

Online Appendix for “A Simple but Powerful Simulated Certainty Equivalent Approximation Method for Dynamic Stochastic Problems”

Yongyang Cai Kenneth Judd

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 $\beta^t \lambda_t$ be a $n \times 1$ column vector representing the Lagrange multipliers of the vector of state transition equations $\mathbf{x}_{t+1} = \mathbf{g}_t(\mathbf{x}_t, \mathbf{a}_t, \epsilon_t)$, and $\beta^t \nu_t$ 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 for the stochastic model (1) at period s is³³

$$\lambda_s = \beta \mathbb{E}_s \left(\nabla_{\mathbf{x}_{s+1}} \mathbf{g}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}, \epsilon_{s+1}) \lambda_{s+1} + \nabla_{\mathbf{x}_{s+1}} \mathbf{f}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}) \nu_{s+1} \right) \quad (\text{A.1})$$

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} ,

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

³³These 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.3 for one example.

is an $n \times n$ matrix,

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

is an $n \times k$ matrix.³⁴ The normalized Euler error is defined as

$$\left\| \beta \mathbb{E}_s \left(\nabla_{\mathbf{x}_{s+1}} \mathbf{g}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}, \epsilon_{s+1}) \lambda_{s+1} + \nabla_{\mathbf{x}_{s+1}} \mathbf{f}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}) \nu_{s+1} \right) / \lambda_s - \mathbf{1} \right\| \quad (\text{A.2})$$

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 non-normalized 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 λ_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 λ_s^i , and save them for accuracy measures.

Given the time- s information $(\mathbf{x}_s^i, \mathbf{a}_s^i, \lambda_s^i)$, the next state in the original stochastic model (1), $\mathbf{g}_s(\mathbf{x}_s^i, \mathbf{a}_s^i, \epsilon_s)$, is random and dependent on $\epsilon_s \in \Theta$, where Θ represents the set of all possible vectorial values of ϵ_s . 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 $(\lambda_{s+1}, \nu_{s+1})$, all of which are random.³⁵ 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 ϵ_s is a discrete random variable vector with a finite number of possible values $\{\epsilon_{s,j}\}$, we compute $\mathbf{x}_{s+1,j} = \mathbf{g}_s(\mathbf{x}_s^i, \mathbf{a}_s^i, \epsilon_{s,j})$ for every $\epsilon_{s,j} \in \Theta$, then solve (2) with the starting time $s + 1$ and starting state $\mathbf{x}_{s+1,j}$ to find the associated optimal

³⁴In 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.
```

³⁵ \mathbf{x}_{s+1} can be nonrandom conditional on time- s information.

decision $\mathbf{a}_{s+1,j}$ and their corresponding multipliers $(\lambda_{s+1,j}, \nu_{s+1,j})$. That is, we solve

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

where \mathbf{x}_{s+1} is given by $\mathbf{x}_{s+1,j}$. We then compute the conditional expectation in equation (A.1). When ϵ_s is a continuous random variable vector, we can use its quadrature nodes $\epsilon_{s,j}$, then implement quadrature rules to estimate the conditional expectation in equation (A.1). Thus, we can approximately estimate the Euler error (A.2) at period s and state \mathbf{x}_s^i as

$$\mathcal{E}_s^i = \left\| \beta \mathbb{E}_s \left(\nabla_{\mathbf{x}_{s+1}} \mathbf{g}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}, \epsilon_{s+1}) \lambda_{s+1} + \nabla_{\mathbf{x}_{s+1}} \mathbf{f}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}) \nu_{s+1} \right) / \lambda_s^i - \mathbf{1} \right\| \quad (\text{A.4})$$

where \mathbf{x}_{s+1} , \mathbf{a}_{s+1} , ϵ_{s+1} , λ_{s+1} and ν_{s+1} are random and obtained from solving (A.3) for all possible states \mathbf{x}_{s+1} or for the quadrature nodes of \mathbf{x}_{s+1} . Here the expectation is conditional on the time- s simulated solution: $(\mathbf{x}_s^i, \mathbf{a}_s^i, \lambda_s^i)$.

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

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

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

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

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

If there are m simulated paths with T^* periods, and J possible values of ϵ_t (or

J quadrature nodes for a continuous ϵ_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 has 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 $\mathbf{x}_{s+1,j}$, $\mathbf{a}_{s+1,j}$, $\lambda_{s+1,j}$ and $\nu_{s+1,j}$ 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 real business cycle problems with 200 countries and $\Delta_s = 50$.

In addition, if we solve all different possible paths of $(\mathbf{x}_s^i, \mathbf{a}_s^i, \lambda_s^i)$ 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 Euler Equations for DSICE with Economic Risk

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 $\beta^s \lambda_s^{\mathbf{K}}$ be the time- s shadow prices of (22), (23), and (24) respectively for DSICE with economic risk. Let

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

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

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

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 $\mathbf{x}_{t+1} = \mathbf{g}_t(\mathbf{x}_t, \mathbf{a}_t, \zeta_t)$ denote the six transition laws of the continuous state variables: (22), (23), and (24). Let $\lambda_t := (\lambda_t^{\mathbf{K}}, \lambda_t^{\mathbf{M}}, \lambda_t^{\mathbf{T}})^\top$. The normalized Euler error at time s is defined as

$$\mathcal{E}_s = \left\| \beta \mathbb{E}_s \left(\nabla_{\mathbf{x}_{s+1}} \mathbf{g}_{s+1}(\mathbf{x}_{s+1}, \mathbf{a}_{s+1}, \zeta_{s+1}) \lambda_{s+1} \right) / \lambda_s - \mathbf{1} \right\|$$

where the division is elementwise.

A.3 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

$$p_{\text{tip},t} = 1 - \exp \left\{ -\varrho \max \left\{ 0, T_{\text{AT},t} - \underline{T}_{\text{AT}} \right\} \right\}, \quad (\text{A.7})$$

where $\rho = 0.0035$ 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 $\bar{J} = 0.05$ 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 $J_t = \bar{J}$ 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

$$K_{t+1} = (1 - \delta)K_t + \Omega(T_{\text{AT},t}, J_t)Y_t - C_t - \Psi_t. \quad (\text{A.8})$$

We maximize the social planner's objective function (26) subject to the transition laws (22)-(23), (A.8), 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 $\mathbb{E}(J_t | J_s)$ 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 $\mathbb{E}(J_t | J_s)$ is also endogenous, unless the tipping event has occurred by the current time s (i.e., $J_s = \bar{J}$), in which case $\mathbb{E}(J_t | J_s = \bar{J}) = \bar{J}$. If $J_s = 0$, i.e., the tipping event has not occurred by the current time s , then

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

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

$$\max_{C_t, \mu_t} \sum_{t=s}^{600} \beta^{t-s} u(C_t/L_t) L_t \quad (\text{A.9})$$

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

$$K_{t+1} = (1 - \delta)K_t + \Omega(T_{\text{AT},t}, \mathbb{E}(J_t | J_s)) Y_t - C_t - \Psi_t, \quad (\text{A.10})$$

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

$$K_{t+1} = (1 - \delta)K_t + \Omega(T_{\text{AT},t}, \bar{J}) Y_t - C_t - \Psi_t, \quad (\text{A.11})$$

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

$$K_{s+1}^i = (1 - \delta)K_s^i + \Omega(T_{\text{AT},s}^i, \bar{J}) Y_s^i - C_s^i - \Psi_s^i,$$

for the i -th simulation; otherwise in the optimization step of Algorithm 1,

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

and in the simulation step of Algorithm 1, we have

$$K_{s+1}^i = (1 - \delta)K_s^i + \Omega(T_{\text{AT},s}^i, 0) Y_s^i - C_s^i - \Psi_s^i,$$

for the i -th simulation.

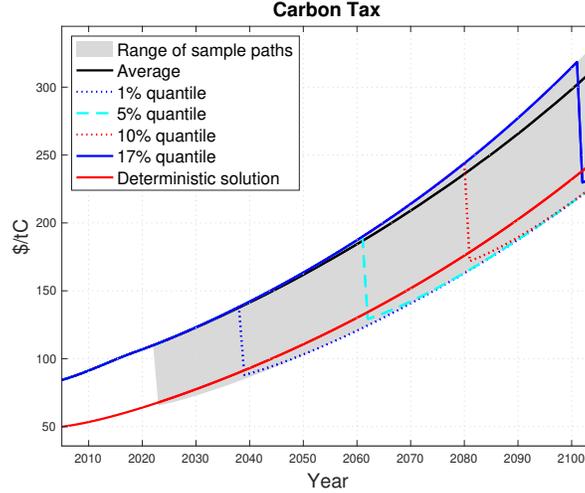
Note that the tipping event is irreversible. If the tipping event has happened by time s , then there are no more stochastic events from time s onwards, so in the optimization step of Algorithm 1 we solve (A.9) subject to the transition laws (22)-(23) and (A.11). The simulation step of Algorithm 1 will 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 ‘‘pre-tipping 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 pre-tipping 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.7), then generate 1,000 or more simulated paths of carbon taxes and other state and decision variables as well as shadow prices of the constraints, using 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 pre-tipping 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.

Figure A.1: Carbon Tax for DSICE with Climate Risk



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.4). 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 $T_{AT} = 1$ °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 pre-tipping path, shown as the upper envelope in Figure A.1. Moreover, we can obtain the tipping probabilities $p_{tip,s}$ at each time s from the pre-tipping path of atmospheric temperature anomalies. Therefore we use the 86 different paths to obtain the Euler errors \mathcal{E}_s , defined in Appendix A.4, at each time s on the pre-tipping 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}^∞ 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}^∞ 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 pre-tipping 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}^∞ error is 0.0096 and the relative \mathcal{L}^1 error is 0.0049. In addition, the relative \mathcal{L}^∞ 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 an accuracy of 0.1-1% 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 two 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 1,000 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 $\bar{J} = 0.1$. Table A.1 displays the Euler errors and the relative errors of the optimal carbon taxes along the pre-tipping 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 \bar{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}^∞ norm for the worst case scenario (both γ and \bar{J} are larger),³⁶ if we treat the VFI solution as the true solution.

A.4 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

³⁶The 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}^∞ 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.1: Errors for Carbon Taxes in the First 100 Years

γ	\bar{J}	Euler Error		Relative Error	
		\mathcal{L}^∞	\mathcal{L}^1	\mathcal{L}^∞	\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 $a \times 10^{-n}$.

the deterministic problem. Thus to estimate Euler errors, we assume that the tipping event has not happened by time s , i.e., $J_s = 0$.

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 $\beta^s \lambda_s^{\mathbf{K}}$ be time- s shadow prices of (22), (23), and (A.8) respectively for the DSICE model with climate risk. Let

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

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

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

subject to the transition laws (22), (23), and (A.11), 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 $s+2$, i.e.,

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

subject to the transition laws (22), (23), and (A.8).

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

$$\begin{aligned}\lambda_s^K &= \beta \mathbb{E}_s \left\{ \lambda_{s+1}^K \left[1 - \delta + \frac{\partial \widehat{Y}_{s+1}}{\partial K_{s+1}} \right] + \beta \lambda_{s+1}^{\text{MAT}} \frac{\partial E_{s+1}^{\text{Ind}}}{\partial K_{s+1}} \right\} \\ \lambda_s^{\text{M}} &= \beta \mathbb{E}_s \left\{ \lambda_{s+1}^{\text{M}} \Phi_{\text{M}} + \left(\lambda_{s+1}^{\text{TAT}} \xi_1 \frac{\partial F_{s+1}}{\partial M_{\text{AT},s+1}}, 0, 0 \right) \right\} \\ \lambda_s^{\text{T}} &= \beta \mathbb{E}_s \left\{ \lambda_{s+1}^{\text{T}} \Phi_{\text{T}} + \left(\lambda_{s+1}^{\text{K}} \frac{\partial \widehat{Y}_{s+1}}{\partial T_{\text{AT},s+1}}, 0 \right) \right\} + \beta^2 (\Lambda_{s+2}, 0)\end{aligned}$$

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.10) conditional on $J_s = 0$.

Given the time- s information with $J_s = 0$, λ_{s+1}^{K} , λ_{s+1}^{M} , λ_{s+1}^{T} , \widehat{Y}_{s+1} , E_{s+1}^{Ind} , 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 $s+1$ ($J_{s+1} = \bar{J}$), 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.11). Let $\lambda_t := (\lambda_t^{\text{K}}, \lambda_t^{\text{M}}, \lambda_t^{\text{T}})^\top$.

The normalized Euler error on the pre-tipping path at time s is defined as

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

where the division is elementwise.

A.5 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, i.e.,

$$\max_{c_t, \ell_t, b_t} \mathbb{E} \left\{ \sum_{t=0}^{\infty} \left(\prod_{i=0}^t \beta_i \right) U(c_t, \ell_t) \right\} \quad (\text{A.12})$$

subject to the budget constraint

$$p_t c_t + \frac{b_t}{1 + r_t} = w_t \ell_t + b_{t-1} + T_t + \Pi_t \quad (\text{A.13})$$

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

$$1 = \mathbb{E}_t \left\{ \beta_{t+1} \frac{1 + r_t}{\pi_{t+1}} \frac{c_t}{c_{t+1}} \right\} \quad (\text{A.14})$$

and

$$w_t = p_t c_t \ell_t^\eta \quad (\text{A.15})$$

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

$$y_t = \left(\int_0^1 y_{i,t}^{\frac{\alpha-1}{\alpha}} di \right)^{\frac{\alpha}{\alpha-1}} \quad (\text{A.16})$$

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

$$y_{i,t} = y_t \left(\frac{p_{i,t}}{p_t} \right)^{-\alpha}. \quad (\text{A.17})$$

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

$$y_{i,t} = \ell_{i,t}. \quad (\text{A.18})$$

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 re-optimizing 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:

$$\max_{p_{i,t}} \mathbb{E}_t \left\{ \sum_{j=0}^{\infty} \left(\prod_{k=0}^j \beta_{t+k} \right) \lambda_{t+j} \theta^j (p_{i,t} y_{i,t+j} - w_{t+j} \ell_{i,t+j}) \right\} \quad (\text{A.19})$$

subject to the constraints $y_{i,t+j} = \ell_{i,t+j}$ from (A.18) and

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

from (A.17) by letting $p_{i,t+j} = p_{i,t}$. Here λ_t is the Lagrange multiplier of the budget constraint (A.13), satisfying

$$\lambda_t = \frac{1}{p_t c_t}, \quad (\text{A.20})$$

from the first-order conditions of the household problem (A.12). Let $\pi_{t,j} = p_{t+j}/p_t$. From (A.15), (A.20), and the first-order condition of the re-optimizing intermediate

firm problem (A.19), for any re-optimizing firm i we have

$$\frac{p_{i,t}}{p_t} \equiv q_t = \frac{\alpha \chi_{t,1}}{(\alpha - 1) \chi_{t,2}} \quad (\text{A.21})$$

where

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

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

$$\chi_{t,1} = y_t \ell_t^\eta + \theta \mathbb{E}_t \left\{ \beta_{t+1} \pi_{t+1}^\alpha \chi_{t+1,1} \right\} \quad (\text{A.22})$$

$$\chi_{t,2} = \frac{y_t}{c_t} + \theta \mathbb{E}_t \left\{ \beta_{t+1} \pi_{t+1}^{\alpha-1} \chi_{t+1,2} \right\} \quad (\text{A.23})$$

From (A.16) and (A.17), 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

$$q_t = \left(\frac{1 - \theta \pi_t^{\alpha-1}}{1 - \theta} \right)^{\frac{1}{1-\alpha}} \quad (\text{A.24})$$

From (A.17), (A.18) and the market clearing condition

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

we get

$$\begin{aligned}
v_{t+1} \equiv \ell_t/y_t &= \int_0^1 \left(\frac{p_{i,t}}{p_t} \right)^{-\alpha} di \\
&= (1 - \theta)q_t^{-\alpha} + \theta \int_0^1 \left(\frac{p_{i,t-1}}{p_t} \right)^{-\alpha} di \\
&= (1 - \theta)q_t^{-\alpha} + \theta\pi_t^\alpha \int_0^1 \left(\frac{p_{i,t-1}}{p_{t-1}} \right)^{-\alpha} di \\
&= (1 - \theta)q_t^{-\alpha} + \theta\pi_t^\alpha v_t
\end{aligned} \tag{A.25}$$

From $v_{t+1} = \ell_t/y_t$, (A.22) 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.14). From (A.23), the steady state of $\chi_{t,2}$ is

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

From (A.21) and (A.24), the steady state of $\chi_{t,1}$ is

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

with

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

and from (A.25) the steady state of v_t is

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

Therefore, from $v_t = \ell_t/y_t$ and (A.22), we get

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

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

$$c_t = (1 - s_g)y_t. \quad (\text{A.26})$$

Thus, (A.23) 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

$$r_t = \max(z_t, 0) \quad (\text{A.27})$$

with

$$z_t = (1 + r^*) \left(\frac{\pi_t}{\pi^*} \right)^{\phi_\pi} \left(\frac{y_t}{y^*} \right)^{\phi_y} - 1 \quad (\text{A.28})$$

where we choose $\phi_\pi = 2.5$ and $\phi_y = 0.25$. Thus, (A.14) can be rewritten as

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