Economics Department Working Paper Series

No. 53-2012

The Effect of Learning on Climate Policy under Fat-tailed Uncertainty

Richard S. J. Tol
Department of Economics, University of Sussex, Brighton, United Kingdom
Institute for Environmental Studies, Vrije Universiteit, Amsterdam, Netherlands
Department of Spatial Economics, Vrije Universiteit, Amsterdam, Netherlands
Email: R.Tol@sussex.ac.uk

In Chang Hwang
Institute for Environmental Studies, Vrije Universiteit, Amsterdam, The Netherlands

Frédéric Reynès
Institute for Environmental Studies, Vrije Universiteit, Amsterdam, The Netherlands
OFCE Sciences Po's Economic Research Centre, Paris, France
TNO - Netherlands Organisation for Applied Scientific Research, Delft, The Netherlands

Abstract: The effect of learning on climate policy is not straightforward when climate policy is concerned. It depends not only on the ways that climate feedbacks, preferences, and economic impacts are considered, but also on the ways that uncertainty and learning are introduced. Deep (or fat-tailed) uncertainty does matter for the optimal climate policy in that it requires more stringent efforts to reduce carbon emissions. However, learning may reveal thin-tailed uncertainty, weakening the case for emission abatement: learning reduces the stringency of the optimal abatement efforts relative to the no learning case even when we account for deep uncertainty. In order to investigate this hypothesis, we construct an endogenous (Bayesian) learning model with fat-tailed uncertainty on climate change and solve the model with stochastic dynamic programming. In our model a decision maker updates her belief on the total feedback factors through temperature observations each period and takes a course of action (carbon reductions) based on her belief. With various scenarios, we find that the uncertainty is partially resolved over time, although the rate of learning is relatively slow, and this materially affects the optimal decision: the decision maker with a possibility of learning lowers the effort to reduce carbon emissions relative to the no learning case. This is because the decision maker fully utilizes the information revealed to reduce uncertainty, and thus she can make a decision contingent on the updated information. In addition, with incorrect belief scenarios, we find

that learning enables the economic agent to have less regrets (in economic terms, sunk benefits or sunk costs) for her past decisions after the true value of the uncertain variable is revealed to be different from the initial belief.

JEL Classification: Q54

Key Words: Climate policy; deep uncertainty; fat-tails; Bayesian learning; integrated assessment; stochastic modeling; dynamic programming

1. Introduction

"The acquisition of information has value, which it would not have in a world of certainty." (Arrow, 1957: 524) Following this notion, economists have investigated the effects of learning on policy, including the irreversibility effect, the value of information, the optimal timing of action, the rate of learning, the direction of learning, and the cost of learning.[1] The answers to these questions, however, are not straightforward especially when climate policy is concerned. They depend not only on the ways that climate feedbacks, preferences, and economic impacts are considered, but also on the ways that uncertainty and learning are introduced.

The general framework for the problem of decision making under uncertainty and learning about climate change is as follows (Pindyck, 2000; 2002). In an economy where the impacts of climate change are uncertain with a possibility of learning, a decision maker encounters conflicting risks: a risk that stringent emissions control today turns out to be unnecessary ex post (sunk costs) and a risk that much stronger efforts are required in the future (sunk benefits). If there is no irreversibility to be considered, the problem becomes trivial since the decision maker can revise her actions as and when required. However, both the investment in emissions control and the accumulation of greenhouse gases are, at least partially, irreversible. In the presence of the irreversibility, the decision maker generally favors an option that preserves flexibility (Arrow and Fisher, 1974; Henry, 1974). As far as climate policy is concerned, since there are two kinds of counteracting irreversibility, the problem becomes complicated. The relative magnitude of the effects of the irreversibility determines the direction and the magnitude of the effect of learning on policy: the irreversibility related to carbon accumulation strengthens abatement efforts whereas the capital irreversibility lowers abatement.

Alternatively, we can think of the problem as an optimal experimentation with emissions in the framework of learning by doing (Arrow, 1961; Grossman et al., 1977). The decision maker confronted with uncertainty and a possibility of learning about the impacts of climate change can be interpreted as a Bayesian statistician who experiments with a level of carbon emissions to gain information about uncertainty. The more emissions (in turn, higher temperature increases) are more informative in the sense that it provides more precise information about uncertain parameters such as climate sensitivity. However, the acquisition of information comes at an implicit cost: higher emissions induce consumption losses (via increased temperature). As a result, the decision maker should choose an optimal level of emissions by comparing gains and losses from the acquisition of information.

In the literature, the possibility of learning generally changes the near-term policy towards higher emissions relative to the case where the uncertainty is not reduced (for the summary of the literature see Ingham et al., 2007). One reason for this is that constraints such as the non-negativity of carbon emissions rarely bite in climate change models (Ulph and Ulph, 1997).[2] In addition, even if they do bind, the effect of the (partially) irreversible accumulation of carbon stocks is smaller than the effect of the irreversible capital investment in emissions control (Kolstad, 1996).[3] If

---

[1] The relevant literature for these issues on climate change is as follows: the irreversibility effect (Kolstad, 1996a; Ulph and Ulph, 1997), the value of information (Peck and Teisberg, 1993; Nordhaus and Popp, 1997), the optimal timing of action (Pindyck, 2000, 2002; Guillerminet and Tol, 2008), the rate of learning (Kolstad, 1996b; Kelly and Kolstad, 1999; Leach, 2007; Webster et al., 2008), and the direction of learning (Oppenheimer et al., 2008).

[2] Regarding this, Webster (2002) argues that if the non-negativity matters, the effect of the irreversible accumulation may outweigh the effect of the irreversible investment.

[3] Some literature finds the case where the irreversibility constraint (i.e. nonnegative emissions) binds with various methods, including the alternative parameterization of some critical equations (Ulph and Ulph, 1997; Webster, 2002), the introduction of catastrophic events (Keller et al., 2004), and the presence of stringent climate targets (Webster et al., 2008). However, their models do not find a case where sunk benefits outweigh sunk costs.

we think of the results in the framework of learning by doing, these results imply that the experimentation with more carbon emissions is more informative in the sense that the decision maker can attain more utility from the experimentation (Blackwell, 1951).

In the current paper we add another perspective on the literature. That is, where previous papers studied thin-tailed distributions, we here focus on fat-tailed ones.[4] Fat-tailed uncertainty (or deep uncertainty) may lead to different results since the marginal damage costs of climate change become far larger, if not arbitrarily large, under deep uncertainty (Tol, 2003; Weitzman, 2009). Consequently, sunk benefits may outweigh sunk costs under the possibility of learning, and this may change climate policy in favor of more stringent efforts to reduce emissions compared to the no learning case.

However, learning may of course reveal thin-tailed uncertainty (about social welfare), weakening the case for emissions abatement.[5] That is, learning reduces the optimal level of emissions control rates even when we account for deep uncertainty. We develop a dynamic model on climate change introducing deep uncertainty and learning to investigate this hypothesis. Learning in our model is endogenous: the decision maker updates her belief about a parameter, expressed in a probability distribution, by the acquisition of information. Our approach on endogenous learning is not new in the literature on the economics of climate change. For instance, Kelly and Kolstad (1999) introduce uncertainty about a climate parameter (linearly related to climate sensitivity) into the DICE model, and then investigate the expected learning time. Leach (2007) follows a similar model and approach, but introduces an additional uncertainty on climate parameters. Webster et al. (2008) investigate the effect of learning on the near term policy using the DICE model with a discrete four-valued climate sensitivity distribution and exogenous learning. In the second part of their paper, they investigate the time needed to reduce the uncertainties about climate sensitivity and heat uptake by using a simplified climate model. They incorporate fat-tailed uncertainty and Bayesian learning into the model but their model does not analyze policy. The current paper is different from the literature in that we investigate the effect of learning on climate policy using an endogenous learning model incorporating fat-tailed uncertainty.[6] In specific, unlike the literature, we investigate the optimal level of emissions control under the possibility of endogenous learning as well as the rate of learning. In addition, through various scenarios such as the incorrect belief cases, we compare the benefits of learning.

The paper proceeds as follows. Section 2 describes the model and computational methods. We revise the DICE model (Nordhaus, 2008) to represent (deep) uncertainty and endogenous learning about the equilibrium climate sensitivity through the framework of feedback analysis (Hansen et al, 1984; Roe and Baker, 2007). We solve the infinite time horizon model using the method of dynamic programming. Section 3 presents the posterior distribution of the total feedback factors and the climate sensitivity. We compare the rate of learning with various sensitivity analyses. Section 4 illustrates the effect of learning on climate policy. We investigate the case where the initial belief on the climate sensitivity is correct in the sense that the true value of the climate sensitivity turns out to be the same as the expected value of the initial belief of the decision maker. Then we compare the results with the other cases where the initial belief of the decision maker turns out to be incorrect. Section 5 investigates the value of learning. Finally, Section 6 concludes.

---

[4] We follow the definition of fat tails as follows: "a PDF [probability density function] has a fat tail when its moment generating function is infinite - that is, the tail probability approaches 0 more slowly than exponentially." (Weitzman, 2009: 2)

[5] Notice that the posterior distributions of climate sensitivity always have fat tails in the specification of our learning model since the variance of the total feedback factors does not become zero (asymptote) (see Section 2). However, the uncertainty on social welfare may have thin tails as the variance of the total feedback factors approaches zero.

[6] Kelly and Tan (2012) do a similar analysis to ours in that they also incorporate fat-tailed uncertainty and learning into a climate-economy model. However, they do not present any policy analysis yet in their preliminary working paper.

2. Model and methods

2.1. The revised DICE model

We revise the DICE model to introduce uncertainty and learning. There are several distinct differences with the original DICE model. First, the current model incorporates (deep) uncertainty. The key uncertain parameter in the model is the equilibrium climate sensitivity. Second, the probability density function (PDF) of the total feedback factors, and thus the PDF of the climate sensitivity, changes over time through temperature observations.[7] As a result, the mean and the variance of the total feedback factors become endogenous state variables (see below). Third, we apply a simple variant of energy balance model for atmospheric temperature evolution to represent the total feedback factors explicitly. Furthermore, we introduce stochastic shocks into the temperature equation for Bayesian updating on the PDF of the total feedback factors (Equation (5)). Temperature shocks reflect uncertainty about observational errors, model's biases to match observations, and the natural variability (Webster et al., 2008). Fourth, we consider an infinite time-horizon problem and apply a solution method suitable for the endogenous learning model (see Section 2.3). To this end, we modify some specifications of the DICE model for computational convenience. In specific, we apply a simple one-box model for the evolution of carbon stocks (Equation (4)). Our model does not consider a backstop technology for the abatement costs function. In addition, unlike the DICE model, the time period t in our model is annual and thus some parameter values including δ_M and the parameters in Equation (5) and (6) are adjusted. There is an upper limit of 6,000 GtC for the accumulated carbon emissions in DICE but the current model does not include this constraint. Finally, the utility of consumption is not weighted by the level of population in the objective function (Equation (1)).[8]

As with the original DICE model, the (partial) irreversibility is represented by the depreciation rates of carbon stocks and capital stocks (Equation (3), (4)) together with the non-negativity of emissions ($\mu_t \leq 1$). We assume that learning is costless.

The model is specified as follows: [9]

$$\max_{\mu_t, I_t} E_0 \sum_{t=0}^{\infty} \beta^t u(C_t, L_t) = E_0 \sum_{t=0}^{\infty} \beta^t \frac{(C_t/L_t)^{1-\alpha}}{1-\alpha} \tag{1}$$

subject to

---

[7] The equilibrium climate sensitivity is a measure of the responsiveness of climate system to radiative forcing. It denotes how much atmospheric temperature changes when carbon concentration doubles. The feedback factor refers to the impact of a physical factor such as water vapor and cloud on radiative forcing in a way that amplifies the response of climate system (Hansen et al., 1984).

[8] Simulations with population-weighted objective function produce the similar results qualitatively as the results in this paper.

[9] Unless otherwise noted, the parameter values and the initial values for the state variables (in the year 2005) are the same as in DICE 2007. The exceptions are the elasticity of marginal utility of consumption (α) and the exponent of the abatement cost function (θ_2). We use 1 for α and 2 for θ_2 (instead of 2 and 2.8, respectively in DICE 2007) because of the computational convenience in deriving the first-order conditions (see Section 2.3 and Appendix A). The value of 1 for α is in the plausible range of the elasticity of marginal utility of consumption (Gollier 2000). A logarithmic utility function (α=1) is used in the older version of DICE, for instance. The lower α implies the less concavity of utility function, and thus the less aversion to risk and the more aversion to inequality. The marginal abatement costs are linear in our model since we use 2 for θ_2. Ulph and Ulph (1997) apply the quadratic abatement costs function, among others.

$$C_t = \left(\frac{1-\Lambda_t}{1/\Omega_t}\right)Q_t - I_t = \left[\frac{1-\theta_1\mu_t^{\theta_2}}{1+\kappa_1 T_{AT_t} + \kappa_2 T_{AT_t}^{\kappa_3}}\right] A_t K_t^\gamma L_t^{1-\gamma} - I_t \tag{2}$$

$$K_{t+1} = (1-\delta_k)K_t + I_t \tag{3}$$

$$M_{t+1} = (1-\mu_t)\sigma_t A_t K_t^\gamma L_t^{1-\gamma} + E_{LAND_t} + (1-\delta_M)(M_t - M_b) + M_b \tag{4}$$

$$T_{AT_{t+1}} = (\zeta_1 f + \zeta_2)T_{AT_t} + \zeta_3\ln(M_t/M_b) + \zeta_4 T_{LO_t} + \zeta_5 RF_t + \varepsilon_{t+1} \tag{5}$$

$$T_{LO_{t+1}} = T_{LO_t} + \xi\{T_{AT_t} - T_{LO_t}\} \tag{6}$$

$$\overline{f_{t+1}} = \frac{\overline{f_t} + \zeta_1 T_{AT_t}[\zeta_1 f T_{AT_t} + \varepsilon_{t+1}](v_t/v_\varepsilon)}{1 + \zeta_1^2 T_{AT_t}^2(v_t/v_\varepsilon)} \tag{7}$$

$$v_{t+1} = \frac{v_t}{1 + \zeta_1^2 T_{AT_t}^2(v_t/v_\varepsilon)} \tag{8}$$

where E_t is the expectation operator given information at point in time t, u is the instantaneous utility of per capita consumption, C_t is consumption, L_t is the level of population (exogenous), A_t is the total factor productivity (exogenous), Λ_t ($=\theta_1\mu_t^{\theta_2}$) is the abatement cost function, Ω_t ($=1/(1 + \kappa_1 T_{AT_t} + \kappa_2 T_{AT_t}^{\kappa_3})$) is the damage function, Q_t ($=A_t K_t^\gamma L_t^{1-\gamma}$) is the production function, μ_t is the emissions control rate, I_t is the gross investment, K_t is the capital stock, M_t is the carbon stock in the atmosphere, T_{AT_t} is the atmospheric temperature deviation (from 1900), T_{LO_t} is the lower ocean temperature deviation (from 1900), f is the total feedback factors normally distributed with mean $\overline{f_t}$ (the initial value is 0.65 following Roe and Baker (2007)) and variance v_t (the initial value is 0.13^2 following Roe and Baker (2007)), ε_t is the stochastic temperature shocks normally distributed with mean 0 and variance v_ε ($=0.11^2$ following Leach (2007)), σ_t is the emission-output ratio (exogenous), RF_t is the radiative forcing from non-CO_2 gases (exogenous), E_{LAND_t} is carbon emissions from the sources other than energy consumption (exogenous), α is the elasticity of marginal utility of consumption (=1), $\beta = 1/(1 + \rho)$ is the discount factor, ρ is the pure rate of time preference (=0.015), γ is the elasticity of output with respect to capital (=0.3), δ_k (= 0.1) and δ_M (=0.00833 following Leach (2007)) are the depreciation rates, respectively of the capital stock and the carbon stock, M_b is the pre-industrial carbon stocks in the atmosphere (=596.4GtC), κ_1 (=0), κ_2 (=0.0028388), κ_3 (=2), θ_1 (=0.0561), θ_2 (=2), ξ (=0.005), ζ_1 (=0.06967), ζ_2 (=0.92373), ζ_3 (=0.12061), ζ_4 (=0.0066), ζ_5 (=0.022) are parameters.

2.2. Bayesian learning

The temperature response model in the current paper is a simple variant of the energy balance model of Baker and Roe (2009). In their model, the temperature of the ocean mixed layer obeys the following equation: $\rho C_p h\frac{dT_t}{dt} + \frac{T_t(1-f)}{\lambda_0} - \kappa\frac{\partial T_{Z_t}}{\partial z}|_{z=0} = RF_t$, where ρ is the density, C_p is the specific heat, h is the depth of the mixed layer, κ is the thermal conductivity, z is the ocean depth below the mixed layer, and T_Z is the temperature at depth z. With a two-box (mixed layer and deep ocean) simplification, the above equation reduces to Equation (5) (without stochastic shocks) and Equation (6) (Marten, 2011).[10]

The decision maker updates her belief on the total feedback factors through temperature observations each time-period in our model. This representation enables us to consider deep uncertainty because the climate sensitivity is related to the total feedback factors as in Equation (9) (Roe and Baker, 2007). In this relation together with the

---

[10] Equation (5) can also be derived from the climate model in DICE using the relationship as in Equation (9).

assumption on the distribution of the total feedback factors described below, the climate sensitivity has a highly skewed distribution to the upper end (see Figure 1).[11]

$$CS = CS_0/(1 - f) \qquad (9)$$

where CS is the equilibrium climate sensitivity, CS_0 (=1.2°C following Roe and Baker (2007)) is the reference climate sensitivity in the grey-body planet (no feedback effects).

We introduce stochastic shocks into the atmospheric temperature equation. The stochastic shocks are assumed to have a normal distribution with mean 0 and a constant variance. Although the shocks are independent, Equations (5) and (6) imply a first-order autoregressive model (Kelly and Kolstad, 1999). In this representation, the decision maker expects that the atmospheric temperature in the next period is determined by the following equation: $(\zeta_1 E_t[f] + \zeta_2)T_{AT_t} + \zeta_3\ln(M_t/M_b) + \zeta_4 T_{LO_t} + \zeta_5 RF_t$. However, the actual realization of the atmospheric temperature is determined by the true value of the total feedback factors (which is not known to the decision maker ex ante with certainty: parametric uncertainty) and the realized stochastic shocks (which is never known to the decision maker before realization: stochasticity). Put differently, there is deviation between the decision maker's expectation and observations, which leads to the modification of the prior belief. The decision maker in our model obtains the posterior distribution of the total feedback factors by the Bayes Rule as follows.

$$p(f|T_{AT}) \propto p(T_{AT}|f) \times p(f) \qquad (10)$$

where $p(f)$ is the prior belief on the total feedback factors, $p(T_{AT}|f)$ is the likelihood function of the observations given f, and $p(f|T_{AT})$ is the posterior belief.

We use an expert prior for the distribution of the total feedback factors: namely, the normal distribution of Roe and Baker (2007). The normal priors have some advantages over the other priors. First, the posterior calculated from the normal prior is also normally distributed, provided that the likelihood function is also normal. In this case, it is easy to calculate the posterior just by investigating the posterior mean and the variance (Cyert and DeGroot, 1974). Second, as Annan and Hargreaves (2011) point out, uniform priors usually used in a Bayesian analysis assign too much probability to extreme parameter values (say, climate sensitivity of 10°C/2xCO$_2$) beyond the current scientific knowledge.[12] This assignment may dominate the calculation of the expected damage costs.

The resulting posterior has the normal distribution with mean $\overline{f_{t+1}}$ and variance v_{t+1} as in Equation (7) and (8).[13] In the subsequent period, the decision maker uses the previously calculated posterior as the prior. In this way, the decision maker updates her belief every time-period.

The variance of the total feedback factors decreases over time. Put differently, the acquisition of information in our model always increases the precision of the decision maker's belief. Furthermore, as the variance gets smaller, the

---

[11] Kelly and Kolstad (1999) and Leach (2007) apply the similar approach to ours, but the decision makers in their models update beliefs on the (normally distributed) climate parameter so they do not account for fat-tails. In addition, updating the PDF of the total feedback factors is more appropriate than updating the PDF of the climate sensitivity in that the observable parameters are not the climate sensitivity but the feedback factors (Allen et al., 2005).

[12] For instance, compare the peer-reviewed climate sensitivity distribution in Solomon et al. (2007) and a uniform distribution such as U(0, 20). The probability assigned to climate sensitivity of 4.5°C/2xCO$_2$ or more in each case is 15% and 77.5%, respectively.

[13] These equations are derived from a direct application of Bayes' Theorem (Equation (10)) with the above mentioned assumptions on the likelihood function and the prior. For more on Bayesian updating methods, see Lee (2012).

mean approaches to the true value of the total feedback factors, on average (Equation (7)). That is, there is no 'negative' learning (Oppenheimer et al., 2008) in this model. One thing to add is that, although the variance becomes smaller as the temperature observations accumulate, the climate sensitivity has still fat tails, unless the variance becomes equal to 0, in the sense that its density diminishes more slowly than exponentially (Kelly and Tan, 2012).

2.3. Computational methods

There are generally two kinds of methods for solving a learning model numerically: stochastic optimization and dynamic programming. The first one is to consider possible states of the world on parameter values of interest with corresponding probability distributions, and solve for the optimal time path of policy variables that maximize the expected value of the objective function over a *finite* time horizon (e.g. Kolstad, 1996; Webster et al., 2008). The second one is to formulate the problem in a recursive way (a functional equation) and then solve the problem over an *infinite* time horizon (e.g. Kelly and Kolstad, 1999; Leach, 2007). The current paper takes the second approach: dynamic programming.[14] We illustrate the general approach of the current paper below and the solution methods in detail are given in Appendix A.

We formulate the problem as the following. The Bellman equation is:

$$W(\boldsymbol{s}_t, \boldsymbol{\theta}_t) = \max_{\boldsymbol{c}_t}[u(\boldsymbol{s}_t, \boldsymbol{c}_t, \boldsymbol{\theta}_t) + \beta E_t W(\boldsymbol{s}_{t+1}, \boldsymbol{\theta}_{t+1})] \tag{11}$$

where $W(\boldsymbol{s}_t, \boldsymbol{\theta}_t)$ is the value of the maximized objective function, Equation (1), starting from period t (value function), $\boldsymbol{c}$ is the vector of control variables (μ, I), $\boldsymbol{s}$ is the vector of state variables (K, M, T_{AT}, T_{LO}, $\bar{f}$, v), and $\boldsymbol{\theta}$ is the vector of uncertain variables (f, ε).

Then we approximate value function W with flexible basis function ψ having a specific analytic form such as polynomials or a logarithmic function.

$$W(\boldsymbol{s}_t, \boldsymbol{\theta}_t) \approx \psi(\boldsymbol{s}_t, \boldsymbol{\theta}_t; \boldsymbol{b}) = \sum_i \psi_i(\boldsymbol{s}_t, \boldsymbol{\theta}_t; \boldsymbol{b}_i) \tag{12}$$

where ψ is the basis function, $\boldsymbol{b}$ is the vector of coefficients for the basis function.

Following this way, we change the maximization problem into the regression problem: finding $\boldsymbol{b}$ of the basis function that minimizes the approximation errors. The algorithm for finding $\boldsymbol{b}$ is summarized as follows. First, choose an initial guess on $\boldsymbol{b}$. Second, simulate a time series satisfying the first order necessary conditions (see Appendix A), the initial conditions (the initial values for the state variables), and the transitional equations (Equation (3) – (8)) with the initial guess.[15] Third, compare the left-hand side and the right-hand side of Equation (11), and then iterate simulations with updated $\boldsymbol{b}$ if the difference is higher than the pre-specified tolerance level (see Appendix A for the updating rule and the stopping rule).[16] Fourth, if the difference meets the stopping rule, stop the iteration. This is the method proposed by Maliar and Maliar (2005), and it reduces the computational burden in that

---

[14] For the general discussion on recursive methods and applications of computational methods for dynamic programming, see Stocky and Lucas (1989), Rust (1996), and Judd (1998).

[15] If the initial guess is chosen, we can calculate the control variables from the first order necessary conditions, the initial values for state variables, and the transitional equations. In this way, the resulting time series depend on $\boldsymbol{b}$.

[16] For the calculation of the conditional expectation, we use the Gauss-Hermite quadrature method (Judd, 1998).

it searches for a solution in a set satisfying the necessary conditions (an ergodic set).[17] Once found, the solution b attains the fixed point of the Bellman equation, which is the unique solution for the maximization problem (Equation (1)) (Stocky and Lucas, 1989).

3. The resolution of uncertainty

3.1. Climate sensitivity distribution

According to the updating procedure presented in the previous section, the belief of the decision maker on the true value of the total feedback factors changes as the (annual) temperature observations accumulate. Figure 1 shows the evolution of the parameter values of the total feedback factors distribution, the corresponding distributions of the climate sensitivity, and the tail probabilities. Following the current scientific knowledge on the total feedback factors distribution (Roe and Baker, 2007), we present the case where the true value of the total feedback factors is 0.65 (except the top left panel: we use 0.70 there for an illustration purpose. From the top left panel, we can observe how the expected value of the decision maker approaches the true value as observations accumulate over time) and the initial belief on the total feedback factors is the normal distribution with mean 0.65 and standard deviation 0.13. Regarding other assumptions for simulation, see Appendix A. Considering random realizations of the temperature shocks, throughout the paper for the learning case, we present the average of 10,000 Monte Carlo simulations.

Since the mean changes, we consider the coefficient of variation -- the standard deviation divided by the mean -- as a measure of uncertainty. We define learning as a decrease in the coefficient of variation as Webster et al. (2008) do.[18] In this definition, the decision maker learns every year as we can see from the top left panel. As argued in the previous section, the mean approaches the true value of the total feedback factors and the variance decreases as the temperature observations accumulate over time. The coefficient of variation also falls over time. The top right and the bottom left panel show the climate sensitivity distributions corresponding to the baseline case: the true value of the total feedback factors is 0.65 and the prior belief on the total feedback factors is normal distribution with mean 0.65 and standard deviation 0.13. The density on the tails becomes much smaller as time goes by, and thus the precision (defined as the reciprocal of variance) of the belief increases. The bottom right panel illustrates the (right) tail property of the climate sensitivity distribution of the reference case. As expected, the tail probability decreases as learning takes place. For instance, the probability that climate sensitivity is above 6°C is 0.12 in 2005, but it decreases to 0.076 and 0.014 in 2050 and 2100, respectively. Since the tail probability gets smaller over time, the value of climate sensitivity below which a significant percent lie (a percentile) also decreases.[19] For instance, 95th percentile of climate sensitivity decreases from 8.8°C in 2005 to 5.0°C in 2100. One of the important questions in climate science is whether we can provide constraints (or an upper bound) on climate sensitivity (for further discussion on this issue, see Knutti et al., 2002; Annan and Hargreaves, 2006). If we use a percentile to impose an upper bound on climate sensitivity, the bottom right panel gives useful information. For instance, based on 95th (respectively, 99th) percentile, it takes 65 years (resp., 100 years) to set an upper bound on climate sensitivity below 6°C.

---

[17] For further discussion on this issue, see Judd et al. (2011)

[18] Kelly and Kolstad (1999) and Leach (2007) define learning as the estimated mean approaching its true value. Learning takes place in their model when the mean of an uncertain variable becomes statistically close to the pre-specified true value (with significance level 0.05).

[19] Of course we are concerned about a high percentile, say 95th percentile in this sentence. Note that a low percentile (say 5th percentile) increases since the left tail also shrinks over time.

3.2. The rate of learning

The rate of learning is as important as the magnitude of learning since slow learning may lead to incapability for us to take appropriate actions on time because of the irreversibility. Especially when we consider the possibility of (discontinuous) climate catastrophes such as a collapse of the West-Antarctic Ice Sheet (Guillerminet and Tol, 2008) and the thermohaline circulation collapse (Keller et al., 2004), we should put more importance on the rate of learning. Furthermore, fast learning enables more efficient allocation of resources. In order to investigate the rate of learning, we normalize the coefficient of variation so that the initial value (in 2005) is equal to 1. Figure 2 shows the results. We present some sensitivity analyses together with the baseline case. First of all, the time needed to reduce the uncertainty (about the total feedback factors) is relatively long in the baseline case: it takes 67 years (respectively, 164 years) to reduce the uncertainty by 50% (resp., 90%).[20] Learning is faster (respectively, slower) when the true value of the total feedback factors is higher (resp., lower) than the initial belief. This is intuitive in that the higher total feedback factors imply the higher temperature increases, resulting in the lower variance (see Equation (5) and (8)). In the similar fashion, the rate of learning is increasing in emissions. In order to verify this, we perturb emissions by one unit (1GtC) per year from the optimal solution.[21] The top right panel shows the results. A unit increase (respectively, decrease) in emissions from the optimal path reduces (resp., increases) the uncertainty. The more deviations in emissions are, the higher deviations in the rate of learning are.

The rate of learning is highly sensitive to the variance of the temperature shocks. For instance, if the standard deviation of temperature shocks is reduced to 0.05 (the baseline case is 0.11), the time needed to reduce the uncertainty is almost halved relative to the baseline case: it takes 32 years (respectively, 88 years) to reduce the uncertainty by 50% (resp., 90%). The bottom right panel illustrates the sensitivity of the rate of learning to the initial level of uncertainty. Since the coefficient of variation differs from case to case, we do not normalize the coefficient of variation in this panel. We observe that the coefficient of variation converges to a certain (low) level during the late 22nd century. In other words, differences in the initial level of uncertainty become irrelevant after 150 years or so in our learning model. This is because the rate of learning is higher in a more uncertain world. Nevertheless, there are substantial differences in uncertainty in the near future.

4. The effect of learning

We investigate the effect of learning on climate policy. To this end, we compare 3 cases: (1) Deterministic, (2) Uncertainty (no learning), and (3) Learning. The deterministic case refers to the case where the decision maker does not consider uncertainty. The uncertainty case refers to the case where the decision maker accounts for uncertainty, but her belief remains unchanged. The information may accumulate, but the decision maker simply ignores the possibility of learning or chooses to ignore the information gathered. Finally, the belief of the decision maker is subject to change in the learning case. The decision maker fully utilizes the information acquired from the temperature observations so that she can make a decision contingent on the information. Throughout this section, the initial belief on the total feedback factors is represented by the normal distribution with mean 0.65 and standard deviation 0.13. The true value of the total feedback factors is 0.65 in Section 4.1.

---

[20] The time needed to reduce the uncertainty is, of course, sensitive to the specification of the model, especially to the assumptions on the prior and the likelihood function. However, notice that the results we present in this section are based on the current scientific knowledge on the distribution of the total feedback factors following Roe and Baker (2007).

[21] We simulate the model with an additional one unit of exogenous carbon emissions and then find the solution. For simulation, we use the solution *b* of the reference case as the initial guess. The other specifications are the same as the reference case.

In our learning model, atmospheric temperature evolves according to the true value of the total feedback factors, but the decision maker conducts a course of action according to her belief. Then what if our belief on uncertain variables turns out to be incorrect? In order to answer to this question we simulate our model with an assumption that the true value of the total feedback factors, ex post, turns out to be different from the decision maker's initial belief in Section 4.2. We consider two scenarios. First, the true value of the total feedback factors turns out to be higher than the initial belief ('H' case: $f=0.70 > E_0 f =0.65$). Second, the true value turns out to be lower than the initial expectation ('L' case: $f =0.60 < E_0 f =0.65$).

4.1. Correct belief

Figure 3 summarizes the effect of learning, which is generally consistent with the literature as we briefly introduced in Section 1. First, fat-tailed uncertainty greatly increases the abatement efforts relative to the deterministic case. This is because the uncertainty model considers the less probable but the more dismal future as well as the most probable (mode) or the expected state of the world. For instance, the optimal emissions control rate around the year 2030 reaches 50% and the irreversibility constraints (nonnegative emissions) start to bind after the year 2180 when we account for fat tailed uncertainty. Whereas in the deterministic case, the optimal level of emissions control rates in 2030 is below 30%. Second, the possibility of learning reduces the abatement efforts relative to the uncertainty case. Although the atmospheric temperature increases more in the learning case than in the uncertainty case (see the right panel, this is because carbon emissions are greater in the learning case as a result of lower emissions control rates), the decision maker attains (slightly) more consumption (in turn, utility) from the learning case. This implies that the experimentation with more emissions (or learning) is beneficial to the decision maker.

4.2. Incorrect belief

Figure 4 illustrates the results for the case of incorrect belief. The first thing we observe is that the optimal decision changes a lot according to the true value of the total feedback factors. Especially in the uncertainty case the emissions control rates are very sensitive to the true values, and thus the resulting temperature increases and the consumption vary a lot more relative to the learning case. Compared to the uncertainty case, the changes in optimal emissions, temperature deviations, and consumption are small in the learning case. This is because learning enables the decision maker to adjust her actions according to the information revealed. Second, the possibility of learning lowers the optimal level of emissions control rates in the incorrect belief case as in the correct belief case. Third, as one would expect, the optimal control rates are lower (respectively, higher) when the true value of the total feedback factors turns out to be lower (resp., higher). Atmospheric temperature deviations change in the opposite direction of the control rates. Fourth, the decision maker attains higher consumption when the true value of the total feedback factors turns out to be lower than expected.

5. The benefits of learning
In this section we investigate the benefits of learning. Table 1 illustrates the optimal carbon tax and the net present value of utility of each scenario in Figure 4. The optimal carbon tax is much higher in the uncertainty case than in the deterministic case and it is decreased in the learning case. When the true value of the uncertain variable turns out to be higher (respectively, lower) than expected the gain from learning becomes higher (resp., lower). We also observe that learning is valuable in that it increases social welfare. The net present value of utility (or social welfare to be maximized by the decision maker) is increased when we account for the possibility of learning. In addition, social welfare is higher (respectively, lower) when the true value turns out to be lower (resp., higher). The only exception is UNC_L case, where the decision maker does not account for the possibility of learning and the true value turns out to be lower than her belief. In this scenario, social welfare is higher than in the learning case

(LRN_L). One of the possible reasons for this is that 1) investment is higher in the learning case than in the uncertainty case (see Appendix B) and 2) as we can see from Figure 4, there are no significant differences in abatement efforts and temperature increases between UNC_L and LRN_L cases (see the green lines in top and bottom panels), hence no significant differences in abatement costs and damage costs (see Figure 5). Since consumption is calculated as production net of investment, damage costs, and abatement costs (see Equation (2)), consumption is higher in UNC_L case than in LRN_L case, and hence the higher utility in the uncertainty case. However, the gross income defined as the sum of consumption and investment (or the gross production net of abatement costs and damage costs) is higher in LRN_L case than in UNC_L case (see Appendix B).

In order to see the value of learning in a different perspective, let us suppose that the decision maker in the uncertainty case somehow choose to change her strategy about learning in a specific time period, say in the year 2100. That is, the decision maker starts to update her belief after the year 2100 based on temperature observations. With this assumption, we can calculate sunk benefits or sunk costs of the decision maker's past decisions: the difference in the total costs (sum of damage costs and the abatement costs) between the uncertainty case and the learning case represents the benefits of learning. Figure 5 shows the results. Considering the differences in the gross production and investment in each case, we present the costs as a fraction of the gross production of each case. As illustrated in the previous section, the optimal emissions control rates are lower in the learning case than in the uncertainty case, and thus the abatement costs are lower but the damage costs are higher in the learning case. The total costs are lower in the learning case than in the uncertainty case. For instance, the total costs are 0.26% point (as a fraction of the gross world output) lower in the learning case than in the uncertainty case in 2100. When the true value of the total feedback factors turns out to be higher (respectively, lower) than expected, the decision maker will find that emissions control rates have been unnecessarily lower (resp., higher) than required, but she cannot revise her past actions. That is, there are sunk benefits in H case and sunk costs in L case. The benefits of learning (or penalties for no-learning) become higher (respectively, lower) when the true value of the uncertain variable turns out to be higher (resp., lower) than the initial belief. Those values in 2100 are 2.10% point and 0.05% point (as a fraction of the gross world output) in H case ($f=0.70 > E_0 f=0.65$) and L case ($f=0.60 < E_0 f=0.65$), respectively.

6. Conclusion

We constructed an endogenous (Bayesian) learning model with a fat-tailed uncertainty on climate change and solved the model with a stochastic dynamic programming. In our model the decision maker updates her belief on the total feedback factors through temperature observations and takes a course of action (carbon reductions) each period based on her belief. With various scenarios, we find that the uncertainty is partially resolved over time, although the rate of learning is relatively slow, and this materially affects the optimal decision. Consistent with the literature, the decision maker with a possibility of learning lowers the effort to reduce carbon emissions relative to the no learning case (or the uncertainty case). This is because the decision maker fully utilizes the information revealed to reduce uncertainty, and thus she can make a decision contingent on the updated information. In addition, with incorrect belief scenarios, we find that learning enables the economic agent to have less regret (in economic terms, sunk benefits or sunk costs) for the past decisions after the true value of the uncertain variable is revealed to be different from the initial belief. The optimal decisions in the learning case are less sensitive to the true value of the uncertain variable than those in the uncertainty case. The reason is that learning lets uncertainty converge to the true value of the state in a sense that the variance approaches 0 as information accumulates. Deep uncertainty does matter for the optimal climate policy in that it requires more stringent efforts to reduce emissions. However, learning reduces such effect of deep uncertainty on climate policy. As one learns more, the effect of uncertainty becomes less.

Finally, we raise some caveats and suggest further researches. First, our model does not take into account the possibility of negative learning. Indeed, as Oppenheimer et al. (2008) argued, learning does not necessarily converge to the true value of the uncertain variable. The effect of the negative learning will be different from the results of this paper, of course. Since it is beyond the scope of this paper, we refer this to further researches. Second, for simplicity, we assume that learning is costless in this paper, but in reality learning comes at a cost. The value and the rate of learning depend on the cost of learning as well as on the benefit of learning. The inclusion of the cost of learning may complicate the model, but the main implications of this paper will hold unless learning costs more than it earns. Third, we only consider learning from temperature observations. In reality, there are more active forms of leaning such as research and development (Kolstad, 1996). An active learning model incorporates the optimal decision on activities such as R&D investment for reducing uncertainty, which is an important issue that should be considered in further research. Fourth, we use the logarithmic utility function and the quadratic abatement costs function for computational convenience. Since a climate-economy model includes many nonlinear functions, it is not easy to derive the optimal policy rules satisfying the first order conditions, which is a necessary step for the solution method of this paper. Approximations of such nonlinear functions to tractable functional forms may help apply the solution method.

References

Allen M., N. Andronova, B. Booth et al., 2006. Observational constraints on climate sensitivity. Cambridge University Press, pp. 281-289.

Arrow K. J., 1962. The economic implications of learning by doing. The review of economic studies, 155-173.

Baker M. B. and G. H. Roe, 2009. The shape of things to come: why is climate change so predictable? Journal of Climate, 22**:**4574-4589.

Bellman R., 1954. The theory of dynamic programming. Bull. Amer. Math. Soc, 60**:**503-515.

Blackwell D. and J. Neyman, 1951. Comparison of experiments. Proceedings of the second Berkeley, 1**:**415-418.

Cyert R. M. and M. H. DeGroot, 1974. Rational expectations and Bayesian analysis. The Journal of Political Economy, 82**:**521-536.

Gollier C., 2000. The economics of risk and time, The MIT press.

Ingham A., J. Ma, and A. Ulph, 2007. Climate change, mitigation and adaptation with uncertainty and learning. Energy Policy, 35**:**5354-5369.

Judd K. L., 1998. Numerical methods in economics, The MIT press.

Judd K. L., L. Maliar, and S. Maliar, 2011. Numerically stable and accurate stochastic simulation approaches for solving dynamic economic models. Quantitative Economics, 2**:**173-210.

Keller K., B. M. Bolker, and D. F. Bradford, 2004. Uncertain climate thresholds and optimal economic growth. Journal of Environmental Economics and Management, 48**:**723-741.

Kelly D. L. and C. D. Kolstad, 1999. Bayesian learning, growth, and pollution. Journal of Economic Dynamics and Control, 23**:** 491-518.

Kelly D. L. and C. D. Kolstad, 2001. Solving infinite horizon growth models with an environmental sector. Computational Economics, 18**:**217-231.

Kelly D. L. and Z. Tan, 2012. Learning, Growth and Climate Feedbacks Does the Fat Tail Wag the Dog?

Kolstad C. D., 1996. Learning and stock effects in environmental regulation: the case of greenhouse gas emissions. Journal of Environmental Economics and Management, 31**:**1-18.

Lee,P.M., 2012. Bayesian statistics: an introduction. Wiley.

Leach A. J., 2007. The climate change learning curve. Journal of Economic Dynamics and Control, 31**:**1728-1752.

Maliar L. and S. Maliar, 2005. Solving nonlinear stochastic growth models: iterating on value function by simulations. Economics Letters, 87**:**135-140.

Nordhaus W. D. and D. Popp, 1997. What is the value of scientific knowledge? An application to global warming using the PRICE model. The Energy Journal, 18**:**1-46.

Nordhaus W. D., 2008. A Question of Balance: Weighing the Options on Global Warming Policies, Yale University Press, New Haven & London.

Oppenheimer M., B. C. O'Neill, and M. Webster, 2008. Negative learning. Climatic Change, 89**:**155-172.

Peck S. C. and T. J. Teisberg, 1993. Global warming uncertainties and the value of information: an analysis using CETA. Resource and Energy Economics, 15**:**71-97.

Roe G. H. and M. B. Baker, 2007. Why is climate sensitivity so unpredictable? Science, 318**:**629.

Rust J., 1996. Numerical dynamic programming in economics. Handbook of computational economics, 1**:**619-729.

Stocky N. L.and R. E. Lucas, with Prescott, Edward C., 1989, Recursive Methods in Economic Dynamics. Harvard University Press.

Tol R. S. J., 2003. Is the uncertainty about climate change too large for expected cost-benefit analysis. Climatic Change, 56: 265-289.

Ulph A. and D. Ulph, 1997. Global warming, irreversibility and learning. The Economic Journal, 636-650.

Weitzman M. L., 2009. On Modeling and Interpreting the Economics of Catastrophic Climate Change. Review of Economics and Statistics, 91**:**1-19.

Yohe G. W. and R. S. J. Tol, 2010. Precaution and a dismal theorem: Implications for climate policy and climate research. In Risk management in community markets, ed. Geman, Helyette, 91-9. New York: Wiley.

Table 1 The optimal carbon tax and the net present value of utility

	Deterministic	Uncertainty			Learning		
		Correct	H	L	Correct	H	L
Optimal carbon tax in 2005 (2005US$/tonC)	23.9	89.0	112.5	83.0	80.8	80.9	78.7
Net present value of utility (arbitrary unit)	0	18.8	13.4	23.9	19.8	17.8	18.4

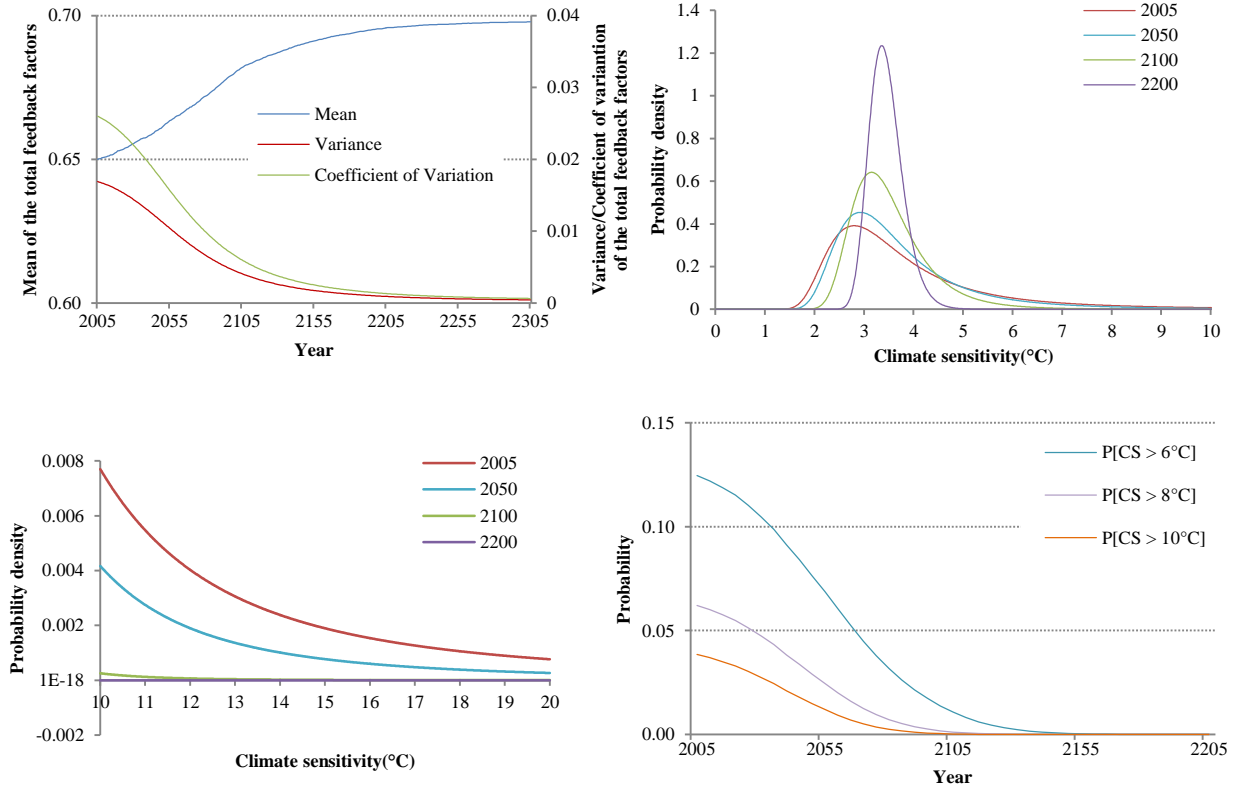Note: the net present values are normalized for the deterministic case to be 0.

Figure 1 The resolution of uncertainty. Top left: parameter values for the total feedback factors distribution. Top right: climate sensitivity distribution (left segment). Bottom left: climate sensitivity distribution (right segment). In 2200, the density approaches 0 far faster than the other cases, and thus it does not show up in the panel. Bottom right: tail probability of climate sensitivity distribution. Throughout the figure, the true value of the total feedback factors is assumed to be 0.65 (except the top left panel: 0.70) and the prior belief on the total feedback factors is normal distribution with mean 0.65 and standard deviation 0.13.
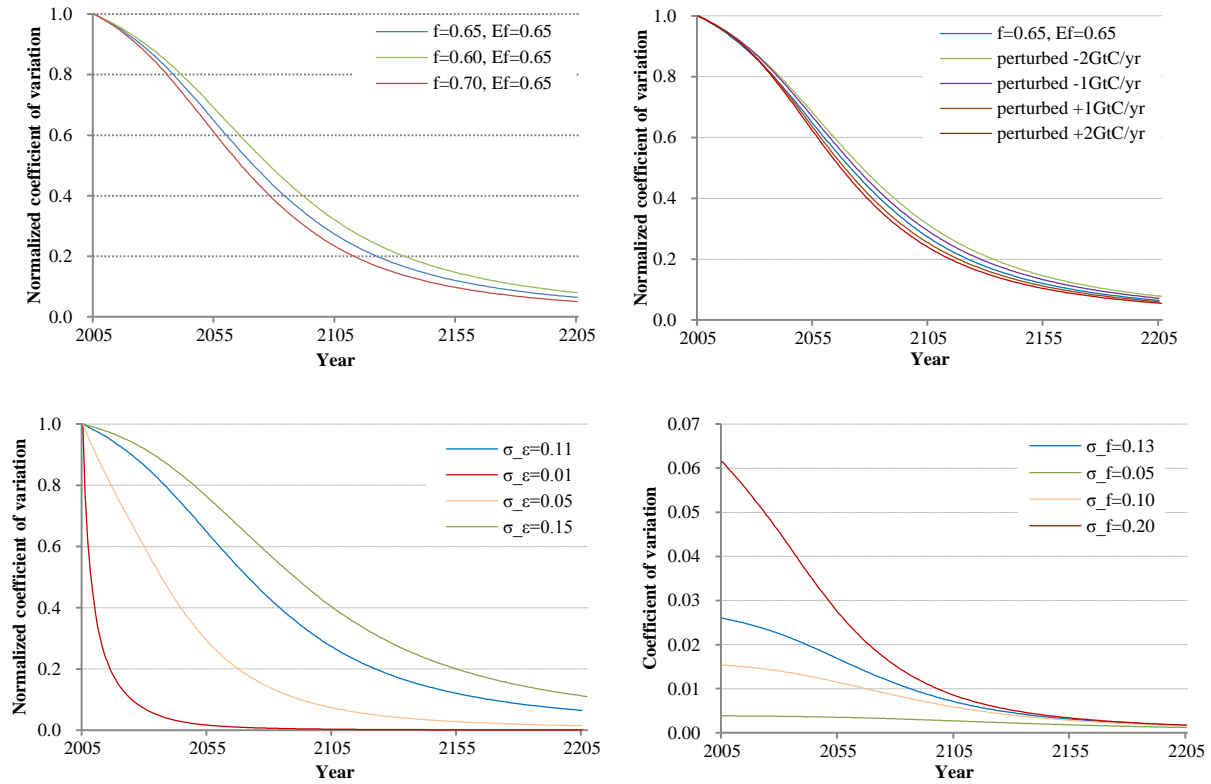
Figure 2 The rate of learning. Top left: sensitivity to the true value of the total feedback factors. Top right: sensitivity to changes in emissions. Bottom left: sensitivity to temperature shocks. Bottom right: sensitivity to the initial level of uncertainty. The true value of the total feedback factors is assumed to be 0.65 (except the top left panel: different true values) and the prior belief on the total feedback factors is normal distribution with mean 0.65 and standard deviation 0.13 (except the bottom right panel: different initial variances). Note that the coefficient of variation is normalized by the initial value (0.026 in 2005) throughout the figure, except the bottom right panel.
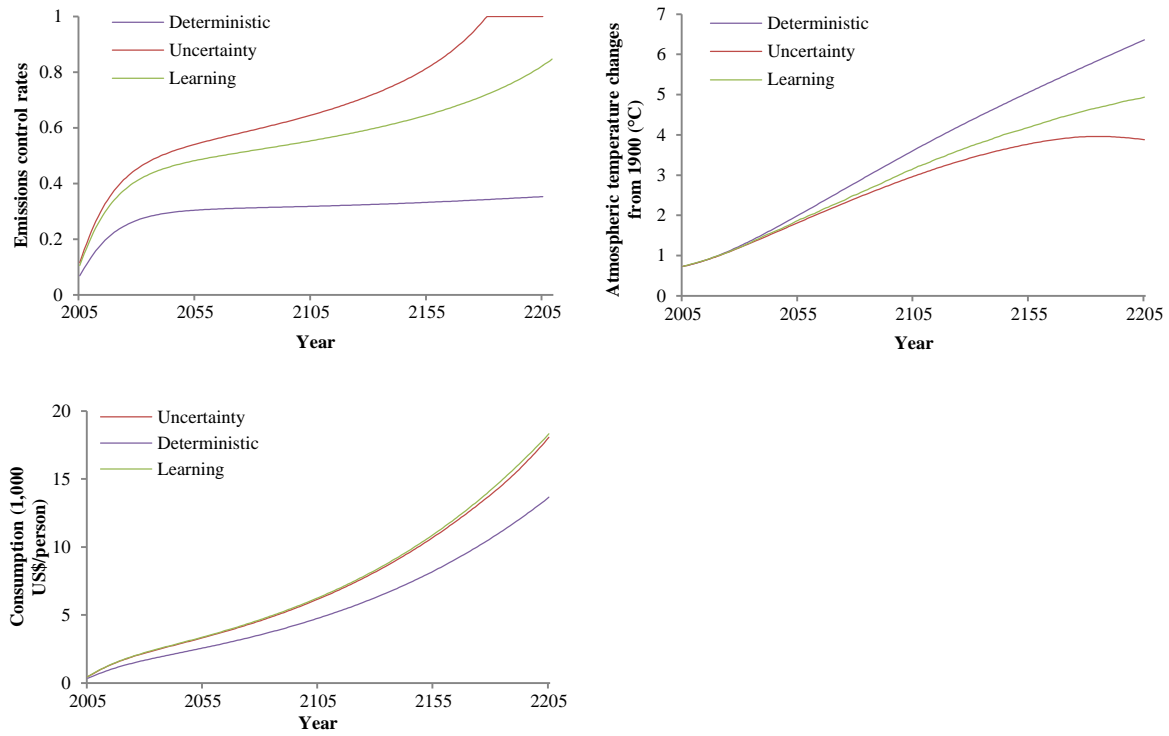
Figure 3 The effect of learning (Correct belief case). Top left: emissions control rates. Top right: atmospheric temperature increases. Bottom: consumption. The true value of the total feedback factors is assumed to be 0.65 and the initial belief on the total feedback factors is normal distribution with mean 0.65 and standard deviation 0.13.
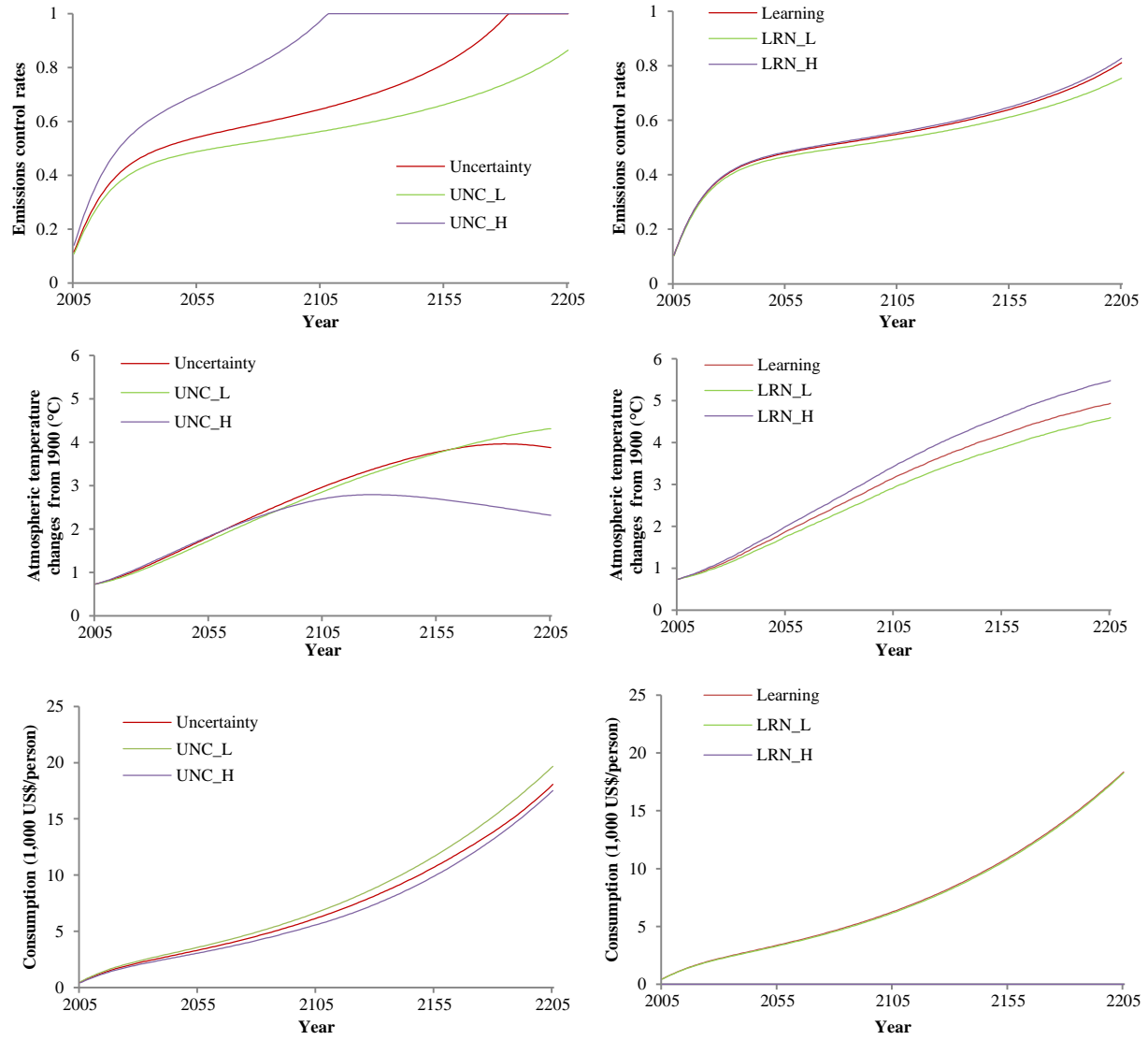
Figure 4 The effect of learning (Incorrect belief case). Top: emissions control rates. Middle: atmospheric temperature evolution. Bottom: consumption. UNC and LRN refer to the uncertainty case and the learning case, respectively. H (L) refers to the case where the true value of the total feedback factors turns to be higher (lower) than expected (H: f=0.70 > $E_0 f$=0.65, L: f =0.60 < $E_0 f$ =0.65). The initial belief on the total feedback factors is normal distribution with mean 0.65 and standard deviation 0.13.
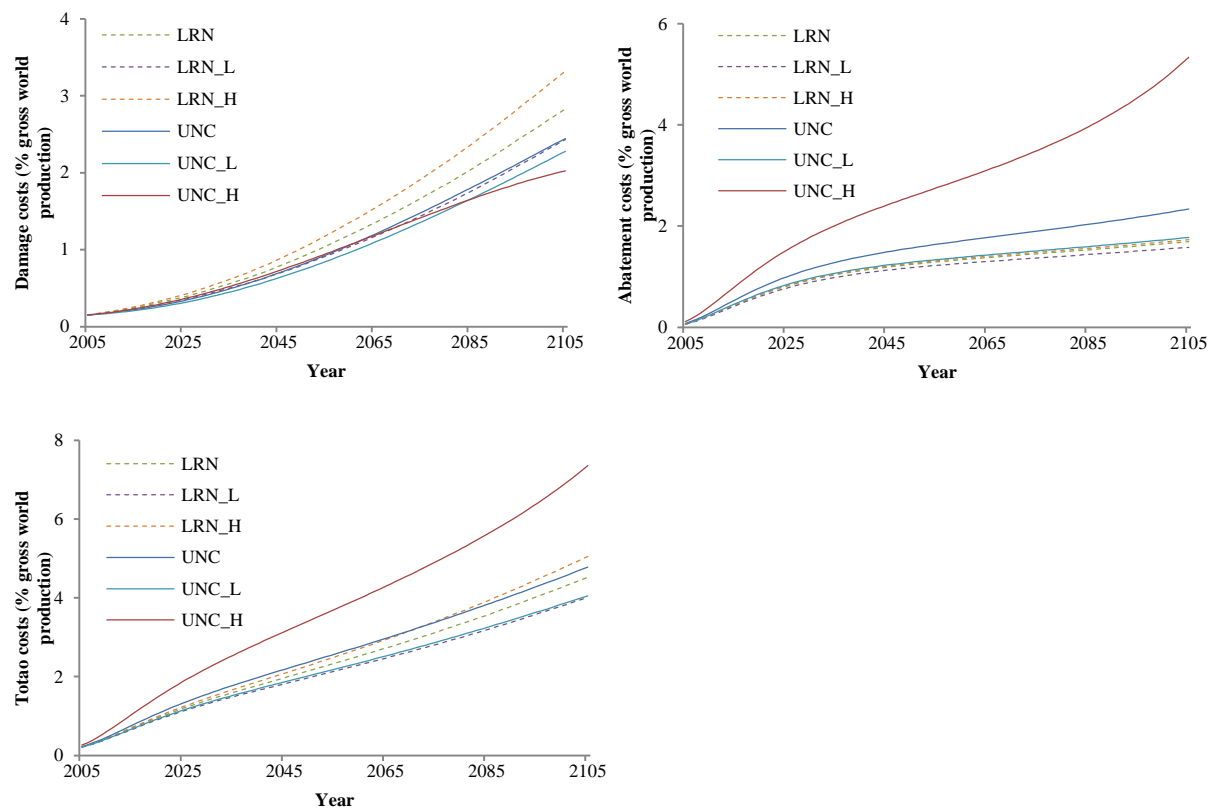
Figure 5 Social cost of climate policy. Top left: damage costs. Top right: abatement costs. Bottom: total costs.

Appendix A: computational methods for the learning model.

This Appendix illustrates the detailed solution methods for the learning model of the current paper. This provides additional information to Section 2.3. We approximate the value function in Equation (11) as the following Equation (A1). We use the logarithmic function as the basis function.[22] Since f represents a parametric uncertainty and ε is a white noise by assumptions, Equation (12) reduces into Equation (A1).

$$W(\boldsymbol{s}_t, \boldsymbol{\theta}_t) = b_0 + b_1 \ln(k_t(f,\varepsilon)) + b_2 \ln(m_t(f,\varepsilon)) + b_3 \ln\big(T_{AT_t}(f,\varepsilon)\big)$$
$$+ b_4 \ln\big(T_{LO_t}(f,\varepsilon)\big) + b_5 \ln\big(\bar{f}_t(f,\varepsilon)\big) + b_6 \ln(v_t(f,\varepsilon)) \tag{A1}$$

where the notations are the same as in the core text except k_t is the per capita capital and m_t is the normalized carbon stock.

In order to avoid an ill-conditioned problem during regression, we normalize the economic variables by the level of population and the carbon stocks by the preindustrial level (M_b). In addition, we apply the least-square method using a singular value decomposition (SVD) (Judd et al., 2011).

The first order necessary conditions for the Bellman equation (Equation 11) are as follows.

$$\frac{\partial u(\boldsymbol{s}_t, \boldsymbol{c}_t, \boldsymbol{\theta}_t)}{\partial \boldsymbol{c}_t} + \beta E_t \frac{\partial g(\boldsymbol{s}_t, \boldsymbol{c}_t, \boldsymbol{\theta}_t)}{\partial \boldsymbol{c}_t} \cdot \frac{\partial W(\boldsymbol{s}_{t+1}, \boldsymbol{c}_{t+1})}{\partial \boldsymbol{s}_{t+1}} = 0 \tag{A2}$$

where g is the law of motions for state variables (Equation (3) – (8)). The resulting policy rules for the emissions control rates and the gross investments are functions of the current state variables and the coefficients $\boldsymbol{b}$ of the basis function as follows (for analytical tractability, we assume that the elasticity of marginal utility of consumption is 1 and θ_2 of the abatement costs function is 2).

$$I_t = \frac{\beta b_1}{\beta b_1 + 1} \frac{A_t k_t^{\gamma}(1 - \theta_1 \mu_t^{\theta_2})}{(1 + \kappa_1 T_{AT_t} + \kappa_2 T_{AT_t}^{\kappa_3})} - \frac{(1 - \delta_k)k_t}{(\beta b_1 + 1)} \tag{A3}$$

$$\mu_t^{\theta_2} + A(t)\mu_t^{\theta_2 - 1} + B(t) = 0, \, where$$
$$A(t) = -\frac{(\beta b_1 + 1)\theta_2[\sigma_t A_t k_t^{\gamma} L_t + E_{LAND_t} + (1 - \delta_M)(M_b m_t - M_b) + M_b]}{[\beta b_2 + \beta \theta_2 b_1 + \theta_2)]\sigma_t A_t k_t^{\gamma} L_t} \tag{A4}$$

---

[22] As a basis function we choose a logarithmic function because it is convenient for finding the solutions (Equation (A3) and (A4)) satisfying the first order conditions. The alternatives as a basis function are ordinary polynomials or Chebyshev polynomials (Judd et al, 2011). The logarithmic function has an advantage in that the objective function of the decision maker has the similar functional form (Note that we use a logarithmic utility function).

$$B(t) = -\frac{\beta b_2 [A_t k_t{}^\gamma + (1 - \delta_k)k_t(1 + \kappa_1 T_{AT_t} + \kappa_2 T_{AT_t}^{\kappa_3})]}{[\beta b_2 + \beta \theta_2 b_1 + \theta_2)]\theta_1 A_t k_t{}^\gamma}$$

We calculate the expectation operator with a deterministic integration method, namely, the Gauss–Hermite quadrature (GH).

$$E_t W(\boldsymbol{s}_{t+1}, \boldsymbol{\theta}_{t+1}) = \sum_{j=1}^{J} w_j \, \psi(\boldsymbol{s}_t, \boldsymbol{\theta}_{t,j}; \boldsymbol{b}) \tag{A5}$$

where j is the integration nodes, w_j is the corresponding weights, J is the total number of integration nodes (we run the model with J from 5 to 10). The integration nodes and the integration weights are calculated from the GH formula (Judd, 1998).

The updating rule and the stopping rule are as follows, respectively.

$$\boldsymbol{b}^{(p+1)} = (1 - \lambda)\boldsymbol{b}^{(p)} + \lambda \widehat{\boldsymbol{b}} \tag{A6}$$

$$\frac{1}{T} \sum_{t=0}^{T-1} \left| \frac{W(\boldsymbol{s}_t, \boldsymbol{\theta}_t)^{(p+1)} - W(\boldsymbol{s}_t, \boldsymbol{\theta}_t)^{(p)}}{W(\boldsymbol{s}_t, \boldsymbol{\theta}_t)^{(p)}} \right| \leq \omega \tag{A7}$$

where $\widehat{\boldsymbol{b}}$ is the vector of coefficients estimated from the regression Equation (12), λ is the damping parameter ($0 < \lambda < 1$) (we choose λ from 0.01), T is the simulation length (we set T at 1,000), ω is the maximum tolerance level (we set ω at 10^{-6}), (p) refers to the pth iteration.

This method finds the equilibrium paths for the economy. For illustration, we present the results for a simple deterministic economic growth model. The problem of the decision maker is to maximize the social welfare defined as in Equation (A8) subject to Equation (A9).

$$\max_{C_t} \sum_{t=0}^{\infty} \beta^t \frac{(C_t/L_t)^{1-\alpha}}{1 - \alpha} \tag{A8}$$

subject to $K_{t+1} = (1 - \delta_k)K_t + F(K_t) - C_t$ \hfill (A9)

where F is the production function. The other notations are the same as our main model. We construct the Bellman equation for this problem (Equation A10) and approximate the value function with a logarithmic function (Equation

21

A11). Applying the above mentioned method, we find the optimal time paths for consumption and capital (the maximum tolerance level ω is 10^{-9}). We use the same parameter values for solving this simple model.

$$W(K_t) = \max_{C_t}[u(C_t/L_t) + \beta E_t W(K_{t+1})]$$ (A10)

$$W(K_t) \approx \psi(K_t; \boldsymbol{b}) = b_0 + b_1 \ln(K_t)$$ (A11)

Figure A1 shows the results. The economy approaches the equilibrium and the variables stabilize. Note that since we are dealing with an infinite time horizon problem, consumption is small relative to income (for more on this point, see Gollier, 2000).
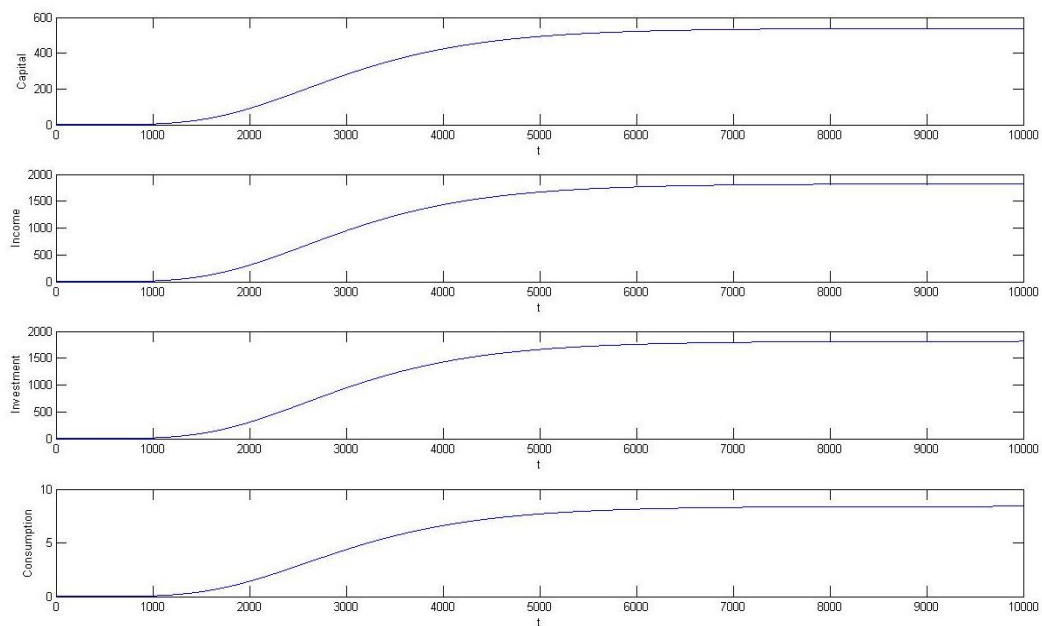
Figure A1 Solutions for the economic growth model (A8-A9). All variables have per capita values (1,000US$ per person).

Appendix B: Investment and gross income flow of UNC_L and LRN_L cases (see Section 5).
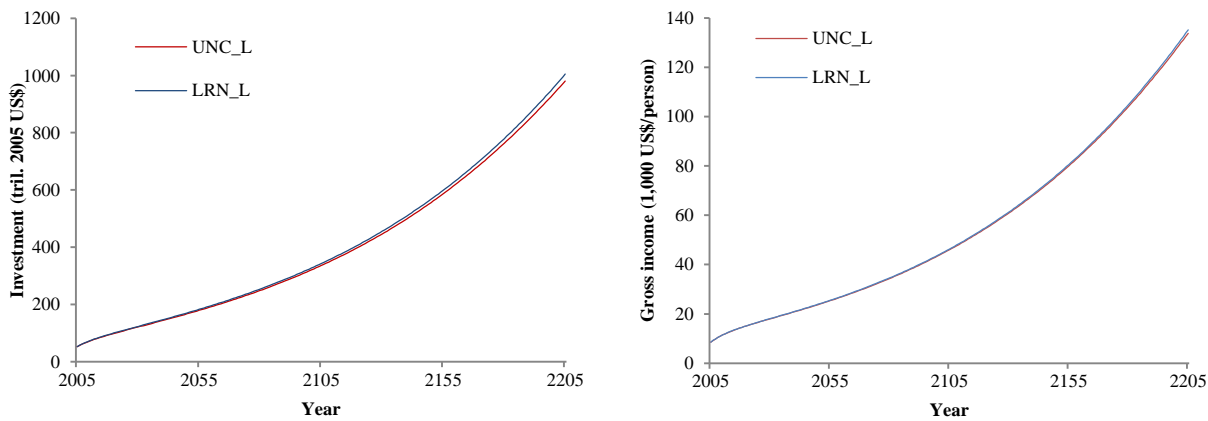
Figure B1 Investment and gross income flow of UNC_L and LRN_L cases (see Section 5)