Supplement to "A simple but powerful simulated certainty equivalent approximation method for dynamic stochastic problems"

(*Quantitative Economics*, Vol. 14, No. 2, May 2023, 651–687)

Yongyang Cai
Department of Agricultural, Environmental, and Development Economics, The Ohio State University

Kenneth L. Judd
Hoover Institution, Stanford University

A.1. Accuracy measure

Let the number of state variables be n (so $\mathbf{g}_t$ is a length-n vector of state transition equations), and the number of feasibility constraints be k (so $\mathbf{f}_t$ is a length-k vector of feasibility constraints). For convenience, we assume that all state variables are continuous. Let $\beta^t \boldsymbol{\lambda}_t$ be a $n \times 1$ column vector representing the Lagrange multipliers of the vector of state transition equations $\mathbf{x}_{t+1} = \mathbf{g}_t(\mathbf{x}_t, \mathbf{a}_t, \boldsymbol{\epsilon}_{t+1})$, and $\beta^t \boldsymbol{\nu}_t$ be a $k \times 1$ column vector representing the Lagrange multipliers of the vector of feasibility constraints $\mathbf{f}_t(\mathbf{x}_t, \mathbf{a}_t) \geq 0$. The system of associated Euler equations at period s for the stochastic model (1) is[34]

$$\boldsymbol{\lambda}_s = \beta \mathbb{E}_s \left\{ \nabla_{\mathbf{x}_{s+1}} \left[\mathbf{g}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}, \boldsymbol{\epsilon}_{s+2}) \boldsymbol{\lambda}_{s+1} + \mathbf{f}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}) \boldsymbol{\nu}_{s+1} \right] \right\}, \tag{A.1}$$

where $\mathbb{E}_s$ is the expectation operator conditioned on time-s information, $\nabla_{\mathbf{x}_{s+1}}$ is the gradient operator over $\mathbf{x}_{s+1}$,

$$\nabla_{\mathbf{x}} \mathbf{g} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_n}{\partial x_n} \end{pmatrix}$$

is an $n \times n$ matrix,

$$\nabla_{\mathbf{x}} \mathbf{f} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n} & \cdots & \frac{\partial f_k}{\partial x_n} \end{pmatrix}$$

Yongyang Cai: cai.619@osu.edu
Kenneth L. Judd: kennethjudd@mac.com

[34] These Euler equations are only valid for problems with exogenous probabilities in state transition. If a model uses an endogenous probability, the Euler equation should be adjusted accordingly. See Section A.6 for one example.

 https://doi.org/10.3982/QE1835

is an $n \times k$ matrix.[35] The normalized Euler error is defined as

$$\left\| \frac{\beta}{\boldsymbol{\lambda}_s} \mathbb{E}_s \left\{ \nabla_{\mathbf{x}_{s+1}} \left[\mathbf{g}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}, \boldsymbol{\epsilon}_{s+2}) \boldsymbol{\lambda}_{s+1} + \mathbf{f}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}) \boldsymbol{\nu}_{s+1} \right] \right\} - \mathbf{1} \right\|, \tag{A.2}$$

where $\mathbf{1}$ is a vector of ones, the division operation is elementwise, and $\| \cdot \|$ is a norm operator. Note that the normalized Euler error is unit free, and thus more suitable for measuring errors than nonnormalized measures.

The common method to estimate Euler errors is to use approximate policy functions in (A.2). Approximating value or policy functions can be challenging if they are nonsmooth or have a high-dimensional state space. We can instead estimate the normalized Euler errors at the simulated solutions only. In the optimization step of Algorithm 1, we can obtain multipliers $\boldsymbol{\lambda}_s^i$ at no cost by solving the deterministic model (2), since these multipliers are explicitly given in the equations in Algorithm 2. Thus, we implement Algorithm 1 or 2 to generate simulated paths of $(\mathbf{x}_s^i, \mathbf{a}_s^i)$ and the associated multipliers $\boldsymbol{\lambda}_s^i$, and save them for accuracy measures.

Given the time-s information $(\mathbf{x}_s^i, \mathbf{a}_s^i, \boldsymbol{\lambda}_s^i)$, the next state in the original stochastic model (1), $\mathbf{g}_s(\mathbf{x}_s^i, \mathbf{a}_s^i, \boldsymbol{\epsilon}_{s+1})$, is random and dependent on $\boldsymbol{\epsilon}_{s+1} \in \Theta$, where Θ represents the set of all possible vectorial values of $\boldsymbol{\epsilon}_{s+1}$. To use the Euler equation (A.1), we have to estimate the next state $\mathbf{x}_{s+1}$ and its associated decision $\mathbf{a}_{s+1}$ as well as shadow prices $(\boldsymbol{\lambda}_{s+1}, \boldsymbol{\nu}_{s+1})$, all of which are random.[36] We approximate them by the solutions from the deterministic model (2) at their corresponding random states $\mathbf{x}_{s+1}$ at time $s+1$. For example, if $\boldsymbol{\epsilon}_{s+1}$ is a discrete random variable vector with a finite number of possible values $\{\boldsymbol{\epsilon}_{s+1}^{(j)}\}$, we compute $\mathbf{x}_{s+1}^{(j)} = \mathbf{g}_s(\mathbf{x}_s^i, \mathbf{a}_s^i, \boldsymbol{\epsilon}_{s+1}^{(j)})$ for every $\boldsymbol{\epsilon}_{s+1}^{(j)} \in \Theta$, then solve (2) with the starting time $s+1$ and starting state $\mathbf{x}_{s+1}^{(j)}$ to find the associated optimal decision $\mathbf{a}_{s+1}^{(j)}$ and their corresponding multipliers $(\boldsymbol{\lambda}_{s+1}^{(j)}, \boldsymbol{\nu}_{s+1}^{(j)})$. That is, we solve

$$\begin{aligned} \max_{\mathbf{a}_t \in \mathcal{D}_t(\mathbf{x}_t)} \quad & \sum_{t=s+1}^{s+\Delta_s} \beta^{t-s-1} u_t(\mathbf{x}_t, \mathbf{a}_t) + \beta^{\Delta_s} V_{s+\Delta_s+1}(\mathbf{x}_{s+\Delta_s+1}) \\ \text{s.t.} \quad & \mathbf{x}_{t+1} = \mathbf{g}_t(\mathbf{x}_t, \mathbf{a}_t, 0), \quad t = s+1, \ldots, s+\Delta_s, \\ & \mathbf{f}_t(\mathbf{x}_t, \mathbf{a}_t) \geq 0, \quad t = s+1, \ldots, s+\Delta_s, \end{aligned} \tag{A.3}$$

where $\mathbf{x}_{s+1}$ is given by $\mathbf{x}_{s+1}^{(j)}$. We then compute the conditional expectation in equation (A.1). When $\boldsymbol{\epsilon}_{s+1}$ is a continuous random variable vector, we can use its quadrature nodes $\boldsymbol{\epsilon}_{s+1}^{(j)}$, then implement quadrature rules to estimate the conditional expectation

[35] In a typical Euler equation, the Lagrange multipliers are often substituted by marginal utilities or some other expressions derived from Karush–Kuhn–Tucker conditions, but in many cases such substitutions are complicated. Instead we can directly obtain the multipliers from numerical optimization solvers. For example, the constrained optimization solver in Matlab, fmincon, has the command

[X,FVAL,EXITFLAG,OUTPUT,LAMBDA] = fmincon(FUN, X0,...)

to report the multipliers in the vector LAMBDA.

[36] $\mathbf{x}_{s+1}$ can be nonrandom conditional on time-s information.

in equation (A.1). Thus, we can approximately estimate the Euler error (A.2) at period s and state $\mathbf{x}_s^i$ as

$$\mathcal{E}_s^i = \left\| \frac{\beta}{\boldsymbol{\lambda}_s^i} \sum_j w_j \mathbb{E}\left\{ \nabla_{\mathbf{x}_{s+1}} \left[\mathbf{g}_{s+1}\left(\mathbf{x}_{s+1}^{(j)}, \mathbf{a}_{s+1}^{(j)}, \boldsymbol{\epsilon}_{s+2}\right) \boldsymbol{\lambda}_{s+1}^{(j)} + \mathbf{f}_{s+1}\left(\mathbf{x}_{s+1}^{(j)}, \mathbf{a}_{s+1}^{(j)}\right) \boldsymbol{\nu}_{s+1}^{(j)} \right] \right\} - \mathbf{1} \right\|, \quad \text{(A.4)}$$

where $\mathbf{x}_{s+1}^{(j)}$ are all possible states or quadrature nodes conditional on $\mathbf{x}_s^i$, w_j are the corresponding probabilities or quadrature weights, $\mathbf{a}_{s+1}^{(j)}$, $\boldsymbol{\lambda}_{s+1}^{(j)}$, and $\boldsymbol{\nu}_{s+1}^{(j)}$ are obtained from solving the next-period optimization model (A.3) with state $\mathbf{x}_{s+1}^{(j)}$, and $\mathbb{E}$ is over $\boldsymbol{\epsilon}_{s+2}$.

If we have m simulated paths, then we can compute the approximate $\mathcal{L}^\infty$ Euler error of the first T^* periods solutions for these simulated paths, defined as

$$\max_{0 \le s < T^*} \left(\max_{1 \le i \le m} \mathcal{E}_s^i \right), \quad \text{(A.5)}$$

where the norm operator in (A.4) for defining $\mathcal{E}_s^i$ is chosen to be the $\mathcal{L}^\infty$ norm, that is, the maximum absolute values of all the vector elements inside the norm operator. Similarly, we can compute the approximate $\mathcal{L}^1$ Euler error of the first T^* period's solutions among the simulated paths, defined as

$$\frac{1}{mT^*} \sum_{0 \le s < T^*, 1 \le i \le m} \mathcal{E}_s^i, \quad \text{(A.6)}$$

where the norm operator in (A.4) for defining $\mathcal{E}_s^i$ is chosen to be the $\mathcal{L}^1$ norm, that is, the average of the absolute values of all the vector elements inside the norm operator.

Note that if the simulated solutions at the first T^* periods do not contain extreme events, then a small approximate $\mathcal{L}^\infty$ or $\mathcal{L}^1$ Euler error in the first T^* periods cannot guarantee an accurate solution (see Jin and Judd (2002), Juillard and Villemot (2011)). This is shown in the climate tipping example in Section A.6, in which the Euler errors at the early periods with zero tipping probabilities are nearly zero, but the Euler errors at the periods with nonzero tipping probabilities are around 0.1–1%. That is, the solution at the early periods should be regarded to have two or three-digit accuracy, although their Euler errors are nearly zero. The reason is that our Euler error measure at period s depends on solutions at both periods s and $s+1$: if the solution at period $s+1$ has 0.1–1% errors, then with a small Euler error at period s, the solution at s would also have around 0.1–1% errors. This is similar to backward VFI with time-varying approximation domains: we may use a linear or quadratic function to approximate value functions and VFI can still converge, and the function approximation errors (and the Euler errors) at the initial periods could be very small if the approximation domains at the initial periods are narrow, but we cannot say the VFI solution at the initial periods is accurate, because of the potential large approximation errors at the later periods with wide approximation domains (see Cai (2019) for more discussion).

If there are m simulated paths with T^* periods, and J possible values of $\boldsymbol{\epsilon}_t$ (or J quadrature nodes for a continuous $\boldsymbol{\epsilon}_t$) in each period, then the accuracy measures need to solve $m \times T^* \times J$ optimization problems (A.3), so the accuracy measure method could

be time consuming. We can instead randomly choose some periods and some simulation paths (e.g., the worst or the best scenario) to measure accuracy, or choose some specific periods (e.g., those with extreme points or those around the time when the occasionally binding constraints become binding) and assess all simulated solutions at those periods.

Moreover, in most cases, the Euler error in the first period can already show if SCEQ can provide a solution with acceptable accuracy. Note that the Euler error in the first period can be directly computed by the solutions from Algorithm 1 or 2 without solving additional optimization problems, as the initial state is observed and unique. That is, starting from the initial state, Algorithm 1 or 2 have already provided m simulated states in the second period and their associated decisions and multipliers, so we can just estimate the expectation in (A.2) as the average across the simulated solutions for $s = 0$. In addition, in many cases the Euler error in the first period is the largest as it depends on the initial state, which could be an extreme point.

Furthermore, the accuracy measure process can be naturally parallelized across the simulated $\mathbf{x}_s^i$ for every s and i, and also its associated next-period nodes $\mathbf{x}_{s+1}^{(j)}$ for every j. That is, if we can have $m \times T^* \times J$ compute cores, we can let each compute core run only one optimization problem to obtain the values of $\mathbf{x}_{s+1}^{(j)}$, $\mathbf{a}_{s+1}^{(j)}$, $\boldsymbol{\lambda}_{s+1}^{(j)}$ and $\boldsymbol{\nu}_{s+1}^{(j)}$ for each $\mathbf{x}_s^i$ and j, and then collect them to compute the Euler errors. Thus, the accuracy measure method could be very fast with parallelism: the running time could be around the time spent to solve one optimization problem (A.3). For example, it took only seconds on a Mac Pro desktop computer to solve one optimization problem (A.3) in all of our examples, including the problems with 200 countries and $\Delta_s = 50$.

In addition, if we solve all different possible paths of $(\mathbf{x}_s^i, \mathbf{a}_s^i, \boldsymbol{\lambda}_s^i)$ using Algorithm 1 or 2, then this accuracy measure has almost no computational cost, as the expectation in (A.4) can be directly computed from the values in the different possible simulated paths.

A.2. Another illustrative example

Here, we solve the same optimal growth model described in Section 4.1, but we let $\beta = 0.96$, $\delta = 0.1$, and $u(c) = -c^{-1}$, in which there is no analytical solution and the certainty equivalent approximation has effect on the solution.We follow the exact same procedure of SCEQ as in Section 4.1 to obtain the new illustrative example's solution. The estimated normalized Euler error at the initial state is 1.9×10^{-4}, and the $\mathcal{L}^\infty$ Euler error over the 1000 simulated paths is 5.7×10^{-4}.

We also solve this simple problem via value function iteration (VFI) to obtain the optimal policy function for consumption, using the 7-node Gauss–Hermite quadrature rule to compute the conditional expectation of the value function in the Bellman equation (Bellman (1957)):

$$V(\mathbf{x}_t) = \max_{c_t} \quad u(c_t) + \beta \mathbb{E}_t\{V(\mathbf{x}_{t+1})\} \tag{A.7}$$

subject to the transition laws of the state variable vector $\mathbf{x}_t = (K_t, \theta_t)$.

To make sure that next period's θ is inside the approximation domain of θ for the value function for all Gauss–Hermite quadrature nodes, we have to choose it from the interval $[0.223, 4.482]$.[37] With this wide range for θ, we have to choose a wide range for K ($[0.04, 29.1]$) so that next period's capital will be inside the approximation domain of K for all possible values of θ.[38] That is, the two-dimensional approximation domain for the state variable vector (K, θ) is $[0.04, 29.1] \times [0.223, 4.482]$, much wider than the ranges of the SCEQ's simulated states ($[0.75, 1.34] \times [0.81, 1.25]$).

This wide approximation domain requires a very high-degree approximation. For this specific example, we can do a nonlinear change of variables so that the value function can be approximated with a lower degree approximation. Using $\ln(K)$ and $\ln(\theta)$ as the state variables, we approximate the value function as a degree-30 complete Chebyshev polynomial of $\ln(K)$ and $\ln(\theta)$ (see, e.g., Cai and Judd (2014)). We also approximate the corresponding consumption policy functions by a degree-30 complete Chebyshev polynomial of $\ln(K)$ and $\ln(\theta)$, denoted $\mathcal{C}^{\text{VFI}}(K, \theta)$. The normalized $\mathcal{L}^\infty$ Euler error of the solution $\mathcal{C}^{\text{VFI}}(K, \theta)$, defined as

$$\max_{K,\theta}\left|1-\beta\mathbb{E}\left[\frac{u'\left(\mathcal{C}^{\text{VFI}}(K_+, \theta_+)\right)}{u'\left(\mathcal{C}^{\text{VFI}}(K, \theta)\right)}\left(1-\delta+\theta_+ A\alpha K_+^{\alpha-1}\right)\,\middle|\,(K, \theta)\right]\right|$$

is 1.2×10^{-4}, where θ_+ is next period's productivity shock conditional on the current period's θ, and K_+ is next period's capital: $K_+ = (1-\delta)K + \theta A K^\alpha - \mathcal{C}^{\text{VFI}}(K, \theta)$. Thus, we see the VFI solution is accurate enough for checking the accuracy of our SCEQ solution.

For comparison with the VFI solution, we use the superscript "SCEQ" to denote the SCEQ solution. We then compare $c_t^{i,\text{SCEQ}}$ and $\mathcal{C}^{\text{VFI}}(K_t^{i,\text{SCEQ}}, A_t^{i,\text{SCEQ}})$ for all $1 \le i \le 1000$ and $t < 20$. We find the $\mathcal{L}^\infty$ relative error of the SCEQ solution, defined as

$$\max_{0\le t<20, 1\le i\le 1000}\frac{\left|c_t^{i,\text{SCEQ}}-\mathcal{C}^{\text{VFI}}\left(K_t^{i,\text{SCEQ}}, A_t^{i,\text{SCEQ}}\right)\right|}{\mathcal{C}^{\text{VFI}}\left(K_t^{i,\text{SCEQ}}, A_t^{i,\text{SCEQ}}\right)}$$

is 0.0016, and the $\mathcal{L}^1$ relative error, defined as

$$\frac{1}{20{,}000}\sum_{0\le t<20, 1\le i\le 1000}\frac{\left|c_t^{i,\text{SCEQ}}-\mathcal{C}^{\text{VFI}}\left(K_t^{i,\text{SCEQ}}, A_t^{i,\text{SCEQ}}\right)\right|}{\mathcal{C}^{\text{VFI}}\left(K_t^{i,\text{SCEQ}}, A_t^{i,\text{SCEQ}}\right)}$$

is 8.0×10^{-4}.[39] The relative errors are close to the estimated Euler errors of the SCEQ solution, so SCEQ's own accuracy measures with Euler errors are also good for checking solution accuracy.

[37] The range of $\ln(\theta)$ is proportional to $\sigma/(1-\rho)$, so if ρ or σ is larger, then the range has to be wider. Moreover, if we use more Gauss–Hermite quadrature nodes, then the range will also be wider.

[38] If we use a narrower range for K, then next period's capital could be binding at its bounds, which creates kinks when approximating the value function, so the solution may be inaccurate (particularly for those states near the bounds). Moreover, the kinks might prevent the value function iteration process from converging.

[39] We view the VFI optimal consumption policy to be the "true" solution as there is no analytical solution for the optimal growth problem.

A.3. A SIMPLE MULTICOUNTRY MODEL

Judd, Maliar, and Maliar (2012) and Maliar and Maliar (2015) uses a multicountry model similar to our Case 3 in Section 4.2.3 for testing their method, but their model is simpler: it does not have kinks, a wide-ranging state space, labor, or adjustment costs. That is, the social planner maximizes the aggregate welfare

$$\max_c \mathbb{E}\left(\sum_{t=0}^{\infty}\beta^t\sum_{j=1}^{N}\tau_j\log(c_{t,j})\right) \tag{A.8}$$

subject to

$$\sum_{j=1}^{N}(c_{t,j}+k_{t+1,j})=\sum_{j=1}^{N}\left((1-\delta)k_{t,j}+\zeta_{t,j}A(K_{t,j})^{\alpha}\right), \tag{A.9}$$

where $\beta=0.99$, $\alpha=0.33$, $\delta=0.025$, $\tau_j\equiv 1$, $A=(1-(1-\delta)\beta)/(\alpha\beta)$, and $\zeta_{t,j}$ follows the same stochastic process (21). The initial state for the jth country is set as

$$K_{0,j}=\exp\left(\ln(K_{\min})+\left(\ln(K_{\max})-\ln(K_{\min})\right)\frac{j-1}{N-1}\right)$$

with $K_{\min}=0.8$ and $K_{\max}=1.2$ for $j=1,\ldots,N$, using the state space in Judd, Maliar, and Maliar (2012) and Maliar and Maliar (2015). Note that the state space is much narrower than what we use in our Case 3 in Section 4.2.3. We apply SCEQ to generate 1000 simulated paths of the first 20 periods.

Table A.1 reports SCEQ's running times (in hours) and Euler errors at the initial state, for the number of countries $N=20$, 40, and 200 with $\Delta_s=50$. We see that SCEQ solves the problems within minutes or hours, at the same order of magnitude with the method of Judd, Maliar, and Maliar (2012) and Maliar and Maliar (2015). However, SCEQ can be much faster, down to seconds or minutes, if we use many compute cores for parallelism, while the method of Judd, Maliar, and Maliar (2012) and Maliar and Maliar (2015) does not have efficient parallelization. The Euler errors of our solution are around 0.5%, less accurate than the method of Judd, Maliar, and Maliar (2012) and Maliar and Maliar (2015) with 0.01–0.1% Euler errors. Thus, for stationary problems without kinks nor a

TABLE A.1. SCEQ's running times and errors for the simple multicountry model.

N	Time (in Hours) for SCEQ	Euler Error at the Initial State
20	0.15	5.3(−3)
40	0.28	5.5(−3)
200	2.58	5.3(−3)

Note: $a(-n)$ means $a\times 10^{-n}$.

wide-ranging state space, the method of Judd, Maliar, and Maliar (2012) and Maliar and Maliar (2015) could be more accurate than SCEQ, which is not surprising because of the certainty equivalent approximation error of SCEQ in exchange of its simplicity, stability, and flexibility in solving problems with kinks, nonstationarity, high dimension, and a wide-ranging state space.

A.4. Parameters in DSICE

This section specifies the values of the parameters and deterministic exogenous processes in Section 4.3, such as L_t (world population), A_t (deterministic total factor productivity), σ_t (carbon intensity of output), $\theta_{1,t}$ (mitigation cost coefficient), $E_{\text{Land},t}$ (annual carbon emissions from biological processes), and $F_{\text{EX},t}$ (exogenous radiative forcing):

$$L_t = 6514e^{-0.035t} + 8600\big(1 - e^{-0.035t}\big), \tag{A.10}$$

$$A_t = 0.0272\exp\big(0.0092\big(1 - e^{-0.001t}\big)/0.001\big), \tag{A.11}$$

$$\sigma_t = 0.13418\exp\big(-0.0073\big(1 - e^{-0.003t}\big)/0.003\big), \tag{A.12}$$

$$\theta_{1,t} = \frac{1.17\sigma_t\big(1 + e^{-0.005t}\big)}{2\theta_2}, \tag{A.13}$$

$$E_{\text{Land},t} = 1.1e^{-0.01t}, \tag{A.14}$$

$$F_{\text{EX},t} = \begin{cases} -0.06 + 0.0036t & \text{if } t \le 100, \\ 0.3 & \text{otherwise.} \end{cases} \tag{A.15}$$

The transition matrix of the carbon cycle is

$$\mathbf{\Phi}_M = \begin{bmatrix} 1 - \phi_{12} & \phi_{21} & 0 \\ \phi_{12} & 1 - \phi_{21} - \phi_{23} & \phi_{32} \\ 0 & \phi_{23} & 1 - \phi_{32} \end{bmatrix}, \tag{A.16}$$

where ϕ_{ij} is the rate at which carbon diffuses from level i to level j, where $i, j = 1, 2, 3$ represent the atmosphere, upper ocean, and lower ocean, respectively. The transition matrix of temperature system is

$$\mathbf{\Phi}_T = \begin{bmatrix} 1 - \varphi_{21} - \xi_2 & \varphi_{21} \\ \varphi_{12} & 1 - \varphi_{12} \end{bmatrix}, \tag{A.17}$$

where φ_{ij} is the heat diffusion rate from level i to level j, where $i, j = 1, 2$ represent the atmosphere and ocean, respectively.

Table A.2 lists the values and definitions of the parameters.

TABLE A.2. Parameters in the deterministic version of DSICE.

$\gamma = 1.45$	risk-aversion parameter
$\beta = 0.985$	utility discount factor
$\alpha = 0.3$	capital share
$\delta = 0.1$	annual capital depreciation rate
$\pi_1 = 0$	climate damage factor parameter
$\pi_2 = 0.0028388$	climate damage factor parameter
$\theta_2 = 2.8$	mitigation cost parameter
$K_0 = 137$	initial capital (in \$ trillions)
$M_{\mathrm{AT},0} = 808.9$	initial carbon concentration in atmosphere (billion tons)
$M_{\mathrm{UO},0} = 1255$	initial carbon concentration in upper ocean (billion tons)
$M_{\mathrm{LO},0} = 18{,}365$	initial carbon concentration in lower ocean (billions tons)
$T_{\mathrm{AT},0} = 0.7307$	global average surface temperature increase above the year 1900 temperature level (in °C)
$T_{\mathrm{OC},0} = 0.0068$	global average ocean temperature increase above the year 1900 temperature level (in °C)
$\phi_{12} = 0.019$	rate of carbon diffusion from atmosphere to upper ocean
$\phi_{23} = 0.0054$	rate of carbon diffusion from upper ocean to lower ocean
$\phi_{21} = 0.01$	rate of carbon diffusion from upper ocean to atmosphere
$\phi_{32} = 0.00034$	rate of carbon diffusion from lower ocean to upper ocean
$\xi_1 = 0.037$	temperature transition parameter
$\xi_2 = 0.047$	rate of $T_{\mathrm{AT},t}$ decrease due to infrared radiation to space
$\varphi_{12} = 0.01\,0$	rate of heat diffusion from atmosphere to ocean
$\varphi_{21} = 0.0048$	rate of heat diffusion from ocean to atmosphere
$\eta = 3.8$	radiative forcing parameter
$M^*_{\mathrm{AT}} = 596.4$	preindustrial atmospheric carbon concentration

A.5. EULER EQUATIONS FOR DSICE WITH ECONOMIC RISK

Let $\beta^s \lambda_s^{\mathbf{M}} \equiv \beta^s(\lambda_s^{\mathrm{MAT}}, \lambda_s^{\mathrm{MUO}}, \lambda_s^{\mathrm{MDO}})$, $\beta^s \lambda_s^{\mathbf{T}} \equiv \beta^s(\lambda_s^{\mathrm{TAT}}, \lambda_s^{\mathrm{TOC}})$, and $\beta^s \lambda_s^{\mathrm{K}}$ be the time-s shadow prices of (22), (23), and (24), respectively, for DSICE with economic risk. Let

$$\widetilde{Y}_s := \Omega(T_{\mathrm{AT},s}) Y_s - \Psi_s$$

with $Y_s = \widetilde{A}_s K_s^{\alpha} L_s^{1-\alpha}$. The system of associated Euler equations is

$$\begin{aligned}
\lambda_s^{\mathrm{K}} &= \beta \mathbb{E}_s \left\{ \lambda_{s+1}^{\mathrm{K}} \left[1 - \delta + \frac{\partial \widetilde{Y}_{s+1}}{\partial K_{s+1}} \right] + \beta \lambda_{s+1}^{\mathrm{MAT}} \frac{\partial E_{s+1}^{\mathrm{Ind}}}{\partial K_{s+1}} \right\}, \\
\lambda_s^{\mathbf{M}} &= \beta \mathbb{E}_s \left\{ \lambda_{s+1}^{\mathbf{M}} \Phi_{\mathbf{M}} + \left(\lambda_{s+1}^{\mathrm{TAT}} \xi_1 \frac{\partial F_{s+1}}{\partial M_{\mathrm{AT},s+1}}, 0, 0 \right) \right\}, \\
\lambda_s^{\mathbf{T}} &= \beta \mathbb{E}_s \left\{ \lambda_{s+1}^{\mathbf{T}} \Phi_{\mathbf{T}} + \left(\lambda_{s+1}^{\mathrm{K}} \frac{\partial \widetilde{Y}_{s+1}}{\partial T_{\mathrm{AT},s+1}}, 0 \right) \right\}.
\end{aligned}$$

Let $\mathbf{x}_t$ denote the vector of continuous state variables $(K, \mathbf{M}_t^{\top}, \mathbf{T}_t^{\top})$, and $\mathbf{a}_t$ the vector of action variables (C_t, μ_t). Let $\mathbf{x}_{t+1} = \mathbf{g}_t(\mathbf{x}_t, \mathbf{a}_t, \boldsymbol{\epsilon}_{t+1})$ denote the six transition laws of the continuous state variables: (22), (23), and (24). Let $\lambda_t := (\lambda_t^{\mathrm{K}}, \lambda_t^{\mathbf{M}}, \lambda_t^{\mathbf{T}})^{\top}$. The normalized

Euler error at time s is defined as

$$\mathcal{E}_s = \left\| \frac{\beta}{\boldsymbol{\lambda}_s} \mathbb{E}_s \left(\nabla_{\mathbf{x}_{s+1}} \mathbf{g}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}, \boldsymbol{\epsilon}_{t+2}) \boldsymbol{\lambda}_{s+1} \right) - \mathbf{1} \right\|,$$

where the division is elementwise.

A.6. DSICE WITH CLIMATE RISK

A climate tipping event may be triggered by global warming and will permanently damage economic output. If the climate tipping event has not occurred by time t, then the probability of the tipping event at time t is

$$p_{\text{tip},t} = 1 - \exp\left\{ -\varrho \max\{0, T_{\text{AT},t} - \underline{T_{\text{AT}}} \} \right\}, \tag{A.18}$$

where $\varrho = 0.0035$ is the hazard rate parameter and $\underline{T_{\text{AT}}}$ is the temperature anomaly for which $p_{\text{tip},t} = 0$. Once the tipping event happens, it is irreversible and incurs a damage of $\overline{J} = 0.05$ to output. That is, the tipping process is a Markov chain with an endogenous tipping probability. We use J_t to measure the possible values of the Markov chain, where $J_t = 0$ if the tipping event has not occurred by time t, and $J_t = \overline{J}$ at all times t after the tipping event.

Output $Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}$ is reduced by the temperature anomaly and a potential tipping event, with a damage factor of

$$\Omega(T_{\text{AT},t}, J_t) = \frac{1 - J_t}{1 + \pi_1 T_{\text{AT},t} + \pi_2 (T_{\text{AT},t})^2}.$$

Thus, the transition law of capital is

$$K_{t+1} = (1 - \delta) K_t + \Omega(T_{\text{AT},t}, J_t) Y_t - C_t - \Psi_t. \tag{A.19}$$

We maximize the social planner's objective function (26) subject to the transition laws (22)–(23), (A.19), and the Markov chain of J_t. This problem is nonstationary and stochastic with seven endogenous state variables (of which six are continuous) and occasionally binding constraints, so it is challenging to solve using standard methods. Here, we apply SCEQ to overcome these challenges, as we do for the model with economic risk.

SCEQ requires replacing the stochastic Markov chain J_t by its conditional expectation $\mathbb{E}(J_t \mid J_s)$ for $t > s$ in the optimization step of Algorithm 1, given the simulated state vector at the current time s. Since the tipping probability is endogenous, the conditional expectation $\mathbb{E}(J_t \mid J_s)$ is also endogenous, unless the tipping event has occurred by the current time s (i.e., $J_s = \overline{J}$), in which case $\mathbb{E}(J_t \mid J_s = \overline{J}) = \overline{J}$. If $J_s = 0$, that is, the tipping event has not occurred by the current time s, then

$$\mathbb{E}(J_t \mid J_s = 0) = \left(1 - \prod_{i=0}^{t-s-1} (1 - p_{\text{tip},s+i}) \right) \overline{J}$$

for any $t > s$. Thus, in the optimization step of Algorithm 1, we are solving

$$\max_{C_t, \mu_t} \quad \sum_{t=s}^{600} \beta^{t-s} u(C_t/L_t) L_t \tag{A.20}$$

subject to the transition laws (22)–(23) and the following deterministic transition equation of capital:

$$K_{t+1} = (1-\delta)K_t + \Omega\big(T_{\mathrm{AT},t}, \mathbb{E}(J_t \mid J_s)\big) Y_t - C_t - \Psi_t, \tag{A.21}$$

for all $t \geq s$. That is, if the tipping event has happened by the current time s, then

$$K_{t+1} = (1-\delta)K_t + \Omega(T_{\mathrm{AT},t}, \overline{J}) Y_t - C_t - \Psi_t, \tag{A.22}$$

for all $t \geq s$, and in the simulation step of Algorithm 1, we have

$$K^i_{s+1} = (1-\delta)K^i_s + \Omega\big(T^i_{\mathrm{AT},s}, \overline{J}\big) Y^i_s - C^i_s - \Psi^i_s,$$

for the ith simulation; otherwise in the optimization step of Algorithm 1,

$$K_{t+1} = \begin{cases} (1-\delta)K_t + \Omega\left(T_{\mathrm{AT},t}, \left(1 - \prod_{i=0}^{t-s-1} (1 - p_{\mathrm{tip},s+i})\right)\overline{J}\right) Y_t - C_t - \Psi_t & \text{if } t > s, \\ (1-\delta)K_t + \Omega(T_{\mathrm{AT},t}, 0) Y_t - C_t - \Psi_t & \text{if } t = s, \end{cases}$$

and in the simulation step of Algorithm 1, we have

$$K^i_{s+1} = (1-\delta)K^i_s + \Omega\big(T^i_{\mathrm{AT},s}, 0\big) Y^i_s - C^i_s - \Psi^i_s,$$

for the ith simulation.

Note that the tipping event is irreversible. If the tipping event has happened by time s, then there are no more stochastic events from time s onwards, so in the optimization step of Algorithm 1 we solve (A.20) subject to the transition laws (22)–(23) and (A.22). The simulation step of Algorithm 1 will not change the solution paths of state and decision variables after s, because there is only one simulation path after time s for each state or decision variable. Therefore, if we are interested in the first T^* years' simulated results, there are T^* different simulated paths for each state or decision variable: the first is where the tipping event does not happen in the first T^* years (the "pretipping path"), the second is where the tipping event happens in the second year (assuming its tipping probability is nonzero), and so on, with the last path having the tipping event happen in year T^*. That is, we just need to solve $2T^* - 1$ optimization problems to obtain the T^* different paths: T^* optimization problems are for the pretipping path, and the other $T^* - 1$ optimization problems are for the cases where the tipping event happens before or at T^* (assuming that the tipping state has not happened in the first period). With these T^* paths of atmospheric temperature anomalies, we can compute their corresponding tipping probabilities (A.18), then generate 1000 or more simulated paths of carbon taxes and other state and decision variables as well as shadow prices of the constraints, using

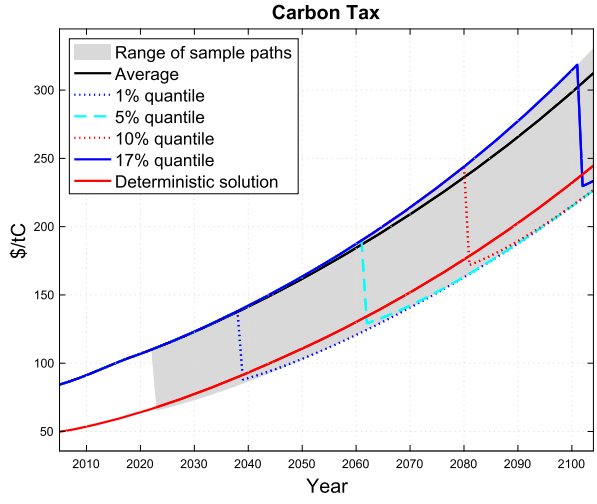

FIGURE A.1. Carbon tax for DSICE with climate risk.

the corresponding carbon taxes saved in the T^* different paths, without solving more optimization problems.

Figure A.1 displays distributions of the optimal carbon taxes in the first $T^* = 100$ years from 10,000 paths simulated from the $T^* = 100$ different paths obtained by SCEQ. The red solid line represents the optimal carbon taxes in the corresponding deterministic model, assuming no tipping event. The upper envelope in Figure A.1 represents the pretipping path of optimal carbon taxes. In contrast, the lower envelope represents the optimal carbon taxes after the tipping event happens.

The initial optimal carbon tax of the stochastic model is significantly higher than that of the deterministic model. Moreover, the optimal carbon tax jumps down significantly once the tipping event occurs, because the tipping probability depends on temperature, so before the tipping event happens there is an incentive to further reduce emissions and lower temperature anomaly to delay or prevent the tipping event from happening. Once the tipping event happens, this incentive is gone. This pattern has been shown in Cai, Judd, and Lontzek (2017) and Cai and Lontzek (2019). Figure A.1 also displays the timing of some sample tipping events. For example, the first drop (which occurs in the year 2023 in Figure A.1) is the first tipping point out of our 10,000 simulations. By the end of the 21st century, 17% of the paths have exhibited a tipping point.

To check the errors of our solution, we compute the Euler errors. Since our problem has an endogenous probability, depending on the atmospheric temperature anomaly, the Euler equations (A.1) and the Euler errors (A.4) should be adjusted (see Appendix A.7). For this specific example, we can compute the errors for all possible paths. We are interested in the solution in the first 100 years, and there are at most 100 different paths in the first 100 years: in our case, there are only 86 different paths because the atmospheric temperature anomalies in the first 14 years are less than $\underline{T_{\text{AT}}} = 1$ °C in our solution, so the corresponding tipping probability is zero. That is, we just need to solve 186 optimization problems to obtain all different paths.

As the tipping event is irreversible, the solution after the tipping event is deterministic, so SCEQ will provide an accurate solution for the periods after the tipping event. Thus, we just need to check the errors on the pretipping path, shown as the upper envelope in Figure A.1. Moreover, we can obtain the tipping probabilities $p_{\text{tip},s}$ at each time s from the pretipping path of atmospheric temperature anomalies. Therefore, we use the 86 different paths to obtain the Euler errors $\mathcal{E}_s$, defined in Appendix A.7, at each time s on the pretipping path. The Euler errors $\mathcal{E}_s$ in the first 14 years are nearly zero, as there is no tipping event in these initial years. But in the 15th year, $\mathcal{E}_s$ becomes 0.0014 in the $\mathcal{L}^\infty$ norm. To obtain a good estimate of Euler errors, we should choose the number of years of interest, T^*, to be large enough to cover at least one period with tipping. We find that the $\mathcal{L}^\infty$ Euler error over the first 100 years is also 0.0014, while the $\mathcal{L}^1$ Euler error is much smaller (1.4×10^{-4}).

We also compare the SCEQ solution with the solution obtained from VFI as in Cai, Judd, and Lontzek (2017) and Cai and Lontzek (2019). We use time-varying approximation domains and degree-6 complete Chebyshev polynomials for approximating the value functions in VFI. We compare the pretipping paths of optimal carbon taxes in the first 100 years across the two solutions. If we treat the VFI solution as the "true" solution, the relative $\mathcal{L}^\infty$ error is 0.0096 and the relative $\mathcal{L}^1$ error is 0.0049. In addition, the relative $\mathcal{L}^\infty$ error is 6.8×10^{-5} and 0.0059 for C_t and μ_t, respectively, and the relative $\mathcal{L}^1$ error is 3.1×10^{-5} and 0.0041 for C_t and μ_t, respectively. These results indicate that SCEQ can obtain a solution with acceptable accuracy for a problem with an endogenous tipping risk, even though the risk has a significant impact on the solution (as shown in Figure A.1).

For this specific example, SCEQ took only 2 minutes on a single compute core. The fast running time of SCEQ is due to the small number of different simulation paths (86) and the corresponding optimization problems (186), so we do not have to run 1000 simulation paths as in the DSICE example with economic risk, which requires solving 100,000 optimization problems for the first 100 years.

We also apply SCEQ to solve cases with a larger risk aversion parameter $\gamma = 2$ or a larger tipping damage $\overline{J} = 0.1$. Table A.3 displays the Euler errors and the relative errors of the optimal carbon taxes along the pretipping path in comparison with corresponding VFI solutions for the first 100 years. A larger degree of risk aversion γ implies larger errors, as SCEQ uses the certainty equivalent approximation ignoring risk aversion, so γ serves as the inverse of intertemporal elasticity of substitution. Also, a larger tipping damage $\overline{J}$ implies a larger variance (which SCEQ ignores), and hence larger errors. However, the SCEQ solution of carbon taxes still has only an error of 4.5% in the $\mathcal{L}^\infty$ norm for the worst case scenario (both γ and $\overline{J}$ are larger),[40] if we treat the VFI solution as the true solution.

[40]The relative errors for the optimal decision variables (C_t, μ_t) are smaller and closer to the Euler errors for the worst case scenario: the relative $\mathcal{L}^\infty$ error is 1.6×10^{-4} and 0.027 for C_t and μ_t, respectively, and the relative $\mathcal{L}^1$ error is 8.4×10^{-5} and 0.019 for C_t and μ_t, respectively.

TABLE A.3. Errors for carbon taxes in the first 100 years.

γ	$\overline{J}$	Euler Error		Relative Error	
		$\mathcal{L}^\infty$	$\mathcal{L}^1$	$\mathcal{L}^\infty$	$\mathcal{L}^1$
1.45	0.05	1.4(−3)	1.4(−4)	9.6(−3)	4.9(−3)
	0.1	3.7(−3)	4.4(−4)	3.5(−2)	1.9(−2)
2	0.05	2.1(−3)	2.2(−4)	1.5(−2)	7.2(−3)
	0.1	6.2(−3)	7.0(−4)	4.5(−2)	2.6(−2)

Note: $a(-n)$ means $a \times 10^{-n}$.

A.7. EULER EQUATIONS FOR DSICE WITH CLIMATE RISK

If the tipping event happens at s, then the stochastic problem becomes deterministic from time s onwards, as the tipping event is irreversible. Our SCEQ solution at times larger than s is therefore accurate as there is no certainty approximation error for the deterministic problem. Thus, to estimate Euler errors, we assume that the tipping event has not happened by time s, that is, $J_s = 0$.

Let $\beta^s \lambda_s^{\mathbf{M}} \equiv \beta^s(\lambda_s^{\text{MAT}}, \lambda_s^{\text{MUO}}, \lambda_s^{\text{MDO}})$, $\beta^s \lambda_s^{\mathbf{T}} \equiv \beta^s(\lambda_s^{\text{TAT}}, \lambda_s^{\text{TOC}})$, and $\beta^s \lambda_s^{\text{K}}$ be time-s shadow prices of (22), (23), and (A.19), respectively, for the DSICE model with climate risk. Let

$$\widehat{Y}_s := \Omega(T_{\text{AT},s}, J_s) Y_s - \Psi_s.$$

Let $V_{\text{tipped},s+2}$ be the aggregate welfare from time $s+2$ onwards, assuming that the tipping event happens in period $s+2$, that is,

$$V_{\text{tipped},s+2} = \max_{C_t, \mu_t} \quad \sum_{t=s+2}^{\infty} \beta^{t-s-2} u(C_t/L_t) L_t$$

subject to the transition laws (22), (23), and (A.22), and let $V_{\text{untipped},s+2}$ be the aggregate welfare from time $s+2$ onwards, assuming that the tipping event has not happened by period $s+2$, that is,

$$V_{\text{untipped},s+2} = \max_{C_t, \mu_t} \quad \mathbb{E}_s \left\{ \sum_{t=s+2}^{\infty} \beta^{t-s-2} u(C_t/L_t) L_t \right\}$$

subject to the transition laws (22), (23), and (A.19).

Thus, the system of associated Euler equations for (26) is

$$\lambda_s^{\text{K}} = \beta \mathbb{E}_s \left\{ \lambda_{s+1}^{\text{K}} \left[1 - \delta + \frac{\partial \widehat{Y}_{s+1}}{\partial K_{s+1}} \right] + \beta \lambda_{s+1}^{\text{MAT}} \frac{\partial E_{s+1}^{\text{Ind}}}{\partial K_{s+1}} \right\},$$

$$\lambda_s^{\mathbf{M}} = \beta \mathbb{E}_s \left\{ \lambda_{s+1}^{\mathbf{M}} \Phi_{\mathbf{M}} + \left(\lambda_{s+1}^{\text{TAT}} \xi_1 \frac{\partial F_{s+1}}{\partial M_{\text{AT},s+1}}, 0, 0 \right) \right\},$$

$$\lambda_s^{\mathbf{T}} = \beta \mathbb{E}_s \left\{ \lambda_{s+1}^{\mathbf{T}} \Phi_{\mathbf{T}} + \left(\lambda_{s+1}^{\text{K}} \frac{\partial \widehat{Y}_{s+1}}{\partial T_{\text{AT},s+1}}, 0 \right) \right\} + \beta^2 (\Lambda_{s+2}, 0),$$

where

$$\Lambda_{s+2} := (1 - p_{\text{tip},s}) \frac{\partial p_{\text{tip},s+1}}{\partial T_{\text{AT},s+1}} (V_{\text{tipped},s+2} - V_{\text{untipped},s+2}).$$

Note that the term $\beta^2(\Lambda_{s+2}, 0)$ is not present in the standard Euler equations (A.1) with exogenous transition probabilities. In computing the Euler error (A.4), $V_{\text{untipped},s+2}$ is approximated by

$$\widetilde{V}_{\text{untipped},s+2} = \max_{C_t, \mu_t} \sum_{t=s+2}^{\infty} \beta^{t-s-2} u(C_t/L_t) L_t$$

subject to the transition laws (22), (23), and (A.21) conditional on $J_s = 0$.

Given the time-s information with $J_s = 0$, λ^{K}_{s+1}, $\lambda^{\mathbf{M}}_{s+1}$, $\lambda^{\mathbf{T}}_{s+1}$, $\widehat{Y}_{s+1}$, E^{Ind}_{s+1}, J_{s+1}, and all decision variables at $s+1$ are random with two possible values: one value for the case that the tipped event happens at $s+1$ ($J_{s+1} = \overline{J}$), the other value for the case that the tipping event does not happen at $s+1$ ($J_{s+1} = 0$). Thus, the conditional expectation operator in the above Euler equations for a random variable X_{s+1} at time $s+1$ is

$$\mathbb{E}_s\{X_{s+1}\} = p_{\text{tip},s} X_{\text{tipped},s+1} + (1 - p_{\text{tip},s}) X_{\text{untipped},s+1},$$

where $X_{\text{tipped},s+1}$ is the value of X_{s+1} for the case that the tipped event happens at time $s+1$, and $X_{\text{untipped,s+1}}$ the value of X_{s+1} for the case that the tipped event does not happen at time $s+1$.

Let $\mathbf{x}_t$ denote the vector of continuous state variables $(K, \mathbf{M}_t^\top, \mathbf{T}_t^\top)$, and $\mathbf{a}_t$ the vector of action variables (C_t, μ_t). Let $\mathbf{x}_{t+1} = \mathbf{g}_t(\mathbf{x}_t, \mathbf{a}_t, J_t)$ denote the six transition laws of the continuous state variables: (22), (23), and (A.22). Let $\lambda_t := (\lambda^{\text{K}}_t, \lambda^{\mathbf{M}}_t, \lambda^{\mathbf{T}}_t)^\top$. The normalized Euler error on the pretipping path at time s is defined as

$$\mathcal{E}_s = \left\| \frac{\beta}{\boldsymbol{\lambda}_s} \left[\mathbb{E}_s \left(\nabla_{\mathbf{x}_{s+1}} \mathbf{g}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}, J_{s+1}) \boldsymbol{\lambda}_{s+1} \right) + \beta(0, 0, 0, 0, \Lambda_{s+2}, 0)^\top \right] - \mathbf{1} \right\|,$$

where the division is elementwise.

A.8. New Keynesian model and equilibrium conditions

The representative household chooses consumption c_t of final goods, labor supply ℓ_t, and government bonds b_t to maximize the total present value of expected utilities, that is,

$$\max_{c_t, \ell_t, b_t} \mathbb{E} \left\{ \sum_{t=0}^{\infty} \left(\prod_{i=0}^{t} \beta_i \right) U(c_t, \ell_t) \right\} \tag{A.23}$$

subject to the budget constraint

$$p_t c_t + \frac{b_t}{1 + r_t} = w_t \ell_t + b_{t-1} + T_t + \Pi_t, \tag{A.24}$$

where p_t is price of the final good, r_t is the nominal interest rate of bonds, w_t is wage for labor supply, T_t is the lump-sum transfer from the government, and Π_t is the total profit

from all firms at period t. The utility function is

$$U(c, \ell) = \ln(c) - \frac{\ell^{1+\eta}}{1+\eta},$$

where $\eta = 1$. The parameters in the stochastic process of β_t are $\beta^* = 0.994$, $\rho = 0.8$, and $\sigma = 0.005$. The first-order conditions of the household problem imply

$$1 = \mathbb{E}_t \left\{ \beta_{t+1} \frac{1+r_t}{\pi_{t+1}} \frac{c_t}{c_{t+1}} \right\} \tag{A.25}$$

and

$$w_t = p_t c_t \ell_t^{\eta}, \tag{A.26}$$

where $\pi_t \equiv p_t / p_{t-1}$ is the gross inflation rate.

The final-good firm buys intermediate goods $y_{i,t}$ with prices $p_{i,t}$ from intermediate firms, and produces a final good y_t with the following production function:

$$y_t = \left(\int_0^1 y_{i,t}^{\frac{\alpha-1}{\alpha}} di \right)^{\frac{\alpha}{\alpha-1}}, \tag{A.27}$$

where $\alpha = 6$. The final-good firm chooses $y_{i,t}$ to maximize its profit:

$$\max_{y_{i,t}} p_t y_t - \int_0^1 p_{i,t} y_{i,t} \, di,$$

which, from the first-order condition, implies that

$$y_{i,t} = y_t \left(\frac{p_{i,t}}{p_t} \right)^{-\alpha}. \tag{A.28}$$

The intermediate firms rent labor supply $\ell_{i,t}$ from the household and have the simple production function

$$y_{i,t} = \ell_{i,t}. \tag{A.29}$$

We assume the Calvo parameter θ is set as 0.9: 90% of the firms in each period keep the same price as the previous period, and 10% of the firms have optimal prices. A reoptimizing intermediate firm $i \in [0, 1]$ chooses its price $p_{i,t}$ to maximize the present value of expected profits over the time when the optimal $p_{i,t}$ remains effective:

$$\max_{p_{i,t}} \mathbb{E}_t \left\{ \sum_{j=0}^{\infty} \left(\prod_{k=0}^{j} \beta_{t+k} \right) \lambda_{t+j} \theta^j (p_{i,t} y_{i,t+j} - w_{t+j} \ell_{i,t+j}) \right\} \tag{A.30}$$

subject to the constraints $y_{i,t+j} = \ell_{i,t+j}$ from (A.29) and

$$y_{i,t+j} = y_{t+j} \left(\frac{p_{i,t}}{p_{t+j}} \right)^{-\alpha}$$

from (A.28) by letting $p_{i,t+j} = p_{i,t}$. Here, λ_t is the Lagrange multiplier of the budget constraint (A.24), satisfying

$$\lambda_t = \frac{1}{p_t c_t}, \tag{A.31}$$

from the first-order conditions of the household problem (A.23). Let $\pi_{t,j} = p_{t+j}/p_t$. From (A.26), (A.31), and the first-order condition of the reoptimizing intermediate firm problem (A.30), for any reoptimizing firm i we have

$$\frac{p_{i,t}}{p_t} \equiv q_t = \frac{\alpha \chi_{t,1}}{(\alpha-1)\chi_{t,2}}, \tag{A.32}$$

where

$$\chi_{t,1} \equiv y_t \ell_t^{\eta} + \mathbb{E}_t \left\{ \sum_{j=1}^{\infty} \left(\prod_{k=1}^{j} \beta_{t+k} \right) \theta^j \pi_{t,j}^{\alpha} y_{t+j} \ell_{t+j}^{\eta} \right\},$$

$$\chi_{t,2} \equiv \frac{y_t}{c_t} + \mathbb{E}_t \left\{ \sum_{j=1}^{\infty} \left(\prod_{k=1}^{j} \beta_{t+k} \right) \theta^j \pi_{t,j}^{\alpha-1} \frac{y_{t+j}}{c_{t+j}} \right\}.$$

We have the recursive formulas for $\chi_{t,1}$ and $\chi_{t,2}$:

$$\chi_{t,1} = y_t \ell_t^{\eta} + \theta \mathbb{E}_t \left\{ \beta_{t+1} \pi_{t+1}^{\alpha} \chi_{t+1,1} \right\}, \tag{A.33}$$

$$\chi_{t,2} = \frac{y_t}{c_t} + \theta \mathbb{E}_t \left\{ \beta_{t+1} \pi_{t+1}^{\alpha-1} \chi_{t+1,2} \right\}. \tag{A.34}$$

From (A.27) and (A.28), we have

$$\begin{aligned} p_t &= \left(\int_0^1 p_{i,t}^{1-\alpha} \, di \right)^{\frac{1}{1-\alpha}} \\ &= \left((1-\theta)(q_t p_t)^{1-\alpha} + \theta \int_0^1 p_{i,t-1}^{1-\alpha} \, di \right)^{\frac{1}{1-\alpha}} \\ &= \left((1-\theta)(q_t p_t)^{1-\alpha} + \theta p_{t-1}^{1-\alpha} \right)^{\frac{1}{1-\alpha}}. \end{aligned}$$

This follows that

$$q_t = \left(\frac{1 - \theta \pi_t^{\alpha-1}}{1-\theta} \right)^{\frac{1}{1-\alpha}}. \tag{A.35}$$

From (A.28), (A.29) and the market clearing condition,

$$\ell_t = \int_0^1 \ell_{i,t} \, di,$$

we get

$$\begin{aligned} v_{t+1} \equiv \ell_t / y_t &= \int_0^1 \left(\frac{p_{i,t}}{p_t} \right)^{-\alpha} di \\ &= (1-\theta) q_t^{-\alpha} + \theta \int_0^1 \left(\frac{p_{i,t-1}}{p_t} \right)^{-\alpha} di \\ &= (1-\theta) q_t^{-\alpha} + \theta \pi_t^{\alpha} \int_0^1 \left(\frac{p_{i,t-1}}{p_{t-1}} \right)^{-\alpha} di \\ &= (1-\theta) q_t^{-\alpha} + \theta \pi_t^{\alpha} v_t. \end{aligned} \tag{A.36}$$

From $v_{t+1} = \ell_t / y_t$, (A.33) can be rewritten as

$$\chi_{t,1} = y_t^{1+\eta} v_{t+1} + \theta \mathbb{E}_t \{ \beta_{t+1} \pi_{t+1}^{\alpha} \chi_{t+1,1} \}.$$

Let π^*, r^*, and y^* be the steady-state gross level of inflation, nominal interest rate, and output, respectively. Let $\pi^* = 1.005$ be given. We have $r^* = \pi^* / \beta^* - 1$ from (A.25). From (A.34), the steady state of $\chi_{t,2}$ is

$$\chi_2^* = \frac{1}{(1-s_g)\left(1 - \theta \beta^* \left(\pi^*\right)^{\alpha-1}\right)}.$$

From (A.32) and (A.35), the steady state of $\chi_{t,1}$ is

$$\chi_1^* = \chi_2^* q^* \frac{\alpha - 1}{\alpha}$$

with

$$q^* = \left(\frac{1 - \theta \left(\pi^*\right)^{\alpha-1}}{1-\theta} \right)^{\frac{1}{1-\alpha}}$$

and from (A.36) the steady state of v_t is

$$v^* = \frac{(1-\theta)\left(q^*\right)^{-\alpha}}{1 - \theta \left(\pi^*\right)^{\alpha}}.$$

Therefore, from $v_t = \ell_t / y_t$ and (A.33), we get

$$y^* = \left(\frac{\chi_1^* \left(1 - \theta \beta^* \left(\pi^*\right)^{\alpha}\right)}{\left(v^*\right)^{\eta}} \right)^{\frac{1}{1+\eta}}.$$

Let government spending g_t be always equal to $s_g y_t$ with $s_g = 0.2$. From the market clearing condition $y_t = c_t + g_t$, we have

$$c_t = (1 - s_g) y_t. \tag{A.37}$$

Thus, (A.34) can be rewritten as

$$1 = \frac{1}{\chi_{t,2}} \left(\frac{1}{1-s_g} + \theta \mathbb{E}_t \{ \beta_{t+1} \pi_{t+1}^{\alpha-1} \chi_{t+1,2} \} \right).$$

Following the Taylor rule, we have the nominal interest rate as

$$r_t = \max(z_t, 0) \tag{A.38}$$

with

$$z_t = (1 + r^*)\left(\frac{\pi_t}{\pi^*}\right)^{\phi_\pi}\left(\frac{y_t}{y^*}\right)^{\phi_y} - 1, \tag{A.39}$$

where we choose $\phi_\pi = 2.5$ and $\phi_y = 0.25$. Thus, (A.25) can be rewritten as

$$1 = \mathbb{E}_t\left\{\beta_{t+1}\frac{1 + \max(z_t, 0)}{\pi_{t+1}}\frac{y_t}{y_{t+1}}\right\}.$$

References

Bellman, Richard (1957), *Dynamic Programming*. Princeton University Press, Princeton. [4]

Cai, Yongyang (2019), "Computational methods in environmental and resource economics." *Annual Review of Resource Economics*, 11, 59–82. [3]

Cai, Yongyang and Kenneth L. Judd (2014), "Advances in numerical dynamic programming and new applications." In *Handbook of Computational Economics*, Vol. 3 (Kenneth L. Judd and Karl Schmedders, eds.), 479–516, Elsevier, Amsterdam. [5]

Cai, Yongyang, Kenneth L. Judd, and Thomas S. Lontzek (2017), "The social cost of carbon with economic and climate risks." Hoover Economics Working Paper 18113, https://www.hoover.org/research/social-cost-carbon-economic-and-climate-risk. [11, 12]

Cai, Yongyang and Thomas S. Lontzek (2019), "The social cost of carbon with economic and climate risks." *Journal of Political Economy*, 127 (6), 2684–2734. [11, 12]

Jin, He-Hui and Kenneth L. Judd (2002), "Perturbation methods for general dynamic stochastic models." Working paper, Stanford University. [3]

Judd, Kenneth L., Lilia Maliar, and Serguei Maliar (2012), "Merging simulation and projection approaches to solve high-dimensional problems." Tehcnical Report w18501, National Bureau of Economic Research, https://www.nber.org/papers/w18501. [6, 7]

Juillard, Michel and Sebastien Villemot (2011), "Multi-country real business cycle models: Accuracy tests and test bench." *Journal of Economic Dynamics and Control*, 35 (2), 178–185. [3]

Maliar, Lilia and Serguei Maliar (2015), "Merging simulation and projection approaches to solve high-dimensional problems with an application to a new Keynesian model." *Quantitative Economics*, 6 (1), 1–47. [6, 7]

Co-editor Tao Zha handled this manuscript.

Manuscript received 8 February, 2021; final version accepted 19 December, 2022; available online 31 January, 2023.