Continuous-Time Methods for Integrated Assessment Models[☆]

Yongyang Cai[*]

Hoover Institution, 424 Galvez Mall, Stanford University, Stanford, CA, 94305

Kenneth L. Judd

Hoover Institution, 424 Galvez Mall, Stanford University, Stanford, CA, 94305

Thomas S. Lontzek

University of Zurich, Moussonstrasse 15, 8044 Zurich

Abstract

Continuous time is a superior representation of both the economic and climate systems that Integrated Assessment Models (IAM) aim to study. Moreover, continuous-time representations are simple to express. Continuous-time models are usually solved by discretizing time, but the quality of a solution is significantly affected by the details of the discretization. The numerical analysis literature offers many reliable methods, and should be used because alternatives derived from "intuition" may be significantly inferior. We

[☆]Cai, Judd and Lontzek gratefully acknowledge NSF support (SES-0951576). Furthermore, financial support for Lontzek was provided by the Zürcher Universitätsverein and the Ecosciencia Foundation. Part of this study was done while Lontzek was visiting the Hoover Institution.

[*]The corresponding author is Dr. Yongyang Cai, Hoover Institution, 424 Galvez Mall, Stanford University, Stanford, CA, 94305. Phone: 650-926-9786. E-mail: yycai@stanford.edu.

Email addresses: yycai@stanford.edu (Yongyang Cai), kennethjudd@mac.com (Kenneth L. Judd), thomas.lontzek@business.uzh.ch (Thomas S. Lontzek)

Preprint submitted to *April 30, 2012*

take the well-known DICE model as an example. DICE uses 10-year time steps. We first identify the underlying continuous-time model of DICE. Second, we present mathematical and computational methods for transforming continuous-time deterministic perfect foresight models into systems of finite difference equations. While some transformations create finite difference systems that look like a discrete-time dynamical system, the only proper way to view the finite difference system is as an approximation of the continuous-time problem. DICE is an example where the usage of finite difference methods from numerical analysis produces far superior approximations than do simple discrete-time systems.
Keywords: Integrated Assessment Model, continuous time DICE, climate change, finite difference method

JEL Classification: Q54, C61, C63, D81

---

1. Introduction

Climate and economy are both continuous-time systems. The mutual interaction between these two systems forms the core of any Integrated Assessment Model. Nevertheless, it is common practice in the IAM literature to specify the climate-economy in discrete time, typically assuming very long discrete time-steps of 5 or 10 years. If IAMs use e.g. decadal time steps, it would be highly desirable that they properly represent the true continuous time dynamics of the underlying system and address the appropriate policies to cope with adverse effects of climate change. The insights obtained from IAMs are frequently used by policy makers to design and evaluate various climate policies, such as carbon taxes and global warming targets. For example the United States government (Interagency Working Group on Social Cost of Carbon, 2010) has recently engaged in determining the social costs of carbon, the dollar value on damages from one more ton of carbon emissions. The DICE model was one the three models used for this analysis. It comes at not surprise that DICE was part of the study. It is well known amongst the climate and economics communities and widely used in the IAM literature. Furthermore, it is well documented and simple.

Because of it simplicity and commendable openness DICE has been used and modified extensively over the last 20 years.[1] Moreover, some attempts

[1]Professor Nordhaus, the author of DICE, has always made available all equations and codes of his model. Furthermore, his calculations are extensively documented and anyone can scrutinize them. Unfortunately, within the IAM community, this is one of the few exceptions, rather than the rule. See Cai, Judd and Lontzek (2012a) for a comment on

have been made to incorporate intrinsic stochasticity into the DICE framework to study optimal climate policies under risk and uncertainty. Most of the modifications and extensions also adopte the 10-year time step formulation, and reduce the state space. We argue that, in particular when studying intrinsic uncertainty within an IAM such as DICE, great care has to be taken. For example, dynamic stochastic general equilibrium (DSGE) models in economics use relatively short time periods; usually at most a year. 10-year time steps are too long and might jeopardize economically plausible and quantitatively reliable policy analysis. If one wants to know how carbon prices should react to business cycle shocks or tipping events, the time period needs to be at most a year. Cai, Judd and Lontzek (2012b) provides such a 1-year time step DSGE version of DICE with stochastic shocks (called DSICE). Lontzek, Cai and Judd (2012) investigates the impact of the tipping point on optimal mitigation policy based on the annual DSICE model. No one would accept a policy that takes ten years to respond to current shocks to economic or climate conditions.

For the reasons outlined above, we develop DICE-CJL, a continuous-time formulation of DICE that allows an analyst to choose among several time period lengths without recalibrating the model for each different period length. A continuous-time formulation allows us to show how to use finite-difference methods from numerical analysis to formulate discrete-time versions that

openness in integrated assessment models.

can be used in computations. We demonstrate that many substantive results depend critically on the time step, strongly supporting our contention that short time periods are necessary for quantitatively reliable analysis.

In addition to the 10-year time-step length, the DICE2007 model is inconsistent in its use of finite difference methods. DICE2007 uses an explicit method for the economic module, but the climate model uses an implicit finite difference method. This mixing of explicit and implicit finite difference methods has no mathematical foundation and may lead to errors. We apply two finite difference methods to the DICE-CJL system: the Euler finite difference method, and the implicit Crank-Nicholson finite difference method. Since the true underlying model is in continuous time, modelers should ask themselves how large can the time-step be and still approximate the solution of the continuous time problem with small errors. They also need some diagnostics that tell them they have made a reliable choice. To address these issues, we apply Richardson extrapolation (Richardson and Gaunt, 1927) to DICE-CJL. We find that an annual version of DICE performs very well using either method. In addition, we compare the annual version to much shorter time-step versions. Indeed, we find that the solution of the annual version is identical to e.g. a weekly version, which de facto can be thought of as a continuous time version of DICE. We also find that a ten-year time step using Crank-Nicholson produces a good approximation. The issue is not really what is the right time step. The question is finding reliable numerical methods, and using appropriate time steps for the method we use. Overall, we find that the optimal policy results from the basic decadal DICE model are significantly distorted due to the use of an inappropriate finite difference method.

This paper proceeds as follows. Section 2 provides a summary of the DICE2007 model. Section 3 critically assesses some core assumptions of the model. Section 4 introduces the continuous-time version of DICE2007. Section 5 provides an example on how to specify the terminal value function for discrete time specifications. Section 6 offers a general discussion on how to apply a feasible finite difference scheme and calibrate parameters on DICE-CJL. Section 7 describes the DICE-CJL models using appropriate finite difference schemes. Section 8 discusses the calibration models for the DICE-CJL models. Section 9 shows the numerical solutions of DICE-CJL. Section 10 focuses on the numerical implementation and evaluation of the DICE-CJL models. Section 11 concludes.

2. DICE2007

DICE2007 (Nordhaus, 2008) maximizes social welfare with tradeoffs between carbon dioxide (CO_2) abatement, consumption, and investment. DICE2007 assumes ten-year time steps, and maximizes total discounted social utility subject to economic and climate constraints.

Nordhaus (2011) claims that DICE2007 has 18 dynamic equations. This description overstates the true complexity of DICE, which de facto has only six state variables.[2]

The annual social utility function is

$$u(c_i, l_i) = \frac{(c_i/l_i)^{1-\gamma} - 1}{1-\gamma} l_i,$$

for $i = 0, 1, \ldots, 59$, where i is the number of decades from 2005, $\gamma = 2$ in DICE2007, c_i is annual consumption and l_i is labor supply at the decade i:

$$l_i = 6514e^{-0.35i} + 8600(1 - e^{-0.35i}),$$

which is inelastic and equals world population in millions of people. Therefore, the total discounted social utility over 600 years in (Nordhaus, 2008) is

$$\sum_{i=0}^{59} \beta^{10i} 10u(c_i, l_i),$$

where $\beta = 1.015^{-1}$ is the annual discount factor.

The production side of DICE2007 is a basic optimal growth model. Output during decade i is produced from capital, k_i (measured in trillions of 2005 U.S. dollars), and labor supply l_i according to the production function

$$f_i(k_i, l_i) = A_i k_i^{\alpha} l_i^{1-\alpha}, \tag{1}$$

where $\alpha = 0.3$ is the capital share, and A_i is total productivity factor defined

[2]DICE2007 includes the cumulative resource stock of carbon as a state variable. It enters the model as a constraint on the cumulative extraction of fossil fuels. This constraint is not binding. Therefore, we omit this redundant state variable throughout our analysis.

by

$$
\begin{aligned}
A_0 &= 0.02722, \\
A_{i+1} &= \frac{A_i}{1 - 0.092e^{-0.01i}}.
\end{aligned} \tag{2}
$$

Global average atmospheric temperature, T_i^{AT} (measured in degrees Celsius above the 1900 temperature), reduces output by a factor

$$
\Omega_i = \frac{1}{1 + \pi_1 T_i^{\mathrm{AT}} + \pi_2 (T_i^{\mathrm{AT}})^2},
$$

where $\pi_1 = 0$ and $\pi_2 = 0.0028388$. Abatement efforts can reduce CO2 emissions at some cost. Therefore, net output during decade i is

$$
\mathcal{Y}_i(k_i, T_i^{\mathrm{AT}}, \mu_i) = \left(1 - \psi_i^{1-\theta_2} \theta_{1,i} \mu_i^{\theta_2}\right) \Omega_i f_i(k_i, l_i), \tag{3}
$$

where $\theta_2 = 2.8$, ψ_i is the participation rate (it is assumed to be equal to 1 after 2010 in DICE2007), $\mu_i \in [0, 1]$ is the emission control rate, and

$$
\theta_{1,i} = \frac{1.17\sigma_i(1 + e^{-0.05i})}{2\theta_2} \tag{4}
$$

is the adjusted cost for backstop, where σ_i is the technology factor following the path

$$
\begin{aligned}
\sigma_0 &= 0.13418, \\
\sigma_{i+1} &= \frac{\sigma_i}{1 + 0.073e^{-0.03i}}.
\end{aligned} \tag{5}
$$

Thus, the next-decade capital is

$$k_{i+1} = (1-\delta)^{10} k_i + 10 \left(\mathcal{Y}_i(k_i, T_i^{\mathrm{AT}}, \mu_i) - c_i \right), \tag{6}$$

where $\delta = 0.1$ is the annual rate of depreciation of capital.

Industrial production processes cause CO2 emissions

$$E_i^{\mathrm{Ind}}(k_i, \mu_i) = \sigma_i (1-\mu_i) f_i(k_i, l_i), \tag{7}$$

so the annual total carbon emissions (billions of metric tons) during decade i is

$$\mathcal{E}_i(k_i, \mu_i) = E_i^{\mathrm{Ind}}(k_i, \mu_i) + E_i^{\mathrm{Land}}, \tag{8}$$

where

$$E_i^{\mathrm{Land}} = 1.1 \times 0.9^i$$

represents annual emissions from biological processes during decade i.

DICE2007 uses a simple box model for the carbon cycle. The CO_2 concentrations for the carbon cycle are modeled by a three-layer model,

$$\mathbf{M}_i = (M_i^{\mathrm{AT}}, M_i^{\mathrm{UP}}, M_i^{\mathrm{LO}})^\top,$$

representing carbon concentration (in billions of metric tons) in the atmosphere (M_i^{AT}), upper oceans (M_i^{UP}) and lower oceans (M_i^{LO}) . The transition system of the CO_2 concentration from decade i to next decade $i+1$ is

$$\mathbf{M}_{i+1} = \mathbf{\Phi}_{\mathrm{DICE2007}}^{\mathrm{M}} \mathbf{M}_i + 10 \left(\mathcal{E}_i(k_i, \mu_i), 0, 0 \right)^\top,$$

where $\mathbf{\Phi}^{\mathrm{M}}_{\mathrm{DICE2007}}$ is the carbon diffusion matrix (flows per decade),

$$\mathbf{\Phi}^{\mathrm{M}}_{\mathrm{DICE2007}} = \begin{bmatrix} 0.810712 & 0.097213 & 0 \\ 0.189288 & 0.852787 & 0.003119 \\ 0 & 0.05 & 0.996881 \end{bmatrix}.$$

The CO_2 concentrations impact the surface temperature of the globe through the radiative forcing (watts per square meter from 1900):

$$\mathcal{F}_i\left(M_i^{\mathrm{AT}}\right) = \eta \log_2\left(M_i^{\mathrm{AT}}/M_0^{\mathrm{AT}}\right) + F_i^{\mathrm{EX}}, \tag{9}$$

where $\eta = 3.8$ and F_i^{EX} is the exogenous radiative forcing:

$$F_i^{\mathrm{EX}} = \begin{cases} -0.06 + 0.036i, & \text{if } i \leq 10, \\ 0.3, & \text{otherwise.} \end{cases}$$

DICE2007 uses a simple box model for the climate. The global mean temperature is represented by a two-layer model,

$$\mathbf{T}_i = (T_i^{\mathrm{AT}}, T_i^{\mathrm{LO}})^{\top},$$

representing temperature (measured in degrees Celsius above the 1900 temperature) of the atmosphere (T_i^{AT}) and lower oceans (T_i^{LO}). The transition system of the global mean temperature from decade i to next decade $i+1$ is

$$\mathbf{T}_{i+1} = \mathbf{\Phi}^{\mathrm{T}}_{\mathrm{DICE2007}}\mathbf{T}_i + 10\left(\xi_1 \mathcal{F}_i\left(M_i^{\mathrm{AT}}\right), 0\right)^{\top},$$

where $\xi_1 = 0.022$, and $\mathbf{\Phi}^{\mathrm{T}}$ is the climate temperature diffusion matrix per decade,

$$\mathbf{\Phi}^{\mathrm{T}}_{\mathrm{DICE2007}} = \begin{bmatrix} 0.787333 & 0.066 \\ 0.05 & 0.95 \end{bmatrix}.$$

Therefore, Nordhaus (2008) solves the problem

$$\begin{aligned} \max_{c_i, \mu_i} \quad & \sum_{i=0}^{59} \beta^{10i} 10 u(c_i, l_i) \\ \text{s.t.} \quad & k_{i+1} = (1-\delta)^{10} k_i + 10 \left(\mathcal{Y}_i(k_i, T_i^{\mathrm{AT}}, \mu_i) - c_i \right), \\ & \mathbf{M}_{i+1} = \mathbf{\Phi}^{\mathrm{M}}_{\mathrm{DICE2007}} \mathbf{M}_i + 10 \left(\mathcal{E}_i \left(k_i, \mu_i \right), 0, 0 \right)^{\top}, \\ & \mathbf{T}_{i+1} = \mathbf{\Phi}^{\mathrm{T}}_{\mathrm{DICE2007}} \mathbf{T}_i + 10 (\xi_1 \mathcal{F}_i \left(M_i^{\mathrm{AT}} \right), 0)^{\top}. \end{aligned}$$

where the social planner has two control variables, c_i and μ_i, and there are six states $(k_i, \mathbf{M}_i, \mathbf{T}_i)$, at each decade i. [3]

3. Numerical Issues with DICE2007

Before we formulate a continuous time version of DICE and proceed with our analysis, we would like to point out some critical issues with the current version of DICE. Users familiar with the DICE2007 version (or other versions of DICE) will find different results when running DICE-CJL. These difference come from using shorter time-steps and the adjustments described in this section.

[3] We omit the redundant constraint on the cumulative stock of carbon.

3.1. Nonphysical Nature of DICE2007

The equations in DICE2007 differ from the corresponding ones in earlier versions of DICE. Nordhaus (2008) explains these changes:

> "The lags in the system are primarily caused by the diffusive inertia of the different layers. We have changed the timing slightly to improve the match of the impulse-response function with climate models. Additionally, we have adjusted the climate sensitivity to the center of the IPCC range of 3°C for an equilibrium CO_2 doubling. The timing is calibrated to match model experiments for the IPCC Third and Fourth Assessment Reports."

The changes implied nonphysical phenomena. In the program of DICE2007,

$$\mathcal{F}_i\left(M_i^{\mathrm{AT}}\right) = \eta \log_2\left(\frac{M_i^{\mathrm{AT}} + M_{i+1}^{\mathrm{AT}}}{2M_0^{\mathrm{AT}}}\right) + F_i^{\mathrm{EX}},$$

and

$$\mathbf{T}_{i+1} = \mathbf{\Phi}_{\mathrm{DICE2007}}^{\mathrm{T}}\mathbf{T}_i + 10\left(\xi_1 \mathcal{F}_{i+1}\left(M_{i+1}^{\mathrm{AT}}\right), 0\right)^{\top}.$$

This implies that

$$\mathbf{T}_{i+1} = \mathbf{\Phi}_{\mathrm{DICE2007}}^{\mathrm{T}}\mathbf{T}_i + 10\left[\xi_1\left(\eta \log_2\left(\frac{M_{i+1}^{\mathrm{AT}} + M_{i+2}^{\mathrm{AT}}}{2M_0^{\mathrm{AT}}}\right) + F_{i+1}^{\mathrm{EX}}\right), 0\right]^{\top}.$$

Thus, T_{i+1}^{AT} depends on M_{i+2}^{AT}. For example, if you set $t = 2015$, then warming between 2015 and 2025 is affected by the stock of atmospheric CO_2 in 2035, implying that emissions between 2025 and 2035 will increase warming during the ten years between 2015 and 2025. Nonphysical specifications of a physical system are problematic in general, but particularly in an optimal control problem, the focus of DICE, where the non-physicality of the climate system will introduce non-Markovian features to what is really a Markov decision problem.

Following standard practice in the IAM literature, Nordhaus (2008) chose the parameters of the climate system so that a particular finite difference method with a ten-year time period matches some target results. This is not a standard way to deal with numerical design questions. The underlying problem, a combination of climate and economic systems, is a system of differential equations. The fundamental parameters are fixed by empirical evidence. If a finite difference scheme is not producing accurate results then it is natural to change it, but only to a theoretically valid finite difference method of the given continuous-time econo-physical system.

Climate models are run at time resolutions far shorter than ten years; some are run at resolutions measured in minutes. Climate models aim to solve continuous-time models. It is unclear that it is possible to produce a specification of the climate system with a time resolution of ten years that can match important features of a continuous-time model. A more standard response to problems with lags due to large time steps would be to solve the model at a finer time resolution. Doing this for DICE is straightforward. It just requires a proportional change of those parameters with time unit dimensions; otherwise, no change in parameters or functional form of the equations is necessary. Therefore, for the DICE diffusivity parameters, it is easy to solve the problem with one-year time steps.

We recomputed DICE2007 for time steps ranging from three months to ten years. The results for the short time periods were very close, but differed substantially from the ten year results. In particular, the ten-year formulation produced carbon taxes 50% greater than the results using shorter time periods. See Cai, Judd and Lontzek (2012a) and the accompanying website.

3.2. Terminal Condition

DICE2007 imposes an additional constraint on terminal capital by assuming that investment at the terminal time must be at least 2% of the capital stock at the terminal time. Moreover, DICE2007 assumes that the terminal value function is 0 everywhere. This is not a reasonable assumption, as this implies that people will consume the capital stock as much as possible before the terminal time and do not control carbon emission or temperature at the last period.

3.3. Flexible Savings Rate

In the program of DICE2007, the savings rate is fixed at 0.22 for standardization. A fixed savings rate is a good approximation here because of the inelastic labor supply and the power utility function over consumption. However, in general, it is inappropriate to take a fixed savings rate as a solution, although it is fine to use a fixed savings rate in preliminary computation to generate an initial guess to feed into the real problem. We cancel this fixed savings rate constraint in our examples.

3.4. Extraneous Variables

DICE2007 looks complicated to some users. Its GAMS code has 1263 equations and 1381 variables. However, there are many variables (and equations) that could be cancelled, as they are, in fact, autonomous and time-dependent parameters. These extra variables and equations make it time-consuming for users to understand the essence of the model. After cleaning up these extra variables and equations, DICE-CJL (with the same 10-year time periods) has only 466 equations and 584 variables, about only one third of the original DICE code.

4. A Continuous-Time Reformulation of DICE

The first step towards understanding the computational points we raise below is to see the true underlying continuous-time model.

First, the total discounted utility is

$$\int_0^\infty e^{-\rho t} u(c(t), l(t)) dt,$$

where $\rho = 0.015$ is the discount rate, $c(t)$ is the consumption function. We assume that labor supply $l(t)$ is inelastic and equal to the population, which evolves according to

$$l(t) = 6514 e^{-0.035t} + 8600(1 - e^{-0.035t}), \tag{10}$$

for any continuous time t in units of years.

Second, in the production function (1), the total productivity factor is computed by the recursive formula (2), but in fact it could be represented by a function of the continuous time:

$$A(t) = A_0 \exp\left(0.0092(1 - e^{-0.001t})/0.001\right), \tag{11}$$

and in the industrial emission function (7), instead of the recursive formula (5), the technology factor is also represented by a function of continuous time:

$$\sigma(t) = \sigma_0 \exp\left(-0.0073(1 - e^{-0.003t})/0.003\right). \tag{12}$$

Moreover, the function (4) for the adjusted cost for backstop becomes

$$\theta_1(t) = \frac{1.17\sigma(t)\left(1 + e^{-0.005t}\right)}{2\theta_2}. \tag{13}$$

With these continuous-time formulas, the net output at time t is

$$\mathcal{Y}(k, T^{\mathrm{AT}}, \mu, t) = \frac{1 - \psi(t)^{1-\theta_2}\theta_1(t)\mu^{\theta_2}}{1 + \pi_1 T^{\mathrm{AT}} + \pi_2 (T^{\mathrm{AT}})^2} A(t) k^{\alpha} l(t)^{1-\alpha}, \tag{14}$$

for any capital $k > 0$, surface temperature T^{AT} and emission control rate $\mu \in [0, 1]$, where $\psi(t)$, the participation rate, is assumed to be equal to 1 in our examples. Thus, the differential equation of capital is

$$\dot{k} = \mathcal{Y}\left(k, T^{\mathrm{AT}}, \mu, t\right) - c - \delta k, \tag{15}$$

for any continuous time t.

Third, with the continuous time formulas of the productivity factor, the technology factor and the adjusted cost for backstop, the rate of carbon emissions at time t becomes

$$\mathcal{E}(k, \mu, t) = \sigma(t)(1-\mu)A(t)k^{\alpha}l(t)^{1-\alpha} + E^{\mathrm{Land}}(t),$$

for any capital $k > 0$ and emission control rate $\mu \in [0, 1]$, where

$$E^{\mathrm{Land}}(t) = 1.1e^{-0.01t}$$

is the rate of carbon emissions from biological processes. Thus, the carbon cycle system is

$$\dot{\mathbf{M}} = \mathbf{\Phi}^{\mathrm{M}}\mathbf{M} + \left(\mathcal{E}\left(k, \mu_t, t\right), 0, 0\right)^{\top}, \tag{16}$$

where

$$\mathbf{\Phi}^{\mathrm{M}} = \begin{bmatrix} -\phi_{12} & \phi_{12}\varphi_1 & 0 \\ \phi_{12} & -\phi_{12}\varphi_1 - \phi_{23} & \phi_{23}\varphi_2 \\ 0 & \phi_{23} & -\phi_{23}\varphi_2 \end{bmatrix}, \tag{17}$$

where $\varphi_1 = M_*^{\mathrm{AT}}/M_*^{\mathrm{UP}}$ and $\varphi_2 = M_*^{\mathrm{UP}}/M_*^{\mathrm{LO}}$, where M_*^{AT}, M_*^{UP} and M_*^{LO} are the preindustrial equilibrium states of the carbon cycle system.

Moreover, the total radiative forcing rate becomes

$$\mathcal{F}\left(M^{\mathrm{AT}}, t\right) = \eta \log_2\left(M^{\mathrm{AT}}/M_0^{\mathrm{AT}}\right) + F^{\mathrm{EX}}(t), \tag{18}$$

for any carbon concentration in the atmosphere M^{AT}, where the exogenous radiative forcing rate is

$$F^{\mathrm{EX}}(t) = \begin{cases} -0.06 + 0.0036t, & \text{if } t \leq 100, \\ 0.3, & \text{otherwise.} \end{cases}$$

Thus, the temperature system satisfies the following differential system

$$\dot{\mathbf{T}} = \mathbf{\Phi}^{\mathrm{T}}\mathbf{T} + \left(\xi_1 \mathcal{F}\left(M^{\mathrm{AT}}, t\right), 0\right)^{\top}, \tag{19}$$

where

$$\mathbf{\Phi}^{\mathrm{T}} = \begin{bmatrix} -\xi_1\eta/3 - \xi_1\xi_3 & \xi_1\xi_3 \\ \xi_4 & -\xi_4 \end{bmatrix}. \tag{20}$$

Therefore, the continuous time model becomes

$$\begin{aligned} \max_{c,\mu} \quad & \int_0^{\infty} e^{-\rho t} u(c, l(t)) dt \\ \text{s.t.} \quad & \dot{k} = \mathcal{Y}\left(k, T^{\mathrm{AT}}, \mu, t\right) - c - \delta k, \\ & \dot{\mathbf{M}} = \mathbf{\Phi}^{\mathrm{M}}\mathbf{M} + \left(\mathcal{E}\left(k, \mu, t\right), 0, 0\right)^{\top}, \\ & \dot{\mathbf{T}} = \mathbf{\Phi}^{\mathrm{T}}\mathbf{T} + \left(\xi_1 \mathcal{F}\left(M^{\mathrm{AT}}, t\right), 0\right)^{\top}, \end{aligned}$$

where there are six continuous time state variables $(k, \mathbf{M}, \mathbf{T})$ and two continuous time control variables (c, μ).

5. Terminal Value Function

DICE2007 solves a 600-year horizon optimization problem. Hence, we change the infinite-horizon continuous time model to a 600-year horizon problem by replacing the integration of discounted utilities from the terminal time (the 600th year) to infinity with a terminal value function. We estimate the terminal value function using the summation of discounted utilities which is discretized over $[600, \infty)$ using a one-year interval. Some might argue that, due to discounting, economic costs and benefits in the far distant future have little value from today's point of view.[4] Nevertheless, our aim is to provide the IAM community with numerical tools which facilitate the design and execution of an IAM in accordance with accepted standards in mathematics and economics.

Assume that at the terminal time, the capital is $\tilde{k}$, the three-layer CO_2 concentration is $\tilde{\mathbf{M}}$, the two-layer global mean temperature is $\tilde{\mathbf{T}}$. For any time $t \geq 600$, we assume that the population is $l(t) = \tilde{l} = 8600$, the total production factor and the adjusted cost for backstop will be the same with the numbers at the terminal time respectively, i.e., $A(t) = 1.7283$ and $\theta_1(t) = 0.00386$. We assume that at terminal time, the world reaches a partial equilibrium: after the terminal time, capital will be the same, and emission control rate will always be 1, so that emission of carbon from industry will always be 0, i.e., $k(t) = \tilde{k}$ and $\mu(t) = 1$, for any time $t \geq 600$. Thus, using the explicit Euler method discussed in Section 7, the one-year discretized dynamics of the climate system becomes

$$\begin{aligned} \mathbf{M}_{t+1} &= \mathbf{M}_t + \mathbf{\Phi}^{\mathrm{M}}\mathbf{M}_t + \left(E^{\mathrm{Land}}(t), 0, 0\right)^\top, \\ \mathbf{T}_{t+1} &= \mathbf{T}_t + \mathbf{\Phi}^{\mathrm{T}}\mathbf{T}_t + \left(\xi_1 \mathcal{F}\left(M_t^{\mathrm{AT}}, t\right), 0\right)^\top, \end{aligned}$$

for any year $t \geq 600$, where $\mathbf{M}_{600} = \tilde{\mathbf{M}}$, $\mathbf{T}_{600} = \tilde{\mathbf{T}}$.

To keep the above partial equilibrium, the consumption at year $t \geq 600$

[4] E.g. a one percent annual discount rate over 600 years results in a discount factor of roughly one quarter of a percent.

is

$$c_t = \mathcal{Y}(\tilde{k}, T_t^{\mathrm{AT}}, 1, t) - \delta \tilde{k}.$$

Therefore, we have our terminal value function:

$$V(\tilde{k}, \tilde{\mathbf{M}}, \tilde{\mathbf{T}}) = \sum_{t=600}^{\infty} e^{-\rho(t-600)} u(c_t, \tilde{l}).$$

To compute the terminal value function, we will use the summation of discounted utilities over 800 years from $t = 600$ to $t = 1399$ with one year as the time interval for each period instead. It will be a very good approximation of the summation of the infinite sequence, because $e^{-800\rho} \approx 6.1 \times 10^{-6}$ is small enough. That is,

$$V(\tilde{k}, \tilde{\mathbf{M}}, \tilde{\mathbf{T}}) \approx \sum_{t=600}^{1399} e^{-\rho(t-600)} u(c_t, \tilde{l}).$$

It would be too time-consuming to use the terminal value function of the above formula in optimizers to compute optimal solutions, so we will use its approximation to save computational time. In our examples, we will use a degree-4 complete Chebyshev polynomial approximation, $\hat{V}(k, \mathbf{M}, \mathbf{T})$, over the 6-dimensional state space where $k \in [50000, 90000]$, $M^{\mathrm{AT}} \in [650, 1050]$, $M^{\mathrm{UP}} \in [1400, 1800]$, $M^{\mathrm{LO}} \in [19000, 21000]$, $T^{\mathrm{AT}} \in [1, 3]$, and $T^{\mathrm{LO}} \in [1.5, 3.5]$. Detailed discussion of complete Chebyshev polynomials can be found in Judd (1998), Cai (2009) and Cai and Judd (2010).

Using the terminal value function, we have the new model:

$$\begin{aligned} \max_{c,\mu} \quad & \int_0^{600} e^{-\rho t} u(c, l(t))dt + e^{-600\rho}\hat{V}\left(k_{600}, \mathbf{M}_{600}, \mathbf{T}_{600}\right), \\ \text{s.t.} \quad & \dot{k} = \mathcal{Y}\left(k, T^{\mathrm{AT}}, \mu, t\right) - c - \delta k, \\ & \dot{\mathbf{M}} = \mathbf{\Phi}^{\mathrm{M}}\mathbf{M} + \left(\mathcal{E}\left(k, \mu, t\right), 0, 0\right)^{\top}, \\ & \dot{\mathbf{T}} = \mathbf{\Phi}^{\mathrm{T}}\mathbf{T} + \left(\xi_1 \mathcal{F}\left(M^{\mathrm{AT}}, t\right), 0\right)^{\top}. \end{aligned} \tag{21}$$

6. Solving and Calibrating Differential Equation Models

Suppose that you have an ODE

$$\frac{dx(t)}{dt} = f(x(t), t),$$

while $x(t_0) = x_0$ is given at the initial time t_0. Integrating the ODE,

$$x(t) = x(t_0) + \int_{t_0}^{t} f(x(s), s)ds.$$

Defining a sequence of times $t_n = t_0 + nh$ where h is the "step size", we have

$$x(t_{n+1}) = x(t_n) + \int_{t_n}^{t_{n+1}} f(x(s), s)ds.$$

We denote by x_n a numerical estimate of $x(t_n)$, i.e., $x_n \approx x(t_n)$, by estimating the integration of f over $[t_n, t_{n+1}]$, for $n = 1, 2, \ldots$. There are many ways to numerically compute the integration. The explicit Euler method has the form

$$x_{n+1} = x_n + hf(x_n, t_n).$$

From numerical analysis (Iserles, 1996), the explicit Euler method is convergent[5] and its numerical error decays as $O(h)$, i.e., linearly when h is halved.

The integration of f over $[t_n, t_{n+1}]$ can be estimated more accurately by the trapezoid rule, so the Crank-Nicholson method is derived:

$$x_{n+1} = x_n + \frac{h}{2} \left(f\left(x_n, t_n\right) + f\left(x_{n+1}, t_{n+1}\right) \right).$$

Since x_{n+1} can not be produced explicitly by knowing x_n and computing a value of f like the explicit Euler method, the Crank-Nicholson method is said to be implicit. From numerical analysis (Iserles, 1996), the Crank-Nicholson method is convergent and its numerical error decays as $O(h^2)$, i.e., quadratically when h is halved.

Suppose one has some unknown parameters of the ODE in the function f and wants to choose them so that the solution hits some target points. Denote a as the unknown parameters, and the function f has the form $f(x, t; a)$. Assume that x_n^* is a given state sequence that we want to match. Then we can write this as a minimum norm problem. We can use $\mathcal{L}^1$ or $\mathcal{L}^2$ objectives.

[5] The method is convergent if there exists a real constant λ such that the function f that maps $\mathbb{R}^d \times [t_0, \infty)$ to $\mathbb{R}^d$ satisfies that

$$\|f(x,t) - f(y,t)\| \leq \lambda \|x - y\| \qquad \text{for all } x, y \in \mathbb{R}^d,\ t \geq t_0,$$

in a given norm $\|\cdot\|$, and the Taylor series of f about every $(x_0, t) \in \mathbb{R}^d \times [t_0, \infty)$ has a positive radius of convergence.

That is,

$$\begin{aligned} \min_{a} \quad & \sum_{n=0}^{N} \left\| x_n - x_n^* \right\| , \\ \text{s.t.} \quad & x_{n+1} = x_n + \frac{h}{2} \left(f\left(x_n, t_n; a\right) + f\left(x_{n+1}, t_{n+1}; a\right) \right), \end{aligned} \tag{22}$$

where $\|\cdot\|$ could be $\mathcal{L}^1$ or $\mathcal{L}^2$ norm.

7. Finite Difference Methods of Continuous Time DICE

In this section, we apply the ideas in the previous section to our continuous-time IAM.

7.1. Explicit DICE-CJL Model

We use the explicit Euler finite difference rule to discretize the continuous time model with any time interval h. First, the total discounted utility over the first 600 years is

$$\sum_{n=0}^{N-1} e^{-\rho n h} u(c_n, l_n) h,$$

where labor supply $l_n = l(nh)$ from the function (10) for $n = 0, 1, \ldots, N-1$.

Second, from the explicit Euler finite difference rule for the continuous time differential equation of capital (15), the next-stage capital is

$$k_{n+1} = \left(1 - \delta h\right) k_n + \left(\mathcal{Y}\left(k_n, T_n^{\mathrm{AT}}, \mu_n, nh\right) - c_n\right) h, \tag{23}$$

where the function $\mathcal{Y}$ is defined in (14).

Third, from the explicit Euler finite difference rule for the continuous time climate system (16) and (19), the next-stage climate state becomes

$$\begin{aligned}\mathbf{M}_{n+1} &= \mathbf{M}_n + \left[\mathbf{\Phi}^{\mathrm{M}}\mathbf{M}_n + \left(\mathcal{E}\left(k_n, \mu_n, nh\right), 0, 0\right)^{\top}\right] h, \\ \mathbf{T}_{n+1} &= \mathbf{T}_n + \left[\mathbf{\Phi}^{\mathrm{T}}\mathbf{T}_n + \left(\xi_1 \mathcal{F}\left(M_n^{\mathrm{AT}}, nh\right), 0\right)^{\top}\right] h.\end{aligned}$$

Therefore, the discretized model with the explicit Euler finite difference method becomes

$$\begin{aligned}\max_{c_n,\mu_n} \quad & \sum_{n=0}^{N-1} e^{-\rho nh} u(c_n, l_n) h + e^{-\rho hN} \hat{V}\left(k_N, \mathbf{M}_N, \mathbf{T}_N\right), \\ \text{s.t.} \quad & k_{n+1} = k_n + \left(\mathcal{Y}\left(k_n, T_n^{\mathrm{AT}}, \mu_n, nh\right) - c_n - \delta k_n\right) h, \\ & \mathbf{M}_{n+1} = \mathbf{M}_n + \left[\mathbf{\Phi}^{\mathrm{M}}\mathbf{M}_n + \left(\mathcal{E}\left(k_n, \mu_n, nh\right), 0, 0\right)^{\top}\right] h, \\ & \mathbf{T}_{n+1} = \mathbf{T}_n + \left[\mathbf{\Phi}^{\mathrm{T}}\mathbf{T}_n + \left(\xi_1 \mathcal{F}\left(M_n^{\mathrm{AT}}, nh\right), 0\right)^{\top}\right] h.\end{aligned} \tag{24}$$

7.2. Trapezoidal DICE-CJL Model

We use the trapezoidal rule (the Crank-Nicholson method) to discretize the continuous time model with any time interval h. First, we use the trapezoidal rule for the integration in the continuous time model (21) to estimate the total discounted utility over the first 600 years, which is

$$\sum_{n=0}^{N} e^{-\rho nh} w_n u(c_n, l_n) h$$

where w_n are the weights with $w_n = 1$ for $n = 1, \ldots, N-1$ and $w_0 = w_N = 0.5$, and labor supply $l_n = l(nh)$ from the function (10) for $n = 0, 1, \ldots, N-1$.

Second, from the Crank-Nicholson method for the continuous time differential equation of capital (15), the next-stage capital satisfies

$$\begin{aligned} k_{n+1} &= k_n + \left[\mathcal{Y}\left(k_{n+1}, T_{n+1}^{\mathrm{AT}}, \mu_{n+1}, (n+1)h\right) - c_{n+1} - \delta k_{n+1}\right] \frac{h}{2} + \\ & \quad \left[\mathcal{Y}\left(k_n, T_n^{\mathrm{AT}}, \mu_n, nh\right) - c_n - \delta k_n\right] \frac{h}{2}, \end{aligned} \tag{25}$$

where the function $\mathcal{Y}$ is defined in (14).

Third, from the Crank-Nicholson method for the continuous time climate system (16) and (19), the next-stage climate state satisfies

$$\begin{aligned} \mathbf{M}_{n+1} &= \mathbf{M}_n + \left[\mathbf{\Phi}^{\mathrm{M}}\mathbf{M}_{n+1} + \left(\mathcal{E}\left(k_{n+1}, \mu_{n+1}, (n+1)h\right), 0, 0\right)^{\top}\right] \frac{h}{2} + \\ & \quad \left[\mathbf{\Phi}^{\mathrm{M}}\mathbf{M}_n + \left(\mathcal{E}\left(k_n, \mu_n, nh\right), 0, 0\right)^{\top}\right] \frac{h}{2}, \end{aligned} \tag{26}$$

$$\begin{aligned} \mathbf{T}_{n+1} &= \mathbf{T}_n + \left[\mathbf{\Phi}^{\mathrm{T}}\mathbf{T}_{n+1} + \left(\xi_1 \mathcal{F}\left(M_{n+1}^{\mathrm{AT}}, (n+1)h\right), 0\right)^{\top}\right] \frac{h}{2} + \\ & \quad \left[\mathbf{\Phi}^{\mathrm{T}}\mathbf{T}_n + \left(\xi_1 \mathcal{F}\left(M_n^{\mathrm{AT}}, nh\right), 0\right)^{\top}\right] \frac{h}{2}. \end{aligned} \tag{27}$$

Therefore, the discretized model with the trapezoidal rule becomes

$$\max_{c_n, \mu_n} \quad \sum_{n=0}^{N} e^{-\rho n h} w_n u(c_n, l_n) h + e^{-\rho N h} \hat{V}\left(k_N, \mathbf{M}_N, \mathbf{T}_N\right), \tag{28}$$

subject to the constraints (25), (26) and (27).

7.3. Detrended Finite Differences

Since the population is growing, the productivity factor is increasing and the adjusted cost for backstop is decreasing, the capital path will be explosive. This causes systematic bias in the standard finite difference methods we discussed above. Curvature in the solution is a source of error when one is using piecewise linear approximations.

Next, we transform the differential system to detrend the capital stock path, and apply a finite difference method to the detrended variable. This path is bounded with less curvature, and should be better for the finite difference method. We transform $k(t)$ to its detrended variable $e^{-\lambda t}k(t)$ where λ is an estimated parameter such that $e^{-\lambda t}k(t)$ has less curvature. One typical choice of λ is $\ln\left(k_N/k_0\right)/(Nh)$ by giving an estimated terminal capital k_N.

From (15), we have

$$\begin{aligned}\frac{d}{dt}\left(e^{-\lambda t}k(t)\right) &= e^{-\lambda t}\left(-\lambda k(t)+\frac{d}{dt}\left(k(t)\right)\right)\\ &= e^{-\lambda t}\left(\mathcal{Y}\left(k_t,T_t^{\mathrm{AT}},\mu_t,t\right)-c_t-\left(\lambda+\delta\right)k(t)\right).\end{aligned}$$

By discretizing the new differential equation using the explicit finite difference formula for the capital, we have

$$e^{-\lambda(n+1)h}k_{n+1}-e^{-\lambda nh}k_n=he^{-\lambda nh}\left(\mathcal{Y}\left(k_n,T_n^{\mathrm{AT}},\mu_n,nh\right)-c_n-\left(\lambda+\delta\right)k_n\right),$$

which implies that

$$e^{-\lambda h}k_{n+1}=\left(1-\left(\lambda+\delta\right)h\right)k_n+h\left(\mathcal{Y}\left(k_n,T_n^{\mathrm{AT}},\mu_n,nh\right)-c_n\right). \tag{29}$$

Similarly, if we use the Crank-Nicholson method to discretize the new differential equation of the detrended capital, we have the detrended implicit finite difference method:

$$\begin{aligned}e^{-\lambda h}k_{n+1} &= k_n+\left[\mathcal{Y}\left(k_{n+1},T_{n+1}^{\mathrm{AT}},\mu_{n+1},(n+1)h\right)-c_{n+1}-\left(\lambda+\delta\right)k_{n+1}\right]\frac{h}{2}\\ &\quad+\left[\mathcal{Y}\left(k_n,T_n^{\mathrm{AT}},\mu_n,nh\right)-c_n-\left(\lambda+\delta\right)k_n\right]\frac{h}{2}.\end{aligned} \tag{30}$$

7.4. Richardson Extrapolation

The true model is a continuous-time model. We hope that the solutions converge to the continuous-time solution as we reduce the time period. Also, our optimizer may not give a good optimal solution with a very small time step. For both reasons, we want to check if our discrete time solutions converge to a common limit.

Richardson extrapolation (Richardson and Gaunt, 1927) is a standard way to check if our solutions are consistent with convergence. We apply Richardson extrapolation to our solutions and find that they are consistent with convergence. Let $x^*_{t,h}$ be the optimal solution at time t of an ODE with h as the time interval. The 3-point Richardson extrapolation of $x^*_{t,h}$, $x^*_{t,h/2}$, and $x^*_{t,h/4}$ is defined as

$$x^{\mathrm{R}}_{t,h/4} = \frac{1}{3}\left(8x^*_{t,h/4} - 6x^*_{t,h/2} + x^*_{t,h}\right).$$

From numerical analysis, we know that the 3-point Richardson extrapolation has less errors than $x^*_{t,h/4}$.

8. Calibration

In the equations of the carbon cycle and temperature systems, there are 6 parameters, ϕ_{12}, ϕ_{23}, ξ_1, ξ_3, ξ_4 and η, that we should calibrate. At first, we generate the Business-As-Usual (BAU) paths of carbon cycle and temperature, which are optimal solutions of the integrated models separating the climate part and the economic part by fixing the emission control rate to be 0 at any time. We denote $\mathbf{M}_{\mathrm{D}}(t)$ and $\mathbf{T}_{\mathrm{D}}(t)$, respectively, as the carbon cycle and the temperature paths of DICE2007, and let $E_{\mathrm{D}}(t)$ be the path of the carbon emission rates. Then we find the parameters so that the paths of the carbon cycle and the temperature of the continuous time model match the BAU paths of DICE2007, by assuming that both models have the same path of the carbon emission rates.

If we choose the $\mathcal{L}^2$ norm in the equation (22), then the calibration model becomes

$$\begin{aligned}\min_{\phi_{12},\phi_{23},\xi_1,\xi_3,\xi_4,\eta} \quad & \sum_{n=1}^{N}\left\{\left(\frac{M_n^{\mathrm{AT}}-M_{\mathrm{D}}^{\mathrm{AT}}(t_n)}{M_{\mathrm{D}}^{\mathrm{AT}}(t_n)}\right)^2+\left(\frac{M_n^{\mathrm{UP}}-M_{\mathrm{D}}^{\mathrm{UP}}(t_n)}{M_{\mathrm{D}}^{\mathrm{UP}}(t_n)}\right)^2+\right.\\ & \left(\frac{M_n^{\mathrm{LO}}-M_{\mathrm{D}}^{\mathrm{LO}}(t_n)}{M_{\mathrm{D}}^{\mathrm{LO}}(t_n)}\right)^2+\left(\frac{T_n^{\mathrm{AT}}-T_{\mathrm{D}}^{\mathrm{AT}}(t_n)}{T_{\mathrm{D}}^{\mathrm{AT}}(t_n)}\right)^2+\\ & \left.\left(\frac{T_n^{\mathrm{LO}}-T_{\mathrm{D}}^{\mathrm{LO}}(t_n)}{T_{\mathrm{D}}^{\mathrm{LO}}(t_n)}\right)^2\right\},\end{aligned}$$

where $\mathbf{M}_n$ and $\mathbf{T}_n$ are the sequences generated from the equality constraints (26) and (27) by assuming $\mathcal{E}(k_n,\mu_n,nh)=E_{\mathrm{D}}(t_n)$.

If we choose the $\mathcal{L}^1$ norm in the equation (22), then the calibration model becomes

$$
\begin{aligned}
\min_{\phi_{12},\phi_{23},\xi_1,\xi_3,\xi_4,\eta} \quad & \sum_{n=1}^{N}\Big\{\epsilon_n^{\mathrm{AT}+}+\epsilon_n^{\mathrm{AT}-}+\epsilon_n^{\mathrm{UP}+}+\epsilon_n^{\mathrm{UP}-}+\epsilon_n^{\mathrm{LO}+}+\epsilon_n^{\mathrm{LO}-}+ \\
& \tau_n^{\mathrm{AT}+}+\tau_n^{\mathrm{AT}-}+\tau_n^{\mathrm{LO}+}+\tau_n^{\mathrm{LO}-}\Big\} \\
\text{s.t.} \quad & M_n^{\mathrm{AT}}-M_{\mathrm{D}}^{\mathrm{AT}}(t_n)=M_{\mathrm{D}}^{\mathrm{AT}}(t_n)\left(\epsilon_n^{\mathrm{AT}+}-\epsilon_n^{\mathrm{AT}-}\right), \\
& M_n^{\mathrm{UP}}-M_{\mathrm{D}}^{\mathrm{UP}}(t_n)=M_{\mathrm{D}}^{\mathrm{UP}}(t_n)\left(\epsilon_n^{\mathrm{UP}+}-\epsilon_n^{\mathrm{UP}-}\right), \\
& M_n^{\mathrm{LO}}-M_{\mathrm{D}}^{\mathrm{LO}}(t_n)=M_{\mathrm{D}}^{\mathrm{LO}}(t_n)\left(\epsilon_n^{\mathrm{LO}+}-\epsilon_n^{\mathrm{LO}-}\right), \\
& T_n^{\mathrm{AT}}-T_{\mathrm{D}}^{\mathrm{AT}}(t_n)=T_{\mathrm{D}}^{\mathrm{AT}}(t_n)\left(\tau_n^{\mathrm{AT}+}-\tau_n^{\mathrm{AT}-}\right), \\
& T_n^{\mathrm{LO}}-T_{\mathrm{D}}^{\mathrm{LO}}(t_n)=T_{\mathrm{D}}^{\mathrm{LO}}(t_n)\left(\tau_n^{\mathrm{LO}+}-\tau_n^{\mathrm{LO}-}\right), \\
& \epsilon_n^{\mathrm{AT}+},\epsilon_n^{\mathrm{AT}-},\epsilon_n^{\mathrm{UP}+},\epsilon_n^{\mathrm{UP}-},\epsilon_n^{\mathrm{LO}+},\epsilon_n^{\mathrm{LO}-},\tau_n^{\mathrm{AT}+},\tau_n^{\mathrm{AT}-},\tau_n^{\mathrm{LO}+},\tau_n^{\mathrm{LO}-}>0.
\end{aligned}
$$

We could also choose a weighted $\mathcal{L}^2$ norm, for example,

$$
\begin{aligned}
\min_{\phi_{12},\phi_{23},\xi_1,\xi_3,\xi_4,\eta} \quad & \sum_{n=1}^{N}\left\{\left(\frac{M_n^{\mathrm{AT}}-M_{\mathrm{D}}^{\mathrm{AT}}(t_n)}{M_{\mathrm{D}}^{\mathrm{AT}}(t_n)}\right)^2+\left(\frac{M_n^{\mathrm{UP}}-M_{\mathrm{D}}^{\mathrm{UP}}(t_n)}{M_{\mathrm{D}}^{\mathrm{UP}}(t_n)}\right)^2+\right. \\
& \left(\frac{M_n^{\mathrm{LO}}-M_{\mathrm{D}}^{\mathrm{LO}}(t_n)}{M_{\mathrm{D}}^{\mathrm{LO}}(t_n)}\right)^2+100\left(\frac{T_n^{\mathrm{AT}}-T_{\mathrm{D}}^{\mathrm{AT}}(t_n)}{T_{\mathrm{D}}^{\mathrm{AT}}(t_n)}\right)^2+ \\
& \left.\left(\frac{T_n^{\mathrm{LO}}-T_{\mathrm{D}}^{\mathrm{LO}}(t_n)}{T_{\mathrm{D}}^{\mathrm{LO}}(t_n)}\right)^2\right\}.
\end{aligned}
$$

Using the above three models to calibrate over 500 years, we have the calibrated parameters listed in Table 1.

9. Numerical Results

We apply the explicit Euler method to discretize the climate and economic dynamic system and then solve the optimization problem (24) with

Table 1: Calibration Parameters

	ϕ_{12}	ϕ_{23}	ξ_1	ξ_3	ξ_4	η
L2	0.0190837	0.005403	0.04217	0.2609	0.0048	3.4145
L1	0.0193931	0.005623	0.04407	0.2178	0.00493	2.8324
weighted L2	0.0192281	0.005407	0.02777	0.4453	0.00479	5.8551

various time steps. We also apply the trapezoid rule and then solve the optimization problem (28) with various time steps. Numerically, if the time step h is large, then the numerical errors may become too large so that the numerical solutions may not be trusted. We use various time intervals to see the difference of solutions of DICE-CJL with these different time steps h. We use CONOPT in the GAMS environment (McCarl, 2011) to solve DICE-CJL with different time period lengths in our GAMS code.

9.1. Starting Point Strategy

It will become more challenging to solve DICE-CJL with a smaller time interval h, because the number of variables and constraints and nonzero elements in the system will increase proportionally.

A good initial guess of solutions will be very helpful for an optimizer to solve such a large-scale optimization problem. In our examples, one good initial guess is the linear interpolation of optimal solutions of DICE-CJL with a larger time interval. For example, we could use the linear interpolation of optimal solutions of the 2-year DICE-CJL as the initial guess for the annual DICE-CJL problem.

Table 2: Running Times of DICE-CJL with Finite Difference Methods

Step Size h	Explicit DICE-CJL	Trapezoidal DICE-CJL
10 years	-	1.1 seconds
5 years	-	6.3 seconds
2 years	6.8 seconds	22 seconds
1 year	0.9 seconds	3.1 seconds
6 months	4.6 seconds	13 seconds
3 months	14 seconds	58 seconds
1 month	133 seconds	241 seconds
2 weeks	334 seconds	-
1 week	1979 seconds	-

9.2. Running times of DICE-CJL

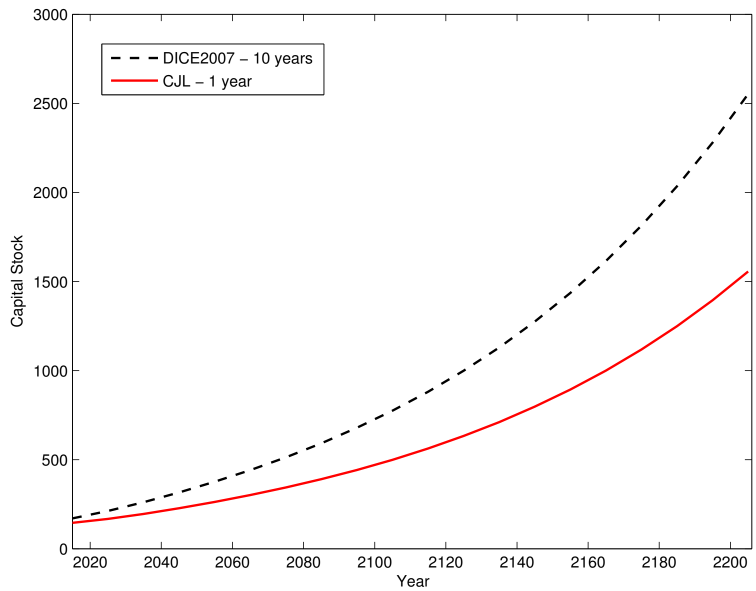
Table 2 lists the running times of DICE-CJL with various time intervals in the GAMS environment on a single-processor Mac. For the explicit model, the one-year version takes less than 1 second and the weekly version only take 33 minutes. And for the trapezoidal model it takes only 4 minutes to solve the monthly version. It seems surprising that the 2-year version (or larger time step for the trapezoidal model) takes more time than 1-year and even 6-month versions. This happens because we do not use the starting strategy for the 2-year version (or larger time step for the trapezoidal model), but for other smaller time steps we use the starting strategy.

9.3. State Paths of DICE-CJL

Figures 1-3 plot the optimal paths of the capital stock, the atmospheric carbon stock, and surface temperature of the explicit DICE-CJL model (24) with 1-year time steps. For other smaller time steps and the solutions of the trapezoidal DICE-CJL model (28) with various time steps from 10 years to

1 month, the paths are close to the 1-year solutions of the explicit model, so we omit them here. We see that the line of DICE2007 using Nordhaus's program diverges strongly from our solutions in each figure.

Figure 1: Capital Stock



9.4. Carbon Tax Results in DICE-CJL

One application of DICE2007 was to compute the optimal carbon tax; see Nordhaus (2007, 2008) and Interagency Working Group on Social Cost of Carbon (2010). Figure 4 displays the optimal path of the carbon tax for three numerical approaches. DICE2007 is the path produced by DICE2007. The dotted line is the path when we use DICE-CJL with the same diffusivity rates as used in DICE2007, and is the path whether we use the non-causal specification for warming in DICE2007 or use the DICE-CJL specification.

Figure 2: Carbon Concentration in the Atmosphere

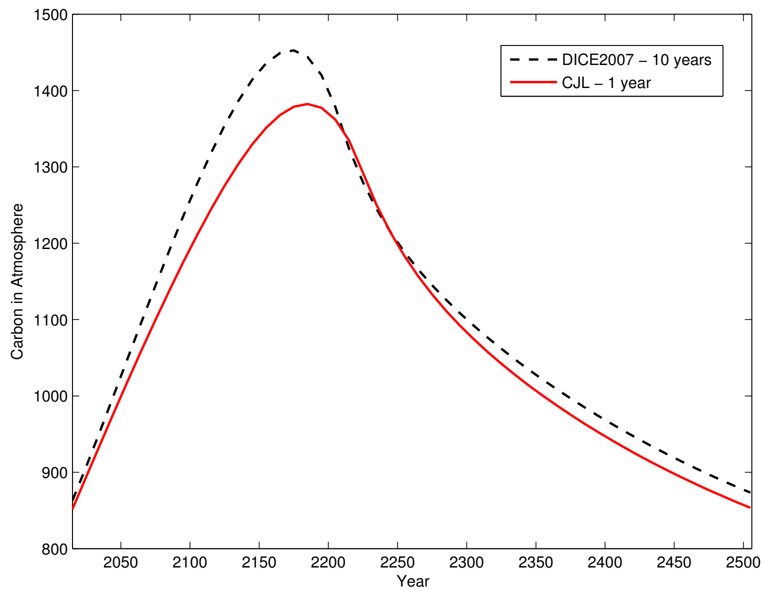


Figure 3: Surface Temperature

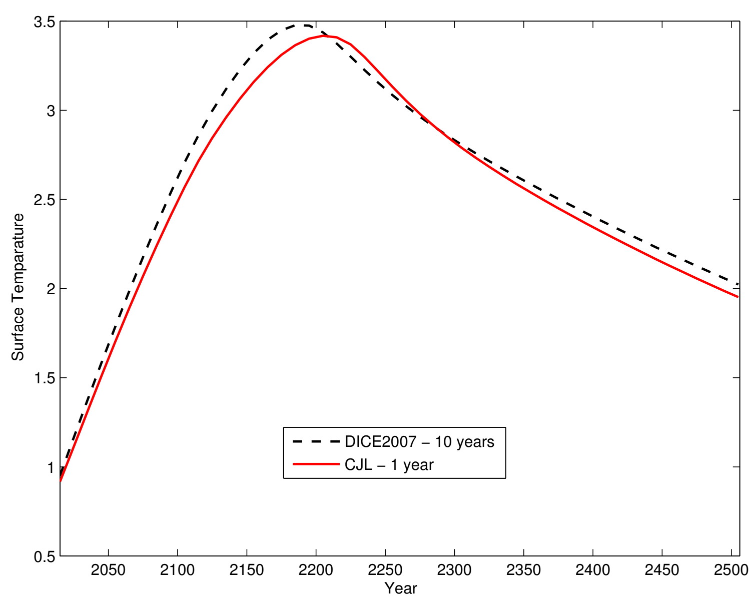
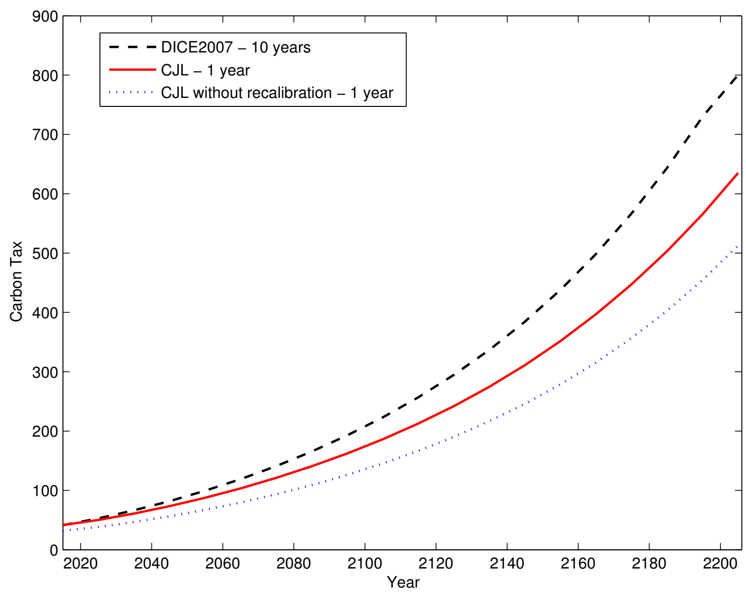


Figure 4: Carbon Tax



The solid red line is DICE-CJL with one-year time steps where the diffusivity parameters are chosen so that the continuous-time climate system matches the results for the climate system in DICE2007. The results are striking. The ten-year time period in DICE2007 consistently produces much higher path for the carbon tax than other models with shorter time periods. This example shows clearly why the continuous-time approach is the only proper foundation for these models, and that finite-difference approximations must be based on numerical methods for differential equations.

10. Error Analysis of DICE-CJL Model

We next analyze and compare the errors of alternative methods.

10.1. Explicit DICE-CJL Model

Figures 5-7 show that our solutions from the explicit finite difference method are good. The vertical axis in each figure is

$$\log_{10}\left(\left|\frac{x^*_{t,h} - x^{\mathrm{R}}_t}{x^{\mathrm{R}}_t}\right| + 10^{-7}\right),$$

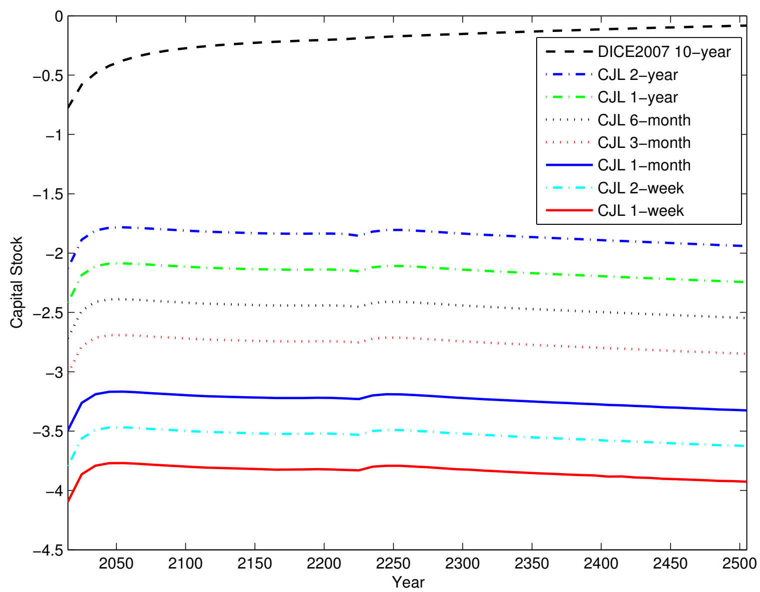
where $x^*_{t,h}$ is the optimal solution at time t of the explicit DICE-CJL model with h being the time interval, x^{R}_t is the 3-point Richardson extrapolation of the optimal solution with 1 month, 0.5 month and 1 week (equals to 1/4 month by our assumption) for the explicit model, i.e.,

$$x^{\mathrm{R}}_t = \frac{1}{3}\left(8x^*_{t,1/48} - 6x^*_{t,1/24} + x^*_{t,1/12}\right).$$

Figures 5-7 plot the relative difference over time intervals of optimal paths of capital, carbon concentration in the atmosphere, and surface temperature respectively. We omit the figures of the other three state variables, as they show a similar pattern. Each line represents the difference between the solution for a time step and the Richardson extrapolant.

From the figures, we see that Nordhaus' solution has a $O(1)$ error in the optimal capital, and $O(10^{-1})$ errors in the atmospheric carbon and surface temperature, and even for the first 100 years the errors are still around $O(10^{-1})$ or even worse. However, the explicit DICE-CJL model has $O(10^{-2})$ errors for each state variable when the time step is 2 years, and then it decays down to $O(10^{-4})$ when the time step drops to 1 week almost uniformly along the time path and along the time step. Moreover, since 1 week (equals to 1/4 month by assumption) is equal to 1/96 of 2 years, and the error of 1-week explicit DICE-CJL is also about 1/96 of the error of 2-year explicit DICE-CJL, we see that the error of each state variable is decreasing linearly as we reduce the time step in the explicit DICE-CJL model.

Figure 5: Relative Errors of Capital from Explicit DICE-CJL



10.2. Trapezoidal DICE-CJL Model

Figures 8-10 show that our solutions from the trapezoidal rule are more accurate than the solutions given by the explicit DICE-CJL model. The vertical axis in each figure is

$$\log_{10}\left(\left|\frac{x^*_{t,h}-x^{\mathrm{R}}_t}{x^{\mathrm{R}}_t}\right|+10^{-7}\right),$$

Figure 6: Relative Errors of Atmospheric Carbon from Explicit DICE-CJL

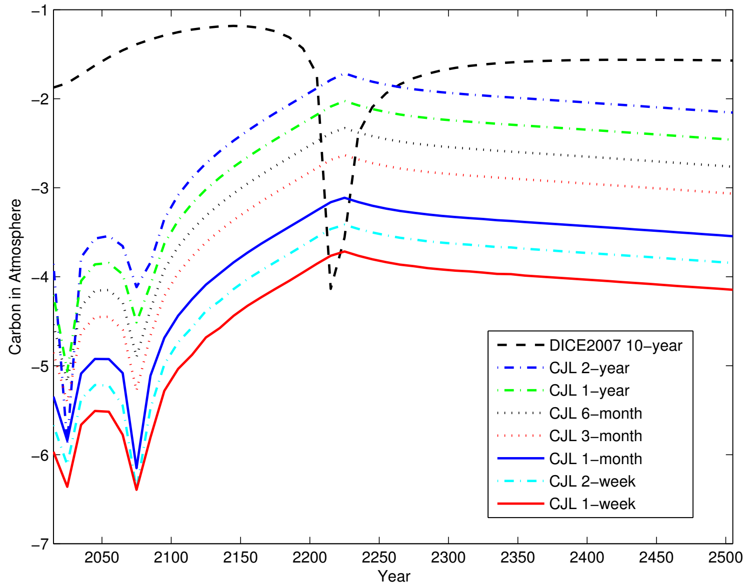
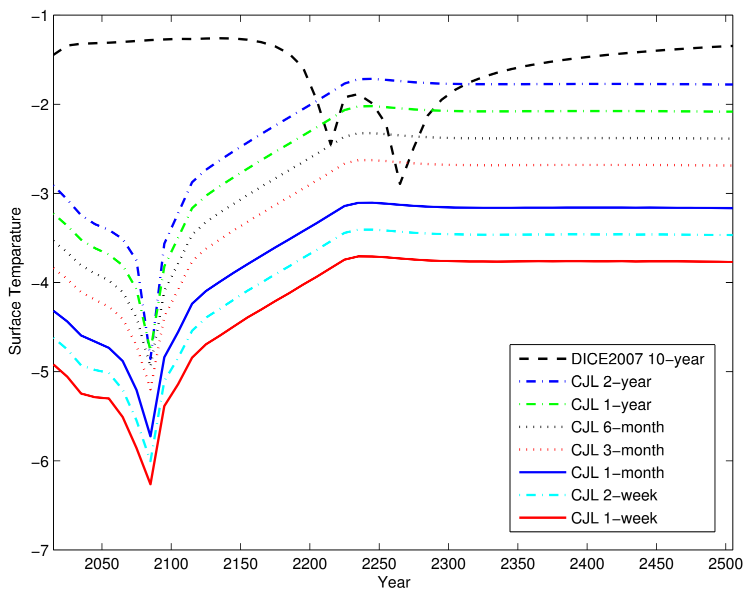


Figure 7: Relative Errors of Surface Temperature from Explicit DICE-CJL



where $x_{t,h}^*$ is the optimal solution at time t of the trapezoidal DICE-CJL model with h as the time interval, x_t^{R} is the 3-point Richardson extrapolation of optimal solution with 1 year, 0.5 year and 0.25 year for the explicit model, i.e.,

$$x_t^{\mathrm{R}} = \frac{1}{3}\left(8x_{t,1/4}^* - 6x_{t,1/2}^* + x_{t,1}^*\right).$$

From the figures, the trapezoidal DICE-CJL model has $O(10^{-2}) \sim O(10^{-3})$ errors for each state variable when the time step is 10 years which has almost the same accuracy with the 1-year explicit DICE-CJL model. When the time step is reduced to 0.25 year, the error has already converged to $O(10^{-5}) \sim O(10^{-6})$. The 1-month trapezoidal DICE-CJL results also show the convergence. Since 0.25 year is equal to 1/40 of 10 years, and the error of 0.25-year trapezoidal DICE-CJL is about 1/1000 of the error of 10-year trapezoidal DICE-CJL, we see that the error of each state variable decays about quadratically as we reduce the time step in the trapezoidal DICE-CJL model.

10.3. Detrended Finite Difference

From Figure 1, we see that the capital path has a high curvature, so we use the detrended finite difference method by detrending the capital path $k(t)$ to $e^{-\lambda t}k(t)$, where we choose $\lambda = \ln\left(70000/k_0\right)/600 = 0.0104$. Figure 11 gives the relative errors of optimal capital under logarithmic scale from the detrended finite difference method using 10-year step size. We see that the detrended explicit finite difference method using (29) (the dotted line) has less errors than the explicit DICE-CJL without detrending (the dot-dashed line). Moreover, the detrended trapezoidal finite difference method using (29) (the dashed line) improves half a digit accuracy than the trapezoidal DICE-CJL without detrending (the dot-dashed line). For other climate states, the detrended DICE-CJL has solutions close to those without detrending because we just do detrending on the economic side.

Figure 12 verifies that the detrended capital has less curvature: it ranges from 125 to 200 while the optimal capital has a much wider range along the time.

11. Conclusion

Both the climate and economy are continuous-time systems. Climate system modelers have always based their work on continuous-time models, but economists have used discrete-time models with long time periods. Using DICE as an example, we show that continuous-time formulations of IAM models are natural, that many reliable methods from numerical analysis are available to solve such models, and that the choice of time step and finite-difference method can have economically significant effects on the answers to basic questions in economic policy.

The DICE example is a simple one which can be solved by a variety of numerical methods that are reliable when one takes very short time steps. However, as we move to multisector and multiregional models that significantly increase the dimensionality of the system of ordinary differential equations, it will become increasingly important to use efficient finite difference methods that allow for time steps of moderate size without losing accuracy. The value of the arguments mentioned in this paper will be even higher in these more realistic models.

[1] Cai, Y. (2009). *Dynamic Programming and Its Application in Economics and Finance*. PhD thesis, Stanford University.

[2] Cai, Y., and K.L. Judd (2010). Stable and efficient computational methods for dynamic programming. *Journal of the European Economic Association*, Vol. 8, No. 2-3, 626–634.

[3] Cai, Y., K.L. Judd and T.S. Lontzek (2012a). Open science is necessary. *Nature Climate Change*, Vol. 2, Issue 5, 299–299.

[4] Cai, Y., K.L. Judd and T.S. Lontzek (2012b). DSICE: A dynamic stochastic integrated model of climate and economy. Working paper.

[5] Interagency Working Group on Social Cost of Carbon (2010). *Social Cost of Carbon for Regulatory Impact Analysis under Executive Order 12866*. United States Government. `http://www.whitehouse.gov/sites/default/files/omb/inforeg/for-agencies/Social-Cost-of-Carbon-for-RIA.pdf`

[6] Iserles, A. (1996). *A First Course in the Numerical Analysis of Differential Equations*. Cambridge University Press.

[7] Judd, K.L. (1998). *Numerical Methods in Economics*. The MIT Press.

[8] Lontzek, T.S., Y. Cai and K.L. Judd (2012). Tipping points in a dynamic stochastic IAM. Working paper.

[9] McCarl, B., et al. (2011). McCarl GAMS User Guide. GAMS Development Corporation.

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

[11] Nordhaus, W. (2011). Integrated economic and climate modeling. Cowles Foundation discussion paper, No. 1839, Cowles Foundation For Research in Economics , Yale University.

[12] Richardson, L.F., and Gaunt, J.A. (1927). The deferred approach to the limit. Philosophical Transactions of the Royal Society of London, Series A 226 (636-646): 299–349.

Figure 8: Relative Errors of Capital from Trapezoidal DICE-CJL

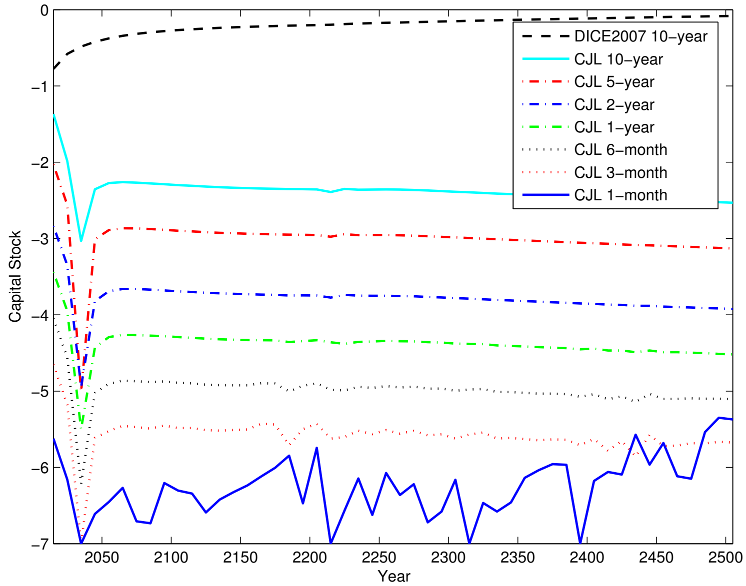


Figure 9: Relative Errors of Atmospheric Carbon from Trapezoidal DICE-CJL

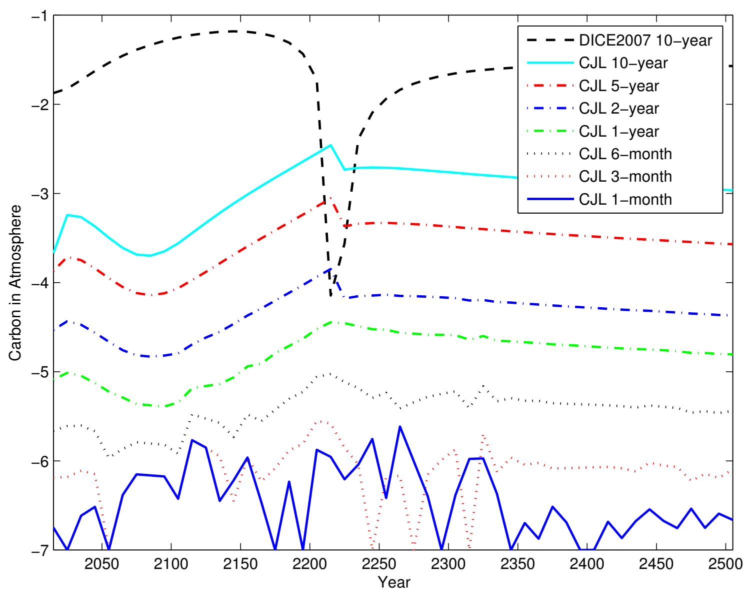


Figure 10: Relative Errors of Surface Temperature from Trapezoidal DICE-CJL

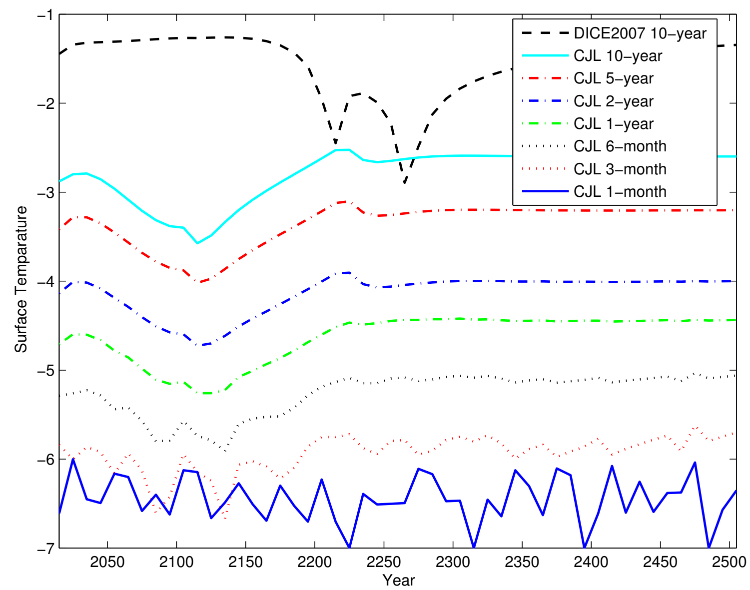


Figure 11: Relative Errors of Capital with Detrended Finite Difference Method

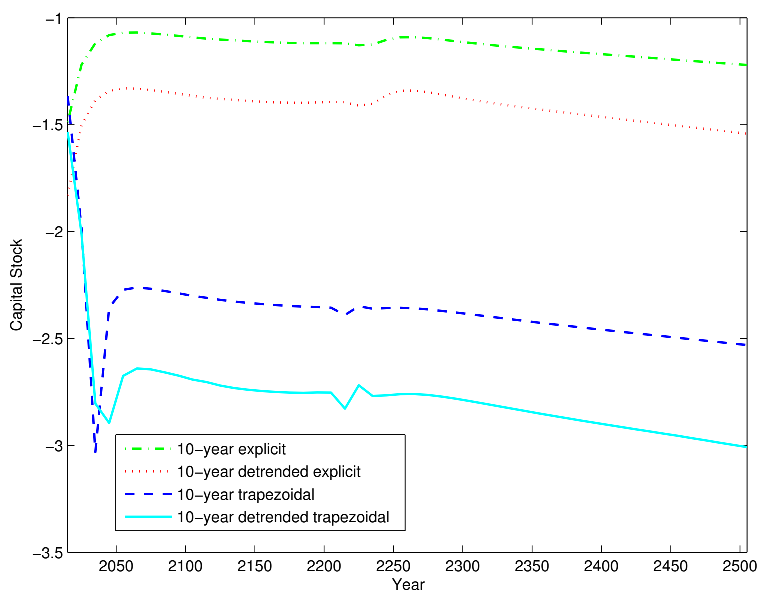


Figure 12: Detrended Capital