpha \to +\infty$; i.e., when there is no concern about model uncertainty. Observe that $\Delta \to 0$ as $\alpha \to +\infty$. Therefore,

$$\begin{aligned}
\lim_{\alpha\to+\infty} \hat{\Lambda}_t^S &= \theta\bar{\gamma}\sum_{j=1}^{+\infty}\left[\phi_L\beta^j + \frac{(1-\phi_L)\phi_0}{1-\phi}[\beta(1-\phi)]^j\right] \\
&= \theta\beta\bar{\gamma}\left[\frac{\phi_L}{1-\beta} + \frac{(1-\phi_L)\phi_0}{1-(1-\phi)\beta}\right] \qquad (17)
\end{aligned}$$

Contrasting this equation with the corresponding equation (12) in GHKT, $\hat{\Lambda}_t^S = \bar{\gamma}\left[\frac{\phi_L}{1-\beta} + \frac{(1-\phi_L)\phi_0}{1-(1-\phi)\beta}\right]$, we identify two differences. First, equation (17) contains an additional term (θ). This is because GHG directly affect aggregate capital instead of output in our model. Second, the externality related to P and T is weighted by β in equation (17). This is because GHG, in our model, affect next period's capital rather than the capital of the current period.

Finally, the first-order condition for E_1 yields

$$\frac{\nu\kappa_1}{E_1^{1-\rho}E^\rho} - \hat{\Lambda}^S = \beta\left[\frac{\nu\kappa_1}{(E_1')^{1-\rho}(E')^\rho} - (\hat{\Lambda}^S)'\right]$$

Note that the operator $\mathbb{E}_t$ does not appear on the right-hand-side, as the planner optimizes under the worst-case scenario, rather than averaging over all cases. As the planner's problem has a similar structure as in the analytical model, it can be shown that analogues of Propositions 1 and 2 hold in this environment. We numerically solve the above problem for the cases where $\alpha = 0.01$ and $\alpha = 100$. We use the same parameter values as in GHKT, except for R_0, which is set to 800, as in Rogner (1997). Figures 4 through 6 plot the computed optimal paths.

Parameter	ϕ	ϕ_L	ϕ_0	θ	ν	β	ρ	$1+g$
Value	0.0228	0.2	0.393	0.3	0.04	0.985^{10}	-0.058	1.02^{10}
Parameter	P_0	T_0	R_0	κ_1	κ_2	$A_{2,0}$	$A_{3,0}$	λ^{-1}
Value	103	699	800	0.5008	0.08916	$7,693$	$1,311$	2.379×10^{-5}

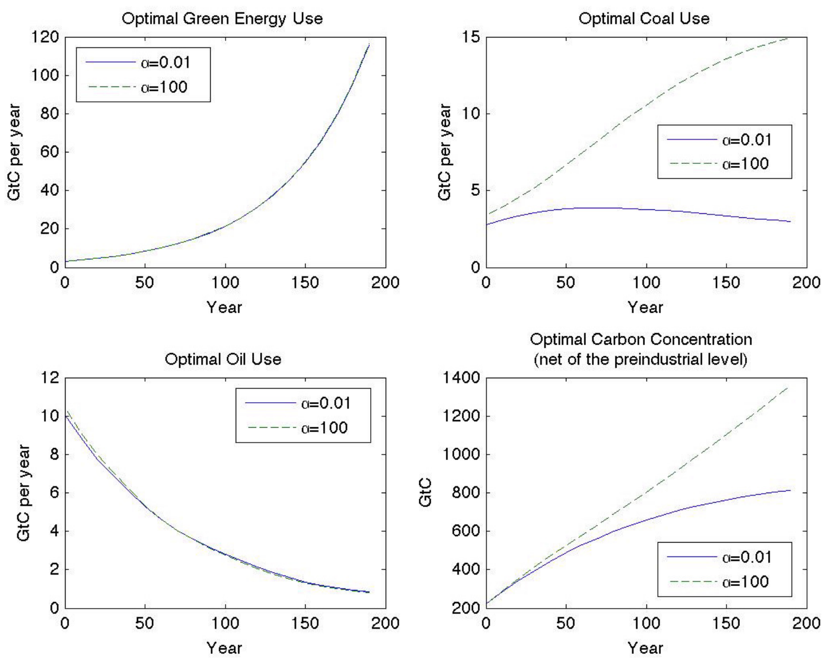

Figure 4: Optimal Use of Energy

Figure 4 describes the optimal paths for the use of green energy, coal, and oil, as well as the resulting carbon concentration in the atmosphere, conditional on different levels of concern about model uncertainty. For simplicity, we refer to the optimal path under $\alpha = 100$ as the "non-robust optimal path," and to the path under $\alpha = 0.01$ as the "robust optimal path." Since the green energy sector does not inject carbon into the atmosphere, the optimal path for the use of green energy does not directly depend on the level of concern about model uncertainty regarding the externality from carbon emissions. However, since green energy, coal, and oil are substitutes, through its impact on the "dirty" energy sectors, model uncertainty considerations do affect the use of green energy indirectly.

We find that an increase in the concern about model uncertainty causes a significant decline in the use of coal. In contrast, the use of oil is delayed, but only slightly. As the supply of oil is finite, the decline rate of oil-use depends not only on model uncertainty, but also on resource scarcity. As we will show in the next section, an initial stock of oil equaling $R_0 = 800GtC$ is low enough so that the resource scarcity effect overwhelms the model uncertainty effect in determining the optimal use of oil in the economy. This explains why we do not observe a sharp decrease in the optimal use of oil when the concern about model uncertainty increases. Finally, straightforward calculation shows that the difference in energy use in the two optimal paths leads to a significant difference in the associated carbon accumulation. Our model predicts that if there is a "small" concern about model uncertainty ($\alpha = 100$), or if model uncertainty is not incorporated into the model ($\alpha = 0.01$), atmospheric carbon concentrations will reach a level as high as $1350GtC$ (net of preindustrial levels) after 180 years. However, this number is reduced by 40 percent to about $800GtC$ if concerns about model uncertainty are incorporated and addressed through the corresponding optimal tax, restoring the optimal energy path under $\alpha = 0.01$.

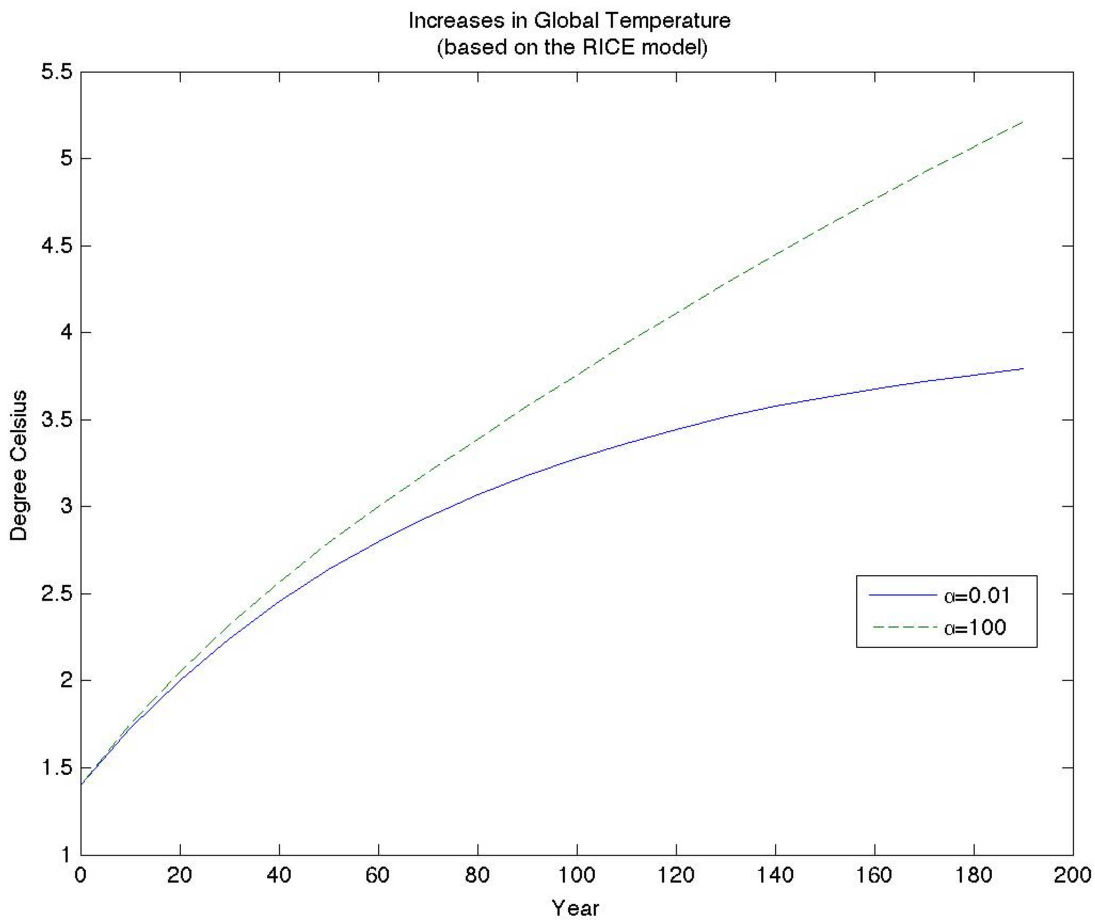

Figure 5: Increases in Global Temperature

Figure 5 demonstrates a direct consequence of the above analysis: based on the mapping from carbon concentrations to global temperatures used in the RICE model, $T(S_t) = 3\ln(\frac{S_t}{\bar{S}})/\ln 2$, the global average temperature will rise by 3.8 degrees Celsius 180 years from now if the concern about model uncertainty is addressed, and by 5.3 degrees Celsius otherwise.

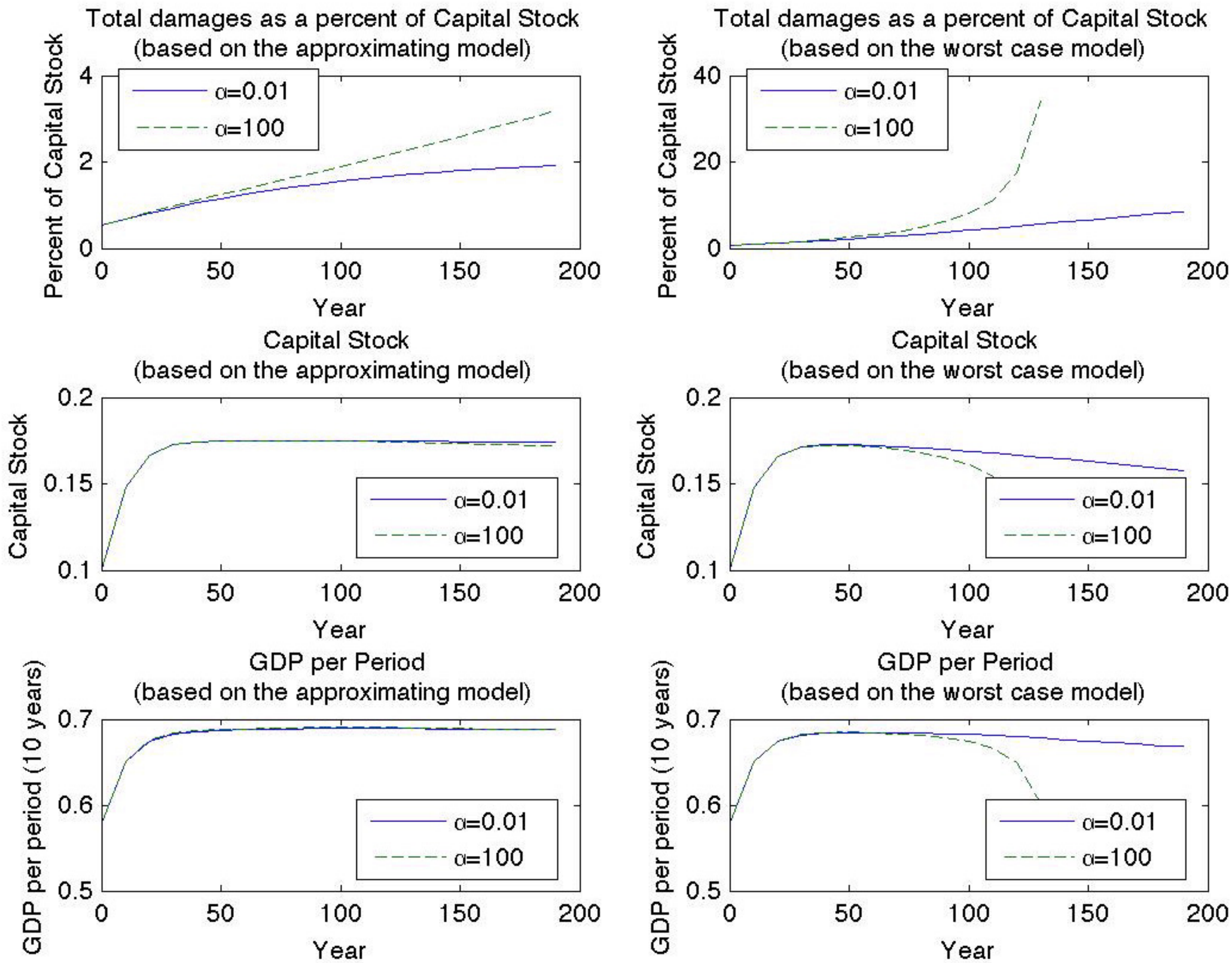

Figure 6: Capital Stock and Output

The graphs in the first (second) column in Figure 6 describe the paths of total damages as a percentage of the capital stock, as a function of the capital stock and output, assuming that the approximating model (worst-case model) for γ is the true model.[11] In each graph, the green-dashed line (blue-solid line) represents the outcome when energy is extracted based on the non-robust (robust) optimal path. The main findings can be summarized as follows. If the approximating model for γ is the true model, pursuing the robust optimal path for energy consumption would further reduce total damages by an additional 1 percent 180 years from now. However, due to a more conservative use of oil and coal in the final good sector, such a policy will also reduce both capital stock and output in the long run. Since utility depends only on consumption (which is proportional to output), this implies that the

[11]To obtain smooth paths, γ is set to be the expected mean of the approximating (worst-case) distribution(s) in each period.

welfare loss from over estimating the concern about uncertainty would be rather small. In contrast, if the true distribution of γ evolves according to the worst-case model in each period (second column of Figure 6), the cost of implementing the non-robust optimal policy is rather large. In fact, the non-robust policy, which overlooks concerns about model uncertainty, will dramatically reduce the entire capital stock in 120 years, resulting in a large reduction in output and welfare.[12]

4.1 Varying the Approximating Distribution

Here we further explore the implications of assumption **(A5)**. To this end, we now assume that the approximating distribution of γ is normal with mean $\bar{\gamma}$ and variance σ^2; i.e., $\pi(\gamma) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(\gamma-\bar{\gamma})^2}{2\sigma^2}}$. This creates two key differences. First, the normal distribution provides us with two degrees of freedom: the mean, $\bar{\gamma}$, reflecting the planner's prior expectation regarding damages, and the variance, σ^2, indicating the prior regarding model uncertainty. In comparison, recall that the exponential distribution only used one parameter, λ, which determined both the mean and the variance of γ. As we shall see below, assuming that γ is normally distributed can also eliminate the "breaking point" for S, which is always present when γ follows an exponential. This is because the exponential distribution has a "fat"tail, thus allowing more room for nature to create a worst-case-scenario given a level of penalty, α. We have

$$\begin{aligned} H(S'; \alpha, \bar{A}) &= -(\bar{\gamma} + \frac{\bar{A}\sigma^2}{2\alpha}S')\bar{A}S' \\ \hat{\pi}^*(\gamma) &\sim \mathcal{N}(\bar{\gamma} + \frac{\bar{A}\sigma^2}{\alpha}S'^2, \sigma^2). \end{aligned}$$

It is straightforward to show that $H(\cdot)$ is strictly negative, strictly increasing in α, and strictly decreasing in both $\bar{\gamma}$ and σ^2. In addition, the worst-case distribution for γ also follows a normal distribution, and $\hat{\pi}^*(\gamma)$ and $\pi(\gamma)$ differ only in their means. That is, when choosing the worst-case model, nature only alters the mean of γ, rather than its variance. As a by-product, the relative entropy of $\hat{\pi}^*(\gamma)$ with respect to $\pi^*(\gamma)$ is given by

$$\rho(\hat{\pi}^*(\gamma), \pi^*(\gamma)) = \frac{\bar{A}^2\sigma^2 S'^2}{2\alpha^2}.$$

To complete the model, we need to replace the term $\alpha\log(1-\Delta S')$ in equation (28) with $-(\bar{\gamma} + \frac{\bar{A}\sigma^2}{2\alpha}S')\bar{A}S'$. Accordingly, the optimality conditions for E_1, E_2, and E_3 remain intact, except that the values of the externality associated with P, T, and E_2 (or E_1), respectively,

[12] The dramatic effects on capital, output, and social welfare are partly due to the assumption that the approximating distribution of γ is exponential. As we discuss next, the losses are somewhat reduced, though still large, if the approximating distribution of γ is assumed to be normal. The exponential distribution is one way to capture the extreme effects in Stern (2013) in the context of our model.

are now as follows:

$$\hat{\Lambda}_t^P = \frac{\beta\theta\bar{\gamma}}{1-\beta} + \frac{\theta\bar{A}\sigma^2}{\alpha}\sum_{j=1}^{+\infty}\beta^j S_{t+j}$$

$$\hat{\Lambda}_t^T = \frac{\beta(1-\phi)\theta\bar{\gamma}}{1-\beta(1-\phi)} + \frac{\theta\bar{A}\sigma^2}{\alpha}\sum_{j=1}^{+\infty}[\beta(1-\phi)]^j S_{t+j}$$

$$\hat{\Lambda}_t^S = \phi_L\hat{\Lambda}_t^P + \frac{(1-\phi_L)\phi_0}{1-\phi}\hat{\Lambda}_t^T.$$

Note that $\hat{\Lambda}_t^S$ reduces to the previous expression as $\alpha \to +\infty$, or as $\sigma^2 \to 0$. That is,

$$\hat{\Lambda}_t^S = \theta\bar{\gamma}\left[\frac{\phi_L\beta}{1-\beta} + \frac{(1-\phi_L)\phi_0\beta}{1-(1-\phi)\beta}\right], \text{ as } \alpha \to +\infty, \text{ or } \sigma^2 \to 0.$$

We will consider three cases regarding the initial stock of fossil fuel: $R_0 = 253.8GtC$, $R_0 = 8000GtC$, and $R_0 = \infty$. While the $R_0 = \infty$ case is for expository purposes only, the other two cases are of interest. Indeed, the total stock of oil and gas is estimated to exceed $8000GtC$ if methane hydrates are included. Estimated resources of methane hydrates vary, but they alone can amount to as much as $2.1 \times 10^4 GtC$.[13] For each case, we numerically solve the above problem for $\alpha = 0.01$ and for $\alpha = +\infty$. To draw an even closer comparison with GHKT, we have rescaled γ by a factor of $1/\theta$, where θ is the share of capital. The reason is that, given a Cobb-Douglas specification in final goods production, and given 100 percent depreciation of capital, a proportional damage of $e^{-\gamma S'}$ on capital is equivalent to a proportional damage of $e^{-\theta\gamma S'}$ on output. Accordingly, the mean and variance of γ in the approximating model are set to $\bar{\gamma} = 7.93 \times 10^{-5}$ and $\sigma^2 = 2.65 \times 10^{-8}$, respectively.

Below we plot the same quantities as those shown in Figure 4 through Figure 6, but under the assumption that the approximating distribution of γ is normal. Our focus here is to compare the effects of model uncertainty on optimal oil use under different values of R_0. As we have discussed earlier, holding other parameters fixed, the optimal path of oil consumption is determined jointly by the resource scarcity effect and the model uncertainty effect. First, note that we can hardly identify a difference between the robust and the non-robust optimal paths for oil-consumption when the scarcity effect dominates, that is, when R_0 is sufficiently small. Figure 7 shows that when $R_0 = 253.8GtC$, the non-robust optimal paths replicate their counterparts in GHKT. In this case, model uncertainty delays the optimal use of oil only slightly. However, Figure 10 displays an altogether different pattern. When R_0 is set to $8000GtC$, although both paths are still decreasing over time, model uncertainty discourages the use of oil substantively. Finally, as R_0 goes to infinity, as shown in Figure 12, we observe a qualitative difference between the two paths. On the one hand, the non-robust optimal path allows the use of oil to grow unboundedly, partially due to the technological progress in the coal and green sectors. On the other hand, the increasing trend in oil consumption is curbed due to the externality caused by carbon emissions.

[13] Of course, only a small fraction of these resources is recoverable using today's technologies. See Boswell and Collett (2011). See also Hartley, Medlock, Temzelides, and Zhang (2012) and references therein.

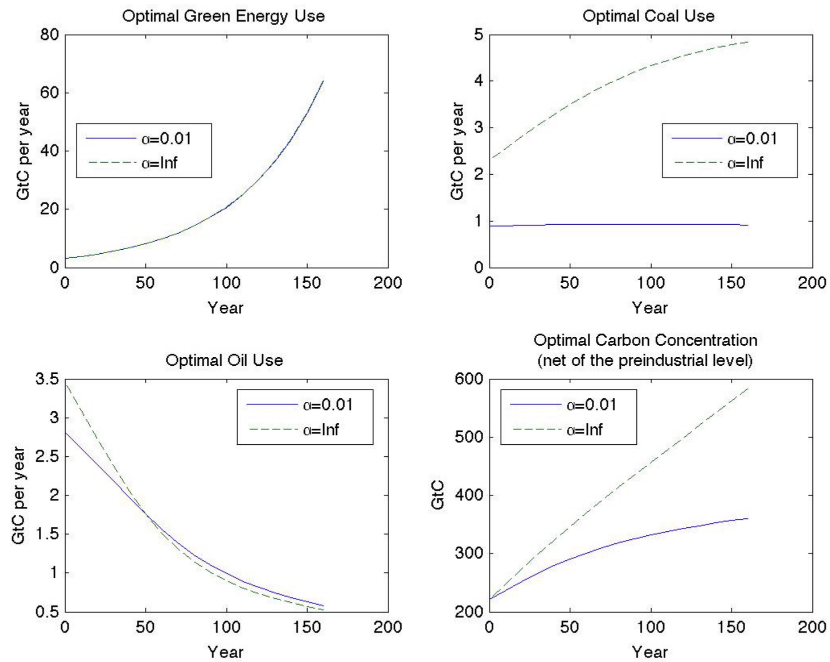

Figure 7: Optimal Use of Energy when $R_0 = 253.8$

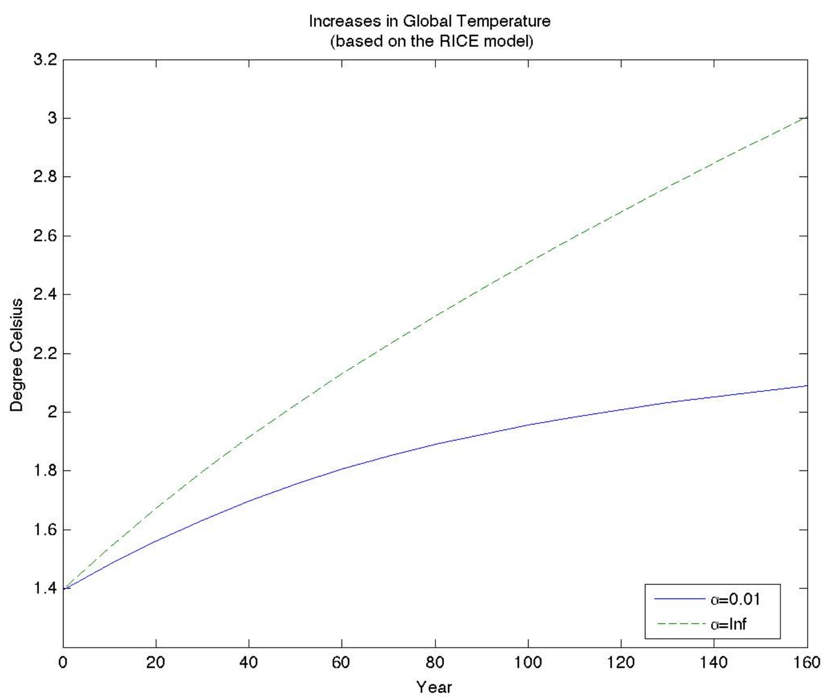

Figure 8: Increases in Global Temperature when $R_0 = 253.8$

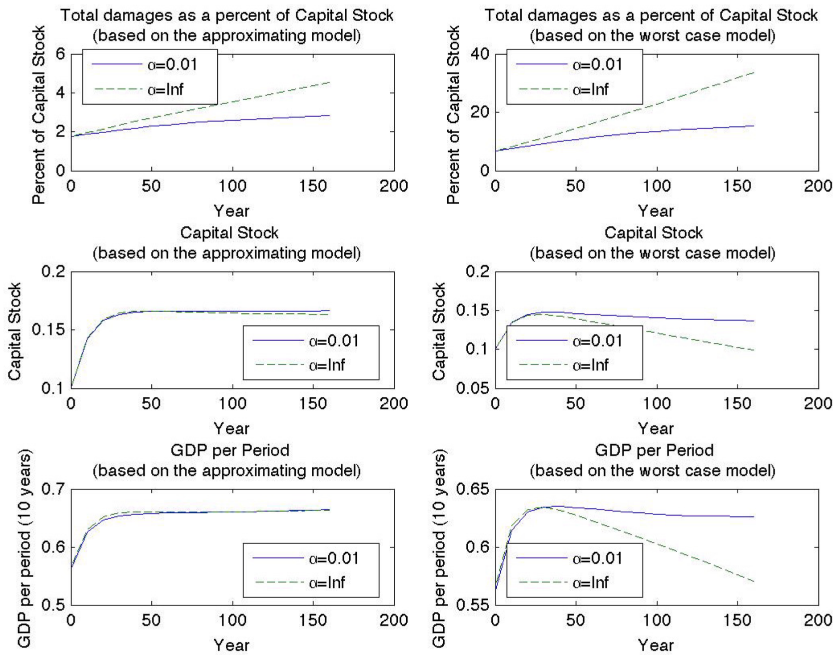

Figure 9: Capital Stock and Output when $R_0 = 253.8$

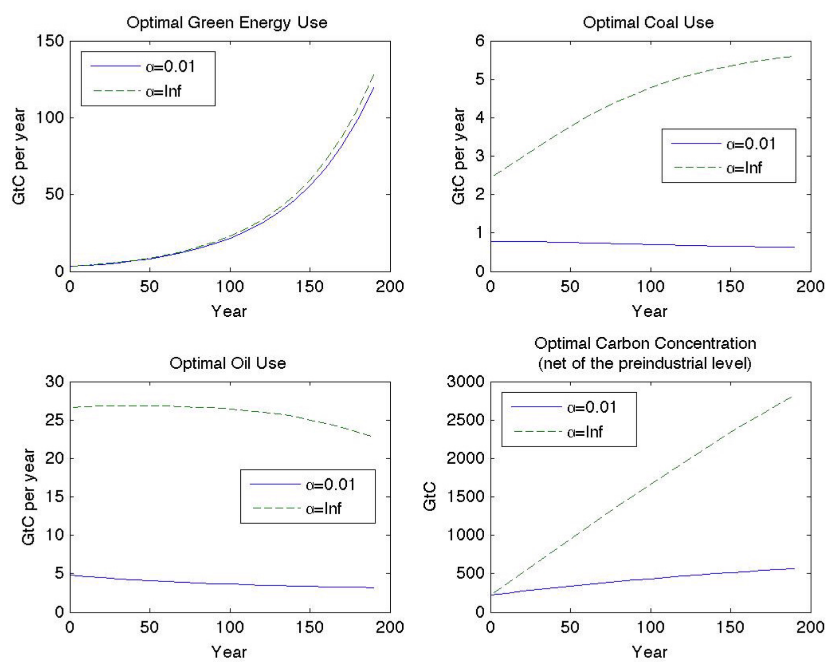

Figure 10: Optimal Use of Energy when $R_0 = 8000$

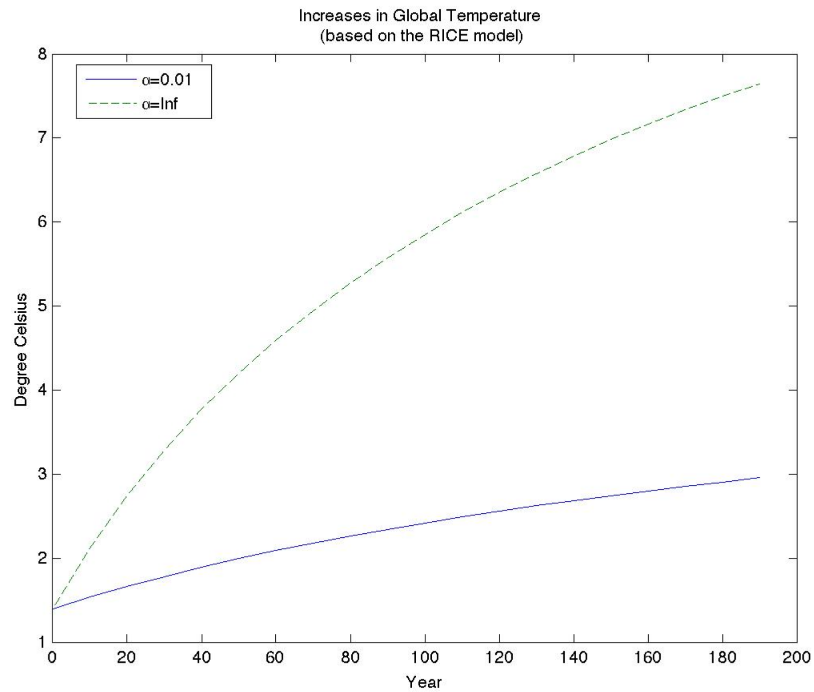

Figure 11: Increases in Global Temperature when $R_0 = 8000$

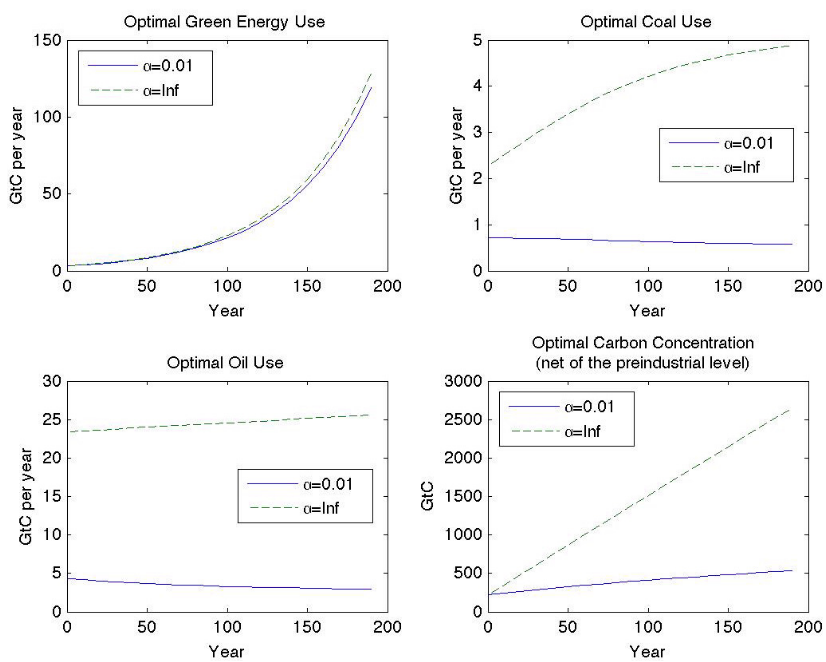

Figure 12: Optimal Use of Energy when $R_0 = \infty$

We now turn to a comparative analysis of the damages resulting from fossil fuel consumption. GHKT assume $R_o = 253.8GtC$ and estimate damages of *\$56.9/ton of carbon* using an annual discount rate of *1.5%* and *\$496/ton* under a rate of *0.1%*. When $\beta = 0.985^{10}$, and if there is no concern about model uncertainty ($\alpha = \infty$), the welfare loss implied by our model equals $0.985^{10} \times 56.4 = \$48.5/ton$. This number is independent of the approximating distribution for γ, the initial stock of oil, and the future path of the GHG concentration. When $\alpha = 0.01$, however, these factors can matter substantially, as seen below. If the approximating distribution is normal, the losses are given in the following table.

R_o/α	0.01	0.1	1	100	∞
253.8 GtC	239.60	70.65	50.85	48.52	48.49
8000 GtC	276.60	90.60	55.08	48.57	48.49
∞	318.70	103.06	63.42	56.49	48.49

4.2 Varying the Resource Feasibility Constraint

In order to further explore the model's implications, we now report the results for the limit case where oil is in infinite supply, while coal is constrained under an initial stock $R_{coal} = 666GtC$. This case demonstrates that the optimal use of oil mimics that of the case in which both oil and coal are in infinite supply. In addition, the use of coal increases steadily at the beginning and then starts to drop.

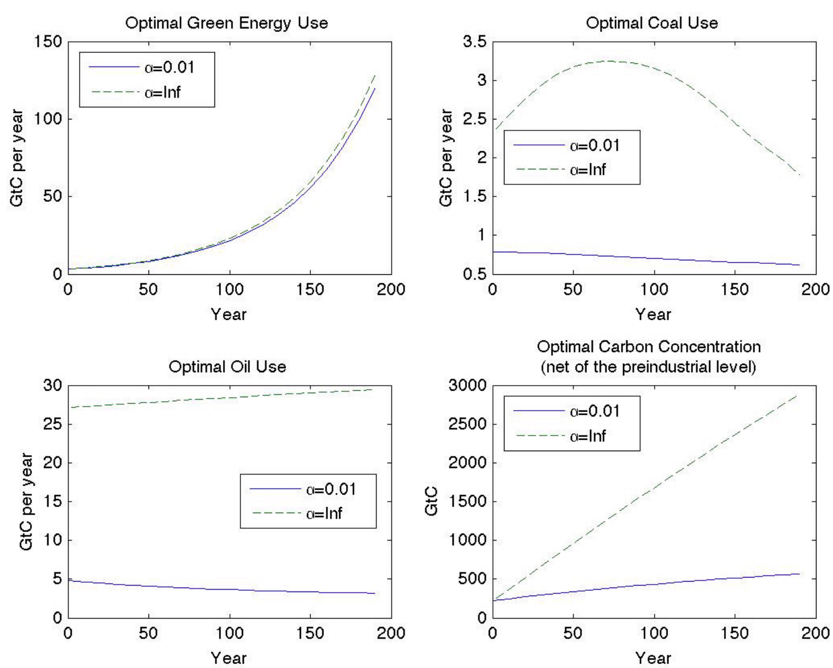

Figure 13: Optimal Use of Energy when $R_{coal} = 666$

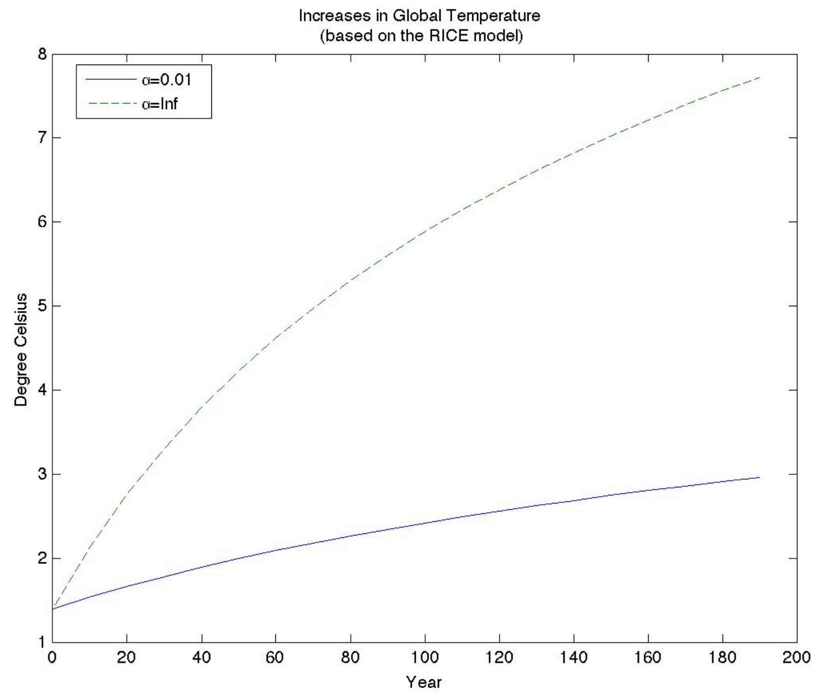

Figure 14: Increases in Global Temperature when $R_{coal} = 666$

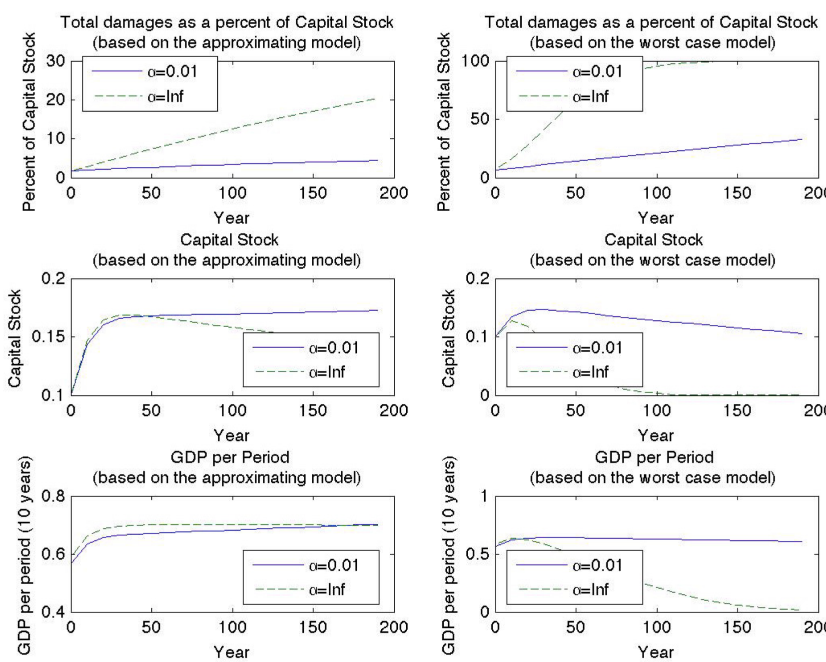

Figure 15: Capital Stock and Output when $R_{coal} = 666$

5 Discussion

We studied optimal taxation in a dynamic stochastic general equilibrium model where agents are concerned about model uncertainty regarding climate change. We used robust control theory in order to model the uncertainty associated with climate change. Our work builds heavily on the model introduced in GHKT. While admittedly restrictive, this framework allows us to derive an analytical solution. In contrast to the existing literature, we used an estimate of fossil fuel that includes methane hydrates as part of the supply of unconventional natural gas. While this huge resource is not readily available with today's technology, we believe that it is appropriate to include it given the long-term modeling that we follow throughout this exercise. Finally, we assumed a fat-tailed distribution of damages as a way to capture the extreme effects discussed in Stern (2013).

We obtained a sharp analytical solution for the implied externality, and we characterized the optimal tax. We found that a small increase in the concern about model uncertainty can cause a significant drop in optimal energy extraction. The optimal tax which restores the socially optimal allocation was shown to be Pigouvian. Under more general assumptions, we developed a recursive method that allowed us to solve the model computationally. We showed that the introduction of uncertainty matters in a number of ways, both qualitatively and quantitatively. This dependence relies heavily on specific assumptions about the magnitude of fossil fuel reserves. As our model is based on GHKT, it is worth discussing some of the main differences in our results.

Several of the variables in the model developed in GHKT can be thought of as being subject to uncertainty. These include the variables governing the dynamics of CO_2 concentration, those governing productivity growth and hence future production, the costs of alternative sources of energy (coal, oil-&-gas, and renewable), and, finally, the damages caused by the concentration of atmospheric CO_2. In this paper we concentrate on the uncertainty associated with damages from CO_2 concentration. As in GHKT, we conclude that the consumption of coal should be constrained. However, as we considere a higher stock of hydrocarbons, we derive different results regarding total consumption of fossil fuel. As a result, we showe that under a less binding resource constraint, hydrocarbon use declines significantly as the concern about model uncertainty increases.

The core theoretical result in GHKT is that, when expressed as a proportion of GDP, the optimal tax on CO_2 emissions depends only on the discount factor, the measure of the expected damage, and the depreciation of atmospheric CO_2. In particular, the tax rate is independent of the stochastic value of future output and the stock atmospheric CO_2. They derive this result based on three main assumptions: (i) logarithmic utility (which implies constant saving rate), (ii) the climate damage is proportional to GDP and has constant elasticity with respect to the level of atmospheric CO_2, and (iii) the stock of CO_2 is linear in past and current emissions. We show that once we consider model uncertainty, the Pigouvian tax can implement the optimal allocation as in GHKT. However, the expected level of damage is *no longer sufficient* for determining the optimal tax. Specifically, the optimal tax rises as the concern about uncertainty increases, even though the expected damages remain unchanged.

Our model can be extended in many ways. For comparison purposes, we tried to stay close to the parametrization used in GHKT. We could study versions of the model under different parametrizations. In the current version, the growth rate of renewables is assumed to be independent from the concern about model uncertainty. It would be interesting to endogenize growth in renewable energy productivity. A related extension could involve using a distortionary tax on labor to subsidize R&D in renewables in order to study the effects on energy composition and growth. Additionally, we could study a benchmark case where coal supply is constrained, while assuming infinite supply of gas and oil. Finally, at the cost of significant additional computational complexity, we could consider more involved climate dynamics.

6 Appendix

6.1 Model Uncertainty and Optimal Energy Extraction

We demonstrate that the optimal level of GHG, E^*, has the following properties: $\frac{\partial E^*}{\partial \delta} < 0$ and $\frac{\partial E^*}{\partial \delta}|_{\delta=0} = -\infty$, where δ is the upper bound for entropy allowed in the constraint game.

Proof. Recall that $E^* = c_E(1-\Delta S)$ and $\delta = \log(1-\Delta S'^*) + \frac{\Delta S'^*}{1-\Delta S'^*}$, where $S'^* = S + \phi_0 c_E(1-\Delta S)$. Define $a = \alpha^{-1}$ and $b = 1 - \Delta S'^* = (1-\Delta\phi_0 c_E)(1-\Delta S)$. It follows immediately that E^* is decreasing in a. In addition, since both Δ and c_E are functions of a, it follows that b is a function of a:

$$b(a) = [1 - \Delta(a)\phi_0 c_E(a)][1 - \Delta(a)S]$$

It is easy to see that b is decreasing in a. Thus, it defines a as an implicit function of b, with a negative slope. Moreover, we can rewrite δ as:

$$\delta = \log b + \frac{1-b}{b}$$

which defines b as an implicit function of δ. Direct calculation shows that $\frac{\partial b}{\partial \delta} = -\frac{b^2}{1-b} < 0$, as $b \in (0,1)$. Thus,

$$\frac{\partial E^*}{\partial \delta} = \frac{\partial E^*}{\partial a}\frac{\partial a}{\partial b}\frac{\partial b}{\partial \delta} < 0$$

Evaluating this at $\delta = 0$, we obtain

$$\frac{\partial E^*}{\partial \delta}|_{\delta=0} = \left(\frac{\partial E^*}{\partial a}|_{a=0}\right)\left(\frac{\partial a}{\partial b}|_{b=1}\right)\left(\frac{\partial b}{\partial \delta}|_{\delta=0}\right)$$

It is straightforward to show that the first two terms on the right hand side in the above expression are strictly negative and finite, and the last term goes to $-\infty$. Therefore, $\frac{\partial E^*}{\partial \delta}|_{\delta=0} = -\infty$. □

6.2 Equivalence Between the Recursive Game and the Date-0 Game

Here we discuss the equivalence between the recursive Stackelberg game and its date-0 counterpart. We concentrate on the one-sector model. In the recursive version of the Stackelberg game, the worst-case model for γ_{t+1} depends on the endogenous state S_t and on the choice variable E_t. This feature can be difficult to interpret.[14] Alternatively, we can construct a date-0 Stackelberg game in which the malevolent player, as the leader of the game, chooses the distorted models of $\{\gamma_{t+1}\}$, $\{\hat{\pi}(\gamma_{t+1})\}$, first. This leads to $\{\hat{\pi}(\gamma_{t+1})\}$ being independent of the endogenous states. We then show that, on the equilibrium path, the worst-case models derived from the date-0 Stackelberg game coincide with those derived from the recursive game. We demonstrate the equivalence by using the "big K little k" result as is in Chapter 7 of Hansen and Sargent (2008).

Consider the date-0 Stackelberg game in which, at date zero, the minimizing player chooses the distorted probability process $\{\hat{\pi}(\gamma_{t+1})\}$, followed by the maximizing player choosing the control process $\{u_t = (C_t, E_t)\}$:

$$\begin{aligned}
\inf_{m\in\mathcal{M}} \sup_{u\in\mathcal{U}} \quad & E\left[\sum_{t=0}^{\infty} \beta^t M_t \left(u(C_t) + \beta\alpha m_{t+1}\log m_{t+1}\right)|S_0, K_0\right] && (18)\\
s.t. \quad & \\
M_{t+1} &= M_t m_{t+1} \\
S_{t+1} &= S_t + \phi_0 E_t && (19)\\
K_{t+1} &= h(S_{t+1}, \gamma_{t+1})[F(K_t, E_t) - C_t] && (20)
\end{aligned}$$

where $\mathcal{U}$ denotes the space of control processes $u = \{u_t : t = 0, 1, ...\}$ and $\mathcal{M}$ denotes the space of likelihood ratio processes $m = \{m_{t+1} = \frac{\hat{\pi}(\gamma_{t+1})}{\pi(\gamma_{t+1})} : t = 0, 1, ...\}$.

We introduce an exogenous state vector process $\{(\hat{S}_t, \hat{K}_t)\}$ which evolves as:

$$\begin{aligned}
\hat{S}_{t+1} &= \hat{S}_t + \phi_0 \hat{E}_t(\hat{S}_t), \\
\hat{K}_{t+1} &= h(\hat{S}_{t+1}, \gamma_{t+1})[F(\hat{K}_t, \hat{E}_t(\hat{S}_t)) - \hat{C}_t(\hat{S}_t, \hat{K}_t)] && (21)
\end{aligned}$$

where $\hat{E}_t(\hat{S}_t) = c_E(1-\Delta\hat{S}_t)$ and $\hat{C}_t(\hat{S}_t, \hat{K}_t) = (1-\beta\theta)\hat{K}_t^{\theta}[\hat{E}_t(\hat{S}_t)]^{\nu}$.[15] Note that $\{\hat{S}_t, \hat{K}_t, \hat{E}_t, \hat{C}_t\}$ are independent of the control variables $\{E_t, C_t\}$. Moreover, we set $(S_0, K_0) = (\hat{S}_0, \hat{K}_0)$.

Define the distorted process $\{\gamma_{t+1}\}$ as

$$\gamma_{t+1} \sim \hat{\pi}(\gamma_{t+1}) = \hat{\lambda}(\hat{S}_t) e^{-\hat{\lambda}(\hat{S}_t)\gamma_{t+1}} \tag{22}$$

where the distorted parameter, $\hat{\lambda}$, is given by $\hat{\lambda}(\hat{S}_t) = \lambda(1 - \Delta\hat{S}_{t+1}) = \lambda(1 - \phi_0 c_E)(1 - \Delta\hat{S}_t)$. The last equality results from equation (48) in the main text. Clearly, u_t does not affect $\hat{S}_{t+1}$, and thus the distorted distribution $\hat{\pi}(\gamma_{t+1})$.

[14] We thank Lars Hansen for bringing this point to our attention and for suggesting the use of the "big K little k" result as a way to bypass this difficulty.

[15] The exogenous processes $\hat{E}_t(\hat{S}_t)$ and $\hat{C}_t(\hat{S}_t, \hat{K}_t)$ are constructed to mimic the optimal control $E_t^*(S_t)$ and $C_t^*(S_t, K_t)$ in equation (47) and equation (46) by replacing the endogenous state (S_t, K_t) by the exogenous state $(\hat{S}_t, \hat{K}_t)$.

Given the above exogenous distorted process, the maximizing player chooses $\{u_t\}$ at date zero to maximize the social welfare given in equation (18). With the aid of the exogenous state, this maximization problem can be expressed in a recursive form as:

$$\tilde{V}(S_t, K_t, \hat{S}_t, \hat{K}_t) = \max_{C_t, E_t} \left\{ u(C_t) + \alpha\beta \int \hat{\pi}(\gamma_{t+1}) \log m_{t+1} d\gamma_{t+1} + \beta \int \tilde{V}(S_{t+1}, K_{t+1}, \hat{S}_{t+1}, \hat{K}_{t+1}) \hat{\pi}(\gamma_{t+1}) d\gamma_{t+1} \right\},$$

subject to equation (19), equation (20), equation (21), and equation (21). The relative entropy $\int \hat{\pi}(\gamma_{t+1}) \log m_{t+1} d\gamma_{t+1}$ equals $\log(\frac{\hat{\lambda}(\hat{S}_t)}{\lambda}) + \frac{\lambda - \hat{\lambda}(\hat{S}_t)}{\hat{\lambda}(\hat{S}_t)}$, as has been shown in the main text. Since $\tilde{V}(\cdot)$ depends on $(\hat{S}_t, \hat{K}_t)$ only through $\hat{\pi}(\gamma_{t+1})$ or, equivalently, $\hat{\lambda}(\hat{S}_t)$, the exogenous state $\hat{K}_t$ is eliminated from $\tilde{V}(\cdot)$. Consequently, the above problem can be rewritten as:

$$\tilde{V}(S_t, K_t, \hat{S}_t) = \max_{C_t, E_t} \left\{ u(C_t) + \alpha\beta \left[\log(\frac{\hat{\lambda}(\hat{S}_t)}{\lambda}) + \frac{\lambda - \hat{\lambda}(\hat{S}_t)}{\hat{\lambda}(\hat{S}_t)} \right] + \beta \int \tilde{V}(S_{t+1}, K_{t+1}, \hat{S}_{t+1}) \hat{\pi}(\gamma_{t+1}) d\gamma_{t+1} \right\}, \quad (23)$$

subject to equation (19), equation (20), and equation (21).

We proceed to find the solution to this date-0 problem of the maximizing agent given the distorted process $\{\gamma_{t+1}\}$ in equation (22). Then we will argue that this solution is identical to the Markov perfect equilibrium of the sequential game defined in the main text. We implement a guess-and-verify method. We first guess that $\tilde{V}(\cdot)$ takes the form

$$\tilde{V}(S_t, K_t, \hat{S}_t) = f(S_t, \hat{S}_t) + \tilde{A} \log(K_t) + \tilde{D}$$

where $\tilde{A}$ and $\tilde{D}$ are undetermined coefficients. The functional form for $f(\cdot)$ will be derived later. Using the analysis above and simplifications in the main text, the problem can be written as

$$\begin{aligned} & f(S_t, \hat{S}_t) + \tilde{A} \log(K_t) + \tilde{D} \\ = & \max_{C_t, E_t} \left\{ \log(C_t) + \alpha\beta \left[\log(\frac{\hat{\lambda}(\hat{S}_t)}{\lambda}) + \frac{\lambda - \hat{\lambda}(\hat{S}_t)}{\hat{\lambda}(\hat{S}_t)} \right] + \beta \left[f(S_{t+1}, \hat{S}_{t+1}) + \tilde{A} \log(F(K_t, E_t) - C_t) + \tilde{D} - \frac{\tilde{A} S_{t+1}}{\hat{\lambda}(\hat{S}_t)} \right] \right\}, \end{aligned}$$

subject to equation (19) and equation (21).

Furthermore, we guess that $f(\cdot)$ takes the form $\tilde{B} \log(1 - \Delta\hat{S}_t) + \frac{\tilde{G} S_t}{1 - \Delta\hat{S}_t}$ where $\tilde{B}$ and $\tilde{G}$ are undetermined coefficients. After some tedious derivations, we obtain

$$\tilde{A} = \frac{\theta}{1 - \beta\theta} \quad (24)$$

$$\tilde{G} = \frac{\beta\theta}{(1 - \beta\theta)\lambda(\beta - 1 + \Delta\phi_0 c_E)} \quad (25)$$

and

$$E_t^{opt} = \frac{\nu(1 - \Delta\hat{S}_{t+1})}{(1 - \beta\theta)\beta\phi_0(\frac{\theta}{(1 - \beta\theta)\lambda} - \tilde{G})} = c_E(1 - \Delta\hat{S}_t) \quad (26)$$

$$C_t^{opt} = \frac{F(K_t, E_t)}{1 + \beta\tilde{A}} = (1 - \beta\theta) K_t^{\theta} (E_t^{opt})^{\nu}. \quad (27)$$

When $(S_0, K_0) = (\hat{S}_0, \hat{K}_0)$, we obtain $E_t^{opt} = \hat{E}_t(\hat{S}_t) = E_t^*(S_t)$, $C_t^{opt} = \hat{C}_t(\hat{S}_t, \hat{K}_t) = C_t^*(S_t, K_t)$, $\hat{S}_{t+1} = S_{t+1}$, and $\hat{K}_{t+1} = K_{t+1}$ for $t = 0, 1, ...$ In addition, $\hat{\pi}(\gamma_{t+1}) = \hat{\lambda}(\hat{S}_t)e^{-\hat{\lambda}(\hat{S}_t)\gamma_{t+1}}$, where $\hat{\lambda}(\hat{S}_t) = \lambda(1 - \phi_0 c_E)(1 - \Delta\hat{S}_t) = \lambda(1 - \phi_0 c_E)(1 - \Delta S_t)$. That is, the optimal choices in the date-0 game coincide with the Markov perfect equilibrium allocation in the recursive game.

6.3 The Numerical Solution for the Model

Here we provide a brief description of our numerical procedure. Assume (i) 100 percent capital depreciation, (ii) Cobb-Douglas production function, and (iii) exponential damage function. Then, it follows from the analysis in Sections 3 and 4 that the value function given in equation (61) takes the form

$$V(K, N, P, T, R) = f(N, P, T, R) + \bar{A}\log(K) + \bar{D}$$

where $\bar{A} = \frac{\theta}{1-\beta\theta}$ and $\bar{D}$ is a constant. The inner loop minimization problem for $\hat{\pi}(\gamma)$ remains the same as in the one-sector model in Section 3. Furthermore, the outer loop maximization problem for E_i, P', T', and R' can be carried out separately from the optimization problem for C and $\tilde{K}'$. The solution to the latter also remains the same as in Section 3; i.e., $C^* = (1 - \beta\theta)Y^*$ and $\tilde{K}'^* = \beta\theta Y^*$, where Y^* denotes the optimal output level. After substituting for C^*, the optimization problem for E_i, P', T', and R' can be simplified, leading to the standard dynamic programming problem below:

$$f(N, P, T, R) = \max_{E_1, E_2, E_3, E, P', T', S', R'}$$
$$\left\{ \frac{1}{1-\beta\theta} \log[(1 - \frac{E_2}{A_2 N} - \frac{E_3}{A_3 N})^{1-\theta-\nu} E^{\nu}] + \beta[f(N', P', T', R') + \alpha \log(1 - \Delta S')] \right\}$$

$$
\begin{aligned}
& s.t. \\
E &= (\kappa_1 E_1^{\rho} + \kappa_2 E_2^{\rho} + \kappa_3 E_3^{\rho})^{1/\rho} \\
N' &= (1+g)N \\
R' &= R - E_1 \geq 0 \\
P' &= P + \phi_L(E_1 + E_2) \\
T' &= (1-\phi)T + (1-\phi_L)\phi_0(E_1 + E_2) \\
S' &= P' + T'
\end{aligned}
$$

We then solve for $f(N, P, T, R)$ using a 4-dimensional Chebyshev polynomial approximation method. The above simplification has significantly reduced the computational burden of solving a dynamic max-min game, allowing us to utilize the parallel toolbox of MATLAB on a 8-processor computer. Table 1 and Table 2 below report the grid specifications used in the complete model, as well as its variations for $\alpha = 0.01$ and $\alpha = \infty$, respectively.

Table 1: Grid Specifications of Chebyshev Polynomial Approximation ($\alpha = 0.01$)

	Exponetial γ $R_0 = 800GtC$ (Oil&Gas)	Normal γ $R_0 = 253.8GtC$ (Oil&Gas)	Normal γ $R_0 = 8000GtC$ (Oil&Gas)	Normal γ $R_0 = \infty GtC$ (Oil&Gas)	Normal γ $R_0 = 666GtC$ (Coal)
# of grid points for P	6	6	6	6	6
# of grid points for T	7	7	7	7	7
# of grid points for R	10	10	10	NA	10
# of grid points for N	10	10	10	10	10
$[P_{min}, P_{max}]$	$[-200, 1000]$	$[-200, 1000]$	$[-200, 1000]$	$[-200, 1000]$	$[-200, 1000]$
$[T_{min}, T_{max}]$	$[-200, 1000]$	$[-200, 1000]$	$[-200, 1000]$	$[-200, 1000]$	$[-200, 1000]$
$[R_{min}, R_{max}]$	$[1, 900]$	$[1, 300]$	$[1, 9000]$	NA	$[1, 750]$
$[N_{min}, N_{max}]$	$[0.8, 100]$	$[0.8, 100]$	$[0.8, 100]$	$[0.8, 100]$	$[0.8, 100]$

Table 2: Grid Specifications of Chebyshev Polynomial Approximation ($\alpha = \infty$)

	Exponetial γ $R_0 = 800GtC$ (Oil&Gas)	Normal γ $R_0 = 253.8GtC$ (Oil&Gas)	Normal γ $R_0 = 8000GtC$ (Oil&Gas)	Normal γ $R_0 = \infty GtC$ (Oil&Gas)	Normal γ $R_0 = 666GtC$ (Coal)
# of grid points for P	6	6	4	6	6
# of grid points for T	7	7	4	7	7
# of grid points for R	10	10	30	NA	10
# of grid points for N	10	10	10	10	10
$[P_{min}, P_{max}]$	$[-200, 2000]$	$[-200, 1000]$	$[-200, 20000]$	$[-200, 20000]$	$[-200, 3000]$
$[T_{min}, T_{max}]$	$[-200, 2000]$	$[-200, 1000]$	$[-200, 20000]$	$[-200, 20000]$	$[-200, 3000]$
$[R_{min}, R_{max}]$	$[1, 900]$	$[1, 300]$	$[1, 9000]$	NA	$[1, 750]$
$[N_{min}, N_{max}]$	$[0.8, 100]$	$[0.8, 100]$	$[0.8, 100]$	$[0.8, 100]$	$[0.8, 100]$

References

[1] Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012): "The Environment and Directed Technical Change," *American Economic Review* 102(1) p. 131-166.

[2] Adao, B., Narajabad B., and T. Temzelides (2012): "A Model with Spillovers in the Adaptation of New renewable Technologies," James A. Baker III Institute for Public Policy Working Paper.

[3] Anderson, E., Hansen, L. P. and T. J. Sargent (2000): "Robustness, Detection and the Price of Risk," Mimeo, University of Chicago.

[4] Boswell, R., and T.S. Collett (2011): "Current Perspectives on Gas Hydrate Resources," *Energy and Environmental Science,* 4, 1206–1215.

[5] Barrage (2013): "Sensitivity Analysis for Golosov, Hassler, Krusell, and Tsyvinski (2013): Optimal Taxes on Fossil Fuel in General Equilibrium," Manuscript, University of Maryland.

[6] Bidder, R.M., and M.E. Smith (2012): "Robust Animal Spirits," *Journal of Monetary Economics*, 59, 738-750.

[7] Chandrasekharan, P. C. (1996): "*Robust Control of Linear Dynamical Systems*," Academic Press.

[8] Cogley, T., Colacito, R., Hansen, L. P. and T. J. Sargent (2008): "Robustness and U.S. Monetary Policy Experimentation," *Journal of Money, Credit and Banking,* 40 (8), 1599-1623.

[9] Dasgupta and Heal (1974): "The Optimal Depletion of Exhaustable Resources," *Review of Economic Studies*, 41.

[10] Dupuis, P., James, M. R. and I. Petersen (1998): "Robust Properties of Risk Sensitive Control," Discussion Paper No. LCDS 98-15, Brown University.

[11] Ellsberg, D. (1961): "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics,* 75, 643–669.

[12] Epstein, L. G. and T. Wang (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62 (3), 283–322.

[13] Funke, M., and M. Paetz (2010): "Environmental Policy Under Model Uncertainty: A Robust Optimal Control Approach," *Climatic Change* 107, no. 3-4: 225-239.

[14] Gars, J., Golosov M., and A. Tsyvinski (2009): "Carbon Taxing and Alternative Energy," Manuscript, Yale University.

[15] Golosov M., J. Hassler, P. Krusell, and A. Tsyvinski (2013): "Optimal Taxes on Fossil Fuel in General Equilibrium," *Econometrica*, 82 (1), 41-88.

[16] Gilboa, I., and D. Schmeidler (1989): "Maxmin Expected Utility with Non-unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

[17] Hansen L., and T. Sargent (2008): *Robustness*, Princeton University Press.

[18] Hansen, L. P., Sargent, T. J., and T. Tallarini (1999): "Robust Permanent Income and Pricing," *Review of Economic Studies,* 66, 873–907.

[19] Hansen, L. P., and T. J. Sargent (2001): "Robust Control and Model Uncertainty," *American Economic Review*, 91 (2), 60–66.

[20] Hansen, L. P., and T. J. Sargent (2003): "Robust Control of Forward-looking Models," *Journal of Monetary Economics,* 50 (3), 581 – 604.

[21] Hansen, L. P., and T. J. Sargent (2005): "Robust Control and Model Misspecification," *Journal of Economic Theory*, 128 (1), 45–90.

[22] Hansen, L. P., and T. J. Sargent (2010): "Wanting Robustness in Macroeconomics," Manuscript.

[23] Hansen, L. P., Sargent, T. J., Turmuhambetova, G. A., and N. Williams (2006): "Robust Control, Min-Max Expected Utility, and Model Misspecification," *Journal of Economic Theory,* 128, 45–90.

[24] Hansen, L. P., Sargent, T. J., and N. E. Wang (2002): "Robust Permanent Income and Pricing with Filtering," *Macroeconomic Dynamics*, 6, 40–84.

[25] Hartley, P., Medlock III K., Temzelides T., and X. Zhang (2012): "Energy Sector Innovation and Growth," James A. Baker III Institute for Public Policy Working Paper.

[26] Hennlock, M. (2008): "A Robust Feedback Nash Equilibrium in a Climate Change Policy Game," Chapter 18 in S. K. Neogy, et al., eds., *Mathematical Programming and Game Theory for Decision Making. Statistical Science and Interdisciplinary Research: Volume 1.* World Scientific Publishing Co. Pte. Ltd.

[27] Hennlock, M. (2009): "Robust Control in Global Warming Management – An Analytical Dynamic Integrated Assessment," Resources for the Future Discussion Paper RFF DP 09-19, May.

[28] Hoel, M. (1978): "Climate Change and Carbon Tax Expectations," *CESifo*Working Paper Series 2966.

[29] Hotelling (1931): "The Economics of Exhaustible Resources," *Journal of Political Economy*, 39:2.

[30] Nordhaus, W., and J. Boyer (2000): "Warming the World: Economic Modeling of Global Warming," *MIT Press, Cambridge, MA*.

[31] Kolstad, C. D. (1996: "Learning and Stock Effects in Environmental Regulation: The Case of Greenhouse Gas Emissions." *Journal of Environmental Economics and Management* 31: 1-18.

[32] Knight, F. H. (1921): *"Risk, Uncertainty and Profit,"* Houghton Mifflin Company.

[33] Krusell P., and A. Smith (2009): "Macroeconomics and Global Climate Change: Transition for a Many-Region Economy," Presentation Slides, Yale University.

[34] Lewis, F. L. (1986): "*Optimal Estimation with an Introduction to Stochastic Control Theory,*" John Wiley & Sons.

[35] Lemoine D.M., and C. Traeger (2011): "Tipping Points and Ambiguity in the Economics of Climate Change," CUDARE Working paper 1111R, Department of Agricultural and Resource Economics, University of California at Berkeley.

[36] Nordhaus W. (2008): "*A Question of Balance: Weighing the Options on Global Warming Policies,"* Yale University Press.

[37] Rogner, H.-H. (1997): "An Assessment of World Hydrocarbon Resources," *Annual Review of Energy and the Environment*, 22.

[38] Savage, L. J. (1954): *"The Foundations of Statistics,"* John Wiley & Sons.

[39] Sinn, H.-W. (2008): "Public Policies against Global Warming: A Supply Side Approach," *International Tax and Public Finance*, 15(4).

[40] Stern, N. (2007): *The Economics of Climate Change: The Stern Review*, Cambridge University Press.

[41] Stern N. (2013): "The Structure of Economic Modeling of the Potential Impacts of Climate Change: Grafting Gross Underestimation of Risk onto Already Narrow Science Models," *Journal of Economic Literature* 51(3).

[42] Sterner, T., and M. Hennlock (2011): "Knightian Uncertainty and Endogenous Growth," Manuscript.

[43] van der Ploeg, F., and C. Withagen (2012): "Too Much Coal, Too Little Oil", *Journal of Public Economics*, 96(1-2), 62-77.

[44] van der Ploeg, F., and C. Withagen (2012): "Growth, Renewables and the Optimal Carbon Tax," *International Economic Review.*

[45] Weitzman, M.L. (2014): "Fat Tails and the Social Cost of Carbon," Manuscript.

[46] Williams, N. (2008): "Robust Control," in *The New Palgrave Dictionary of Economics*, 2nd Edition (S. Durlauf and L. Blume, eds.), Palgrave Macmillan.