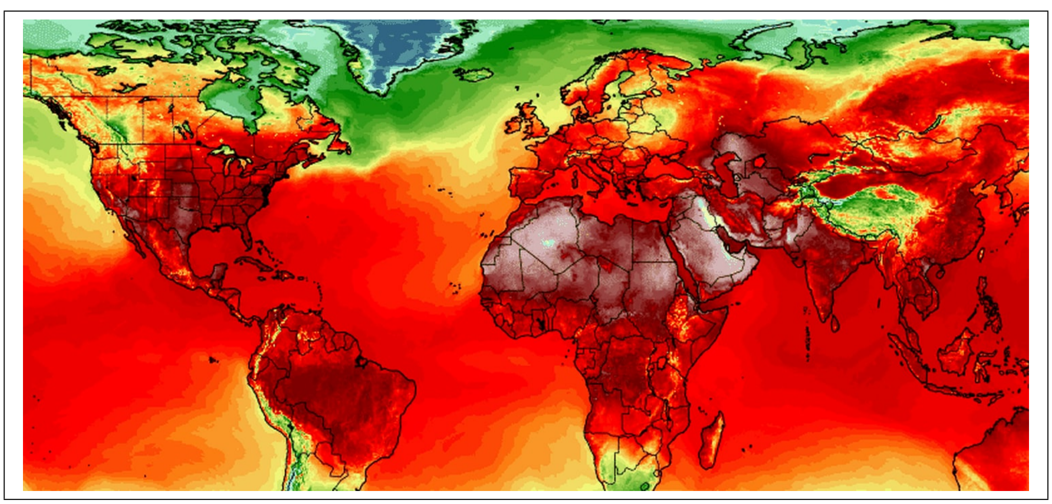
Universiteit Utrecht

Opleiding Natuur- en Sterrenkunde

Performance of Bayesian Learning Applied on the Climate Sensitivity in Economic Climate Modeling

Bachelor Thesis

Samuel Zeegers



Supervisors:

Dr. C.E. Wieners
IMAU; Institute for Marine and Atmospheric research Utrecht
CCSS; Center for Complex System Studies

Prof. Dr. Ir. H.A. Dijkstra
IMAU; Institute for Marine and Atmospheric research Utrecht
CCSS; Center for Complex System Studies

June 12, 2019

Acknowledgements

First of all I want to thank Huib de Swart with helping find a thesis subject when I did not a good idea of what I wanted to do. He introduced me to Claudia Wieners, who helped me so much all throughout my project. I was not skilled at all in economic climate modeling, I never heard about it before. She always was available for questions and always replied to my questions as soon as she could. I would like to thank her a lot for that, without her I would not have been able to deliver this thesis. My interest into economic climate modeling has grown because of her. I am really happy that I could work with her and I have learned a tremendous amount from this project and from her. Finally I would like to thank Henk Dijkstra to make it possible to do my bachelor thesis with him. And also for making time for me in his busy schedule.

Abstract

In the last years economic climate modeling has attracted a lot of attention because of its fundamental interest and applications to policy-making. The most famous economic climate model is DICE, made by William Nordhaus. In his model he assumed that the climate sensitivity is constant. Motivated by the results of the IPCC reports, which over the years show big deviations in estimates of the climate sensitivity, this thesis develops a framework to model the economy and the climate with uncertain climate sensitivity. This is done by implementing Bayesian learning, the results in this thesis show that the framework works and that Bayesian learning has a significant impact on the welfare. It is found that Bayesian learning will enhance the performance of a policy by 476%.

Contents

1 Introduction

Climate change is a hot topic in research right now. Since the first article that carbon dioxide (CO_2) is warming up the climate[1], a lot has changed. And the amount of research in the topic has tremendously increased. Besides that, there is a scientific consensus on the topic.[2] Despite all that, the number of people believing in climate change has remained constant over the years.[3] The future does not seem all so bright, especially when one thinks about the fact that we are emitting more CO_2 than the worst-case scenario stated in 1999. To make it even worse the 1.5 degree Celsius heating mark has already been passed.[4] We do not even understand the consequences of global warming and we might not even want to know what would happen when the global temperature rises 3 degrees Celsius. Therefore, it would be wise to try and lower the CO_2 concentration. But as far, for now, this is not happening. Policies of countries play an important role in this. For example, the presidential election of Donald Trump in the United States or Jair Bolsonaro in Brazil, who both completely disregard climate change as an issue and even state publicly that climate change is not happening at all. Brazil had one of the cleanest electricity portfolios in the world. This is assumed to decrease under the policy making of Bolsonaro.[5] This depicts how changing a policy can affect the climate, climate change is a worldwide problem. Therefore, choosing the correct policy can make a big difference.

Since the meat industry and the big companies are the biggest emitters[6] of CO_2, a carbon tax, for those who emit the most, seems more than reasonable. It is always beneficial for the economy to get a lot of energy for the lowest price. This rises a problem, there is always going to be a trade-off between green energy being good for the climate and grey energy being good for the economy. Therefore you cannot expect firms to lower their emissions just for the sake of climate change, you need to obligate them and make green energy economically more attractive for them. This means that without any policies on emissions, the main energy source will stay fossil fuels. This is why it is important to link the economy to climate change. More specifically, to look if it is economically efficient to 'care' about climate change. This topic was first addressed by W. Nordhaus in his famous DICE (Dynamic Integrated model of Climate and the Economy) model which won him the Nobel price in economics. He found that the best policy would lead a to global warming of about 3 degrees Celsius.[7] This article got criticism, because of the hand-wavy assumptions he made, especially regarding the climate. After this, a few improvements were made and a lot of research has been done on this model. However, improvements in the climate side of DICE and the policy-making side are still behind. A big step on the economy side has been made by F. Lamperti et all with an agent-based model which models the economy called the DSK (Dystopian Schumpeter meeting Keynes) model. The climate and policy part of this model, however, is fairly slim.[8]

One of the biggest problems that all these models have, considering the climate side, is that they model the climate's response to CO_2 as just a variable. This variable is known by everyone and therefore also by the policymaker in the model. In the real world, this is not the case, the climate's response to CO_2 is related to the increase in equilibrium temperature after a doubling of CO_2. This so-called climate sensitivity is very unclear, which can be seen from disagreement on that matter in the IPCC (Intergovernmental Panel on Climate Change) reports over the years.[9] This uncertain climate has been studied in the DICE model over the years, where some results were more disastrous than others.[10] However, it has never been done using an agent-based model of the economy. In this thesis, the first step of introducing uncertain climate sensitivity in an agent based model is done. Motivated by the DSK model and the DICE model a new non-agent based model is created called the Hot Small World Model, most of the model is based on DICE but the ultimate goal is to improve it and make an agent based model from it. The purpose is to find the 'best' policy under uncertain climate sensitivity in this model. The 'best' policy is the policy that maximizes welfare. This is important to study because the climate is uncertain and we do not know how the climate would change due to emissions. Motivated by the papers of Traeger and Lemoine et all [11, 12] the Hot Small World Model uses Bayesian' learning to accomplish a learning effect on climate sensitivity.

2 Theory

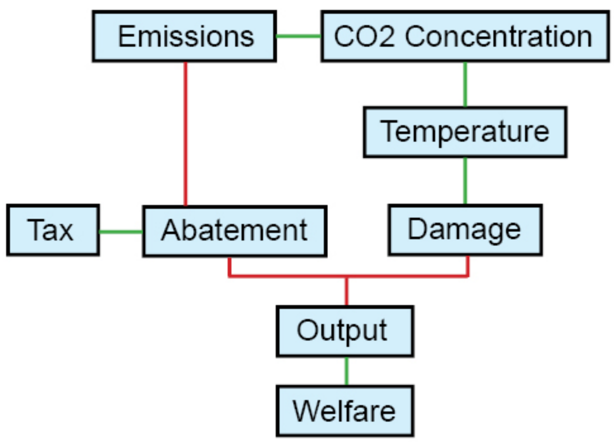


Figure 1: The simplified schematic chart of the Hot Small World Model. Green lines mean positive feedback, while red lines mean negative feedback

The Hot Small World Model tries economic climate policy under uncertain climate sensitivity. The model is build up out of three boxes, the climate box, the economy box, and the policy box. All those boxes have an influence on one another and together they form the total model. The policy box will make the policy and therefore controls the climate and the economy box. In this section, the working of these boxes will be explained, as well as how they work together. Finally, the implementation of uncertainty in the model is explained. [1] In figure 1. a simplified chart of the Hot Small World Model is displayed. Red lines indicate negative feedback while Green lines indicate positive feedback.

[1]All variables and constants with description can be found in the appendix

2.1 The Climate Box

The climate is modeled in the climate box, this is done using C-ROADS.[13] C-ROADS models the climate by distributing the carbon and temperature over different resevoirs. You can divide the model into two parts, there is a Carbon cycle in the model and there is a part that accounts for the radiative forcing and warming.

2.1.1 The Carbon Cycle

The carbon Cycle again can be divided into two parts. Namely the carbon in the atmosphere, biosphere, and hummus and the carbon in the ocean.

Atmospheric Carbon First, there will be a net primary production, this is the production of atmospheric carbon from the plants. It is logical to assume that the production of carbon decreases linearly with the global temperature because most biological processes increase or decrease linearly with temperature. The plants namely get stressed with heat and dry out, therefore they will not produce as much carbon. It also makes sense that there will be more bioproduction when the atmospheric CO_2 increases because there will be more photosynthesis. This is believed to be a logarithmic relation. Hence the primary production is given by:

$$P_N = P_{N_0}\left[1 + p_{fert}\log\left(\frac{C_{at}}{C_{at_0}}\right)\right]\left[1 + p_{heat}\Delta T\right] \tag{2.1.1}$$

Where P_N, P_{N_0}, C_{at}, C_{at_0} and ΔT are the net production, the initial net production, the CO_2 concentration in the atmosphere, the initial CO_2 concentration in the atmosphere and the change in temperature from initial conditions respectively. Furthermore the constants p_{fert} and p_{heat} denote the CO_2 fertilization constant and the heat-stress respectively. So the plant material is formed by this net production but it also dies out with a residence time of τ_{bio}, therefore the carbon concentration in the biosphere is given by:

$$\frac{dC_{bio}}{dt} = P_N - \frac{C_{bio}}{\tau_{bio}} \tag{2.1.2}$$

Note that P_N is still indirectly a function of time because it is a function of ΔT and C_{at}. Therefore the carbon concentration in the biosphere only follows an equilibration towards $\frac{P_N}{\tau}$ for constant net production. A fraction p_{hum} of the died plants will be incorporated in the humus layer, also the humus layer dies out with a residence time of τ_{hum}. The humus layer is the layer of organic material from the plants and animals that form when they die. So the carbon concentration in the humus is given by:

$$\frac{dC_{hum}}{dt} = p_{hum}\frac{C_{bio}}{\tau_{bio}} - \frac{C_{hum}}{\tau_{hum}} \tag{2.1.3}$$

The net production removes carbon from the atmosphere and puts it in the biosphere, but the carbon from decaying humus and a fraction of $1 - p_{hum}$ of the decaying biomass carbon will return in the atmosphere. In addition emissions, E enter the atmosphere mostly due to burning fossil fuels. Then there are also ocean processes that have an influence on the atmospheric carbon. This results in the following equation for the atmospheric carbon:

$$\frac{dC_{at}}{dt} = -P_N + (1 - p_{hum})\frac{C_{bio}}{\tau_{bio}} + \frac{C_{hum}}{\tau_{hum}} + E + \text{OCEAN} \tag{2.1.4}$$

Oceanic Carbon In this model the ocean consists of $n_{lay} = 5$ layers, there is a well-mixed upper layer and $n_{lay} - 1$ lower layers; the depth of layer k is represented as $d_{lay}(k)$. The layers exchange carbon due to diffusion, more specifically eddy diffusion. The upper-layer also exchanges, due to winds and waves, carbon with the atmosphere to reach equilibrium. The mixing-related carbon flux through two layers is related to the vertical derivative of the carbon concentration. The ocean is assumed to be homogeneous in the horizontal direction so the carbon concentration in layer k is just given by $\frac{C_{on}(k)}{d_{lay}(k)}$, $C_{on}(k)$ is the oceanic carbon in layer k. In this way the carbon flux between two ocean layers k and $k+1$ is:

$$M_{k,k+1} = \chi_{eddy} \frac{C_{on}(k)/d_{lay}(k) - C_{on}(k+1)/d_{lay}(k+1)}{\frac{1}{2}(d_{lay}(k) - d_{lay}(k+1))} \tag{2.1.5}$$

As can be noted equation 2.1.5 is positive if the carbon flux is in the downward direction. So now in the lower layers $k \neq 1$ the carbon concentration can be written as:

$$\frac{dC_{on}(k)}{dt} = M_{k-1,k} - M_{k,k+1} \tag{2.1.6}$$

Only in the lowest and the upper layer, these equations are not valid, in the lowest layer there is no layer below it so equation 2.1.6 results in:

$$\frac{dC_{on}(n_{lay})}{dt} = M_{n_{lay}-1,n_{lay}} \tag{2.1.7}$$

The upper layer is exposed to the atmosphere and will exchange carbon with the atmosphere. This means that the two will be in equilibrium i.e. they satisfy the equilibrium condition:

$$C_{on}(1) = C_{mix,ref}(1 - R_T \Delta T)\left(\frac{C_{at}}{C_{at_0}}\right)^{\gamma} \tag{2.1.8}$$

Where $C_{mix,ref}$ is the reference carbon concentration in the mixing layer. R_T is the temperature dependence of the carbon concentration in the mixed layer and γ is given by:

$$\gamma = \frac{1}{R_0 + R_C \log\left(\frac{C_{at}}{C_{at_0}}\right)} \tag{2.1.9}$$

Where R_0 is the Revelle factor and R_C is the carbon dependent Revelle factor. Satisfying this equilibrium is achieved in the model by first calculating the total carbon in the two resevoirs $C_{tot} = C_{at} + C_{on}(1)$, then C_{tot} is given by:

$$\frac{dC_{tot}}{dt} = -M_{1,2} - P_N + (1 - p_{hum})\frac{C_{bio}}{\tau_{bio}} + \frac{C_{hum}}{\tau_{hum}} + E \tag{2.1.10}$$

This total carbon now gets split over the ocean and the atmosphere so that it satisfies equation 2.1.8. In this way, the carbon in the first layer and in the atmosphere are modeled.

2.1.2 Heating and Radiative Forcing

Climate sensitivity is a key parameter of the C-ROADS model. Climate sensitivity relates the equilibrium temperature increase to the CO_2 increase. The equilibrium climate sensitivity is defined as the increase in equilibrium temperature when the CO_2 concentration is doubled leaving everything else the same. As you can guess this is a very hard parameter to measure precisely. The outgoing radiation per Kelvin is the energy intensity that the earth emits per K. The climate sensitivity is related to the outgoing radiation as follows:

$$\lambda = \frac{f_{\mathrm{CO_2}} \log(2)}{\kappa_{out}} \tag{2.1.11}$$

Where λ and κ_{out} are the climate sensitivity and the outgoing radiation respectively. Where $f_{\mathrm{CO_2}}$ is the scale of the radiative forcing, which is related to the radiative forcing ($F_{\mathrm{CO_2}}$) as follows.

$$F_{\mathrm{CO_2}} = f_{\mathrm{CO_2}} \log\left(\frac{C_{at}}{C_{at_0}}\right) \tag{2.1.12}$$

The value for $f_{\mathrm{CO_2}}$ is well established and is believed to be $f_{\mathrm{CO_2}} = 5.35\ \mathrm{Wm}^{-2}$. The additional forcing due to outgoing radiation due to global warming is approximated as:

$$F_{out} = \kappa_{out} T \tag{2.1.13}$$

When earth's temperature has reached equilibrium the outgoing radiation is equal to the radiative forcing, so the equilibrium temperature can be found if the radiative forcing is known.

Air has a low heat capacity due to its low density and the solid land has a low heat capacity because heat can only be transported by conduction. The ocean however has a high heat capacity, due to this it takes some time for earth to reach equilibrium temperature. As shown in figure 2 the surface temperature of the earth, that starts at 0 K, overshoots after an emission schock of 800 gigatonne and then slowly reaches equilibrium, . This means that the equilibrium temperature can not just be measured and therefore there is a complication in measuring the climate sensitivity. This is why researchers now are not certain about the climate sensitivity.

The heat transfer in the ocean is modeled similar to the carbon transfer in the ocean. So again now for the heat flux through from layer k to layer $k+1$:

$$\mathcal{H}_{k,k+1} = \chi_{eddy} \frac{H_{on}(k)/d_{lay}(k) - H_{on}(k+1)/d_{lay}(k+1)}{\frac{1}{2}(d_{lay}(k) - d_{lay}(k+1)} \tag{2.1.14}$$

With $H_{on}(k)$ the heat in layer (k) and χ_{eddy} the eddy diffusivity. Equation 2.1.14 will lead to equation 2.1.5 but then for the heat and the heat flux instead of the carbon concentration and the carbon flux. Again the upper layer of the ocean is exposed to the atmosphere so besides mixing there is also absorption of heat in this layer, hence:

$$\frac{dH_{on}(1)}{dt} = -\mathcal{H}_{1,2} + \frac{(F - F_{out})\alpha_{yr}}{A_{sea}} \tag{2.1.15}$$

Where A_{sea} is the area of the sea and α_{yr} is the conversion factor that converts $\frac{dH_{on}}{dt}$ to be measured in years instead of seconds like F. Furthermore the A_{sea} is needed because this

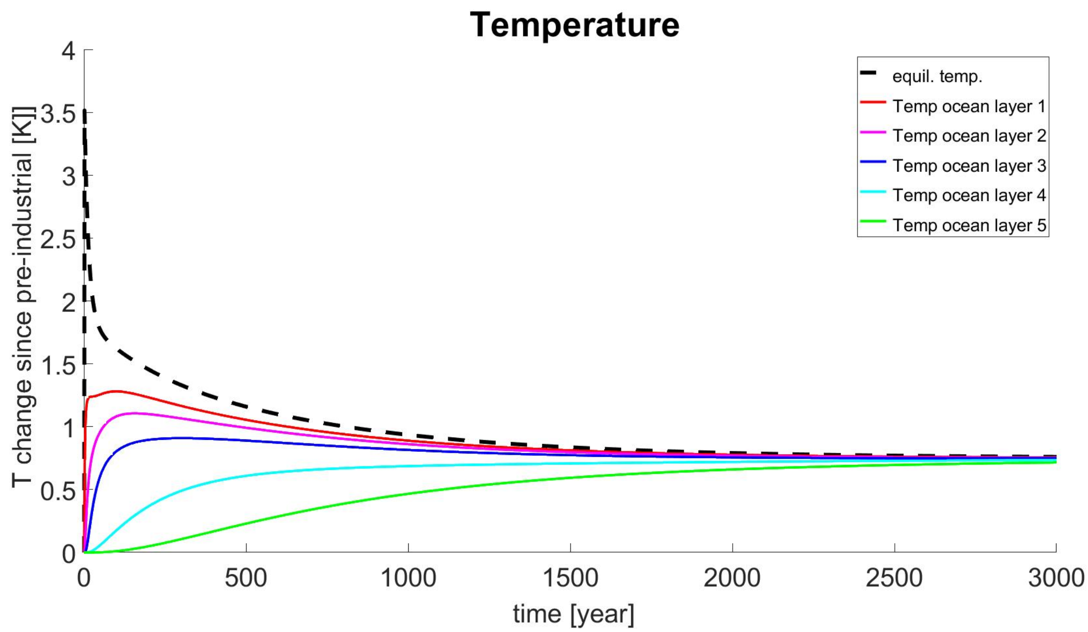


Figure 2: Temperature of the ocean layers and the equilibrium temperature after an emission shock of 800 Gt at $t = 0$. The initial equilibrium temperature was 0 K. This plot is obtained using the C-ROADS model.

equation describes the heat in the ocean and F is measured over the entire earth surface. Now the temperature can be in layer k is described by:

$$T_{on}(k) = \frac{H_{on}(k)}{cd_{lay}(k)} \tag{2.1.16}$$

where c is the heat capacity per volume.

2.2 The Economy Box

In Hot Small World the economy is modeled by using some principles of DICE2013 and making it fit the RCP8.5 scenario.[7, 15] The driving force in the economy are the firms, the model can have n firms. These firms are free to compete or to work together. This competition can all be modeled in a competition matrix, which describes the relations between different firms. This means that in this matrix one could model the growth of a firm due to the growth of a firm that has relations with it. Every firm generates an output, has a certain capital and has emissions. The economy box just like the climate box can be split up in different parts, the economic production, interaction with the climate, climate policy and the cost of abatement.

2.2.1 Economic Production

Every firm will generate an output, the gross output of a firm is a function of its capital its labor force and the productivity:

$$Q_{gross} = A_0 \exp\Big(\frac{2.66\arctan(t/100)}{\pi}\Big) K^{\gamma} L^{1-\gamma} \tag{2.2.1}$$

where time t is in years K is the capital, L is the labor force, A_0 is a constant that determines the initial output and γ determines how much each factor contributes. This is based on the DICE function for the output, but it is made to fit the RCP8.5 scenario when no abatement is done. Namely when no abatement is done the output grows with a factor 6 over the first 100 years, which is consistent with the RCP8.5 scenario.[15] Equation 2.2.1 shows a decreasingly rising productivity with the $A_0 e^{\arctan(t/100)}$ factor. The net output of a firm is dependent on the climate damage that happened see section 2.2.2, the abatement a firm did, the cost of this abatement see section 2.2.4 and the carbon tax see section 2.2.3. Then the net output is given by:

$$Q(t,\mu,T,E) = Q_{gross}(t) - c_\mu(t,\mu) - D(T(t), Q_{gross}(t)) - (1-\mu)E_{gross}(t)\tau(t) \qquad (2.2.2)$$

Where μ is the abatement which is the factor of the gross emissions that the firm has not emitted. D is the damage, E_{gross} is the gross emission of the firm and c_μ the costs for abatement.

The capital in equation 2.2.1 naturally depreciates because machines are getting worse and need to be replaced. On the other hand capital also increases because a fraction s of the net output is saved up. To import this in the model the capital is given by:

$$K(t) = K_0(1-\Delta K)^t + Q(t)s \qquad (2.2.3)$$

where ΔK is the capital depreciation per year, t is the time in years, s is the saving rate per year and Q is the net output. Knowing the net output and the saving rate the consumption per person can be given by:

$$C_{pp}(t) = C_{pp}(t) = \frac{Q(t)(1-s)}{N(t)} \qquad (2.2.4)$$

Where $N(t)$ is the number of people which is also equal to the labor force for the Hot Small World Model. Now the "happiness" (utility) can be defined just as in DICE2013 [14]:

$$U(t) = \frac{C_{pp}^{1-\chi_{elas}}(t)}{1-\chi_{elas}} N(t) \qquad (2.2.5)$$

The utility obeys 2.2.5, i.e. it increases monotonously with consumption, but less than linearly because a person already consuming much will gain less additional happiness from 1 extra unit of consumption than a person consuming little. where χ_{elas} is the elasticity of the market. This utility can then be used to calculate the welfare. The welfare is important because the welfare needs to be optimized in order to get an optimal policy. The welfare is defined just like in DICE2013[14]:

$$W(t) = \int U(t)\Big(\frac{1}{1+\rho(t)}\Big)^t dt \qquad (2.2.6)$$

Here $\rho(t)$ is the rate of impatience and is given by $\rho(t) = \rho_0 e^{-g_\rho t}$. The rate of impatience states that you would rather take 100 euros than 100 euros in a few years.

2.2.2 Interaction With The Climate

The interaction between the economy box and the climate box is modeled by firms emissions and the interaction between the climate box and economy box is modeled by damages. The emissions will higher the carbon concentration and therefore the temperature, the damages are assumed to be temperature dependent and will cause a decrease in output see equation 2.2.2. Every firm has some emissions (E) this emission depends on three factors. First of all the firms abatement if the firm does more abatement their emissions will decrease. Then there is the net output it is believed that when the output increases the emissions also increase. Lastly it is dependent on time, when the time grows the firms will get more technological advanced and will emit less for the same abatement. The latter is modeled in a variable called the technological advancement factor. Taking all this into account the emissions of a firm are given by:

$$E(t) = (1 - \mu(t))E_0 e^{-\alpha t} Q(t) \tag{2.2.7}$$

Where E_0 is the initial emission, α is the technology advancement factor and μ is the abatement at time t. $E_0 e^{-\alpha t} Q(t)$ are then defined as the gross emissions, the emissions when $\mu = 0$. The technology advancement factor(α) states that firms will get more technological advanced overtime and due to this the emissions for a given output decreases. These factors can be made different per firm. These emissions will heat up the planet via the C-ROADS model which causes damage. The damage due to heating is modeled like DICE2013, however this can be easily changed to for example the damage function of Weitzman [16]. The damage function that is used in DICE2013 is:

$$D(T, Q) = (1 - \Big(\frac{1}{1 + \psi_0 T(t)^2}\Big) Q_{gross}(t) \tag{2.2.8}$$

The Weitzman Damage function is given by:

$$D(T, Q) = (1 - \Big(\frac{1}{1 + \psi_1 T(t)^2 + \psi_2 T(t)^{6.76}}\Big) Q_{gross}(t) \tag{2.2.9}$$

2.2.3 Climate Policy Box

Due to the damages in equation 2.2.8 and 2.2.9 the output of the firms decreases, and therefore decreases the welfare. So there has to be some sort of policy that does not let the temperature increase too much. This is the chess piece that the policymaker has to control the climate while optimizing the welfare. The firms only increase the damage by emissions, so when the firms would not emit at all they would not cause any harm. So the policy maker needs to control the emissions of the firms, as can be seen in equation 2.2.7 there are a few possibilities, the policymaker could let the firms travel in time, as for now there is no possible way to time travel so we sadly have to discard this option. Next the policymaker could regulate the output of the firms, but decreasing the output would also directly decrease the welfare. The last option is to regulate abatement this, in fact, would work. So the policymaker has to make abatement attractive for the firms, there are many ways to do that, like subsidies for green firms or a tax on the firms emissions. Here the latter is assumed, the policymaker can tax the emissions of a firm, the tax τ is collected at the end of the year and put into a fund f. At the beginning of next year the firms are asked to do abatement and can receive subsidy for it this subsidy can maximally be equal to the money in the fund or to half of the abatement costs (see section 2.2.4). The leftover parts are used to compensate damage again only half of the damage can be compensated at most because some damages are non reparable like the extinction of a certain plant. If there is any money left over after this it will be added to the fund next year. This is essentially what the policy maker does in the Hot Small World Model.

2.2.4 Cost of Abatement

Sadly enough abatement is not free. In DICE the cost of abatement depends on time the with the same argument as the technological advancement. The Hot Small World Model takes another approach it uses a learning factor called l_μ. This learning factor models a learning effect on the abatement the more abatement a firm does the more advanced it becomes with the technologies of abatement and the cheaper the abatement becomes. The learning effect is given by:

$$l_\mu = g_0 e^{\mu \log(g_l)} \tag{2.2.10}$$

Here g_0 is the abatement cost when there is no abatement and therefore when there was not any learning and g_l is the learning factor.

It makes sense that the abatement costs will be higher when the abatement is higher, because an abatement of 0.1 is easily realizable by for example just turning the lights of when there is enough light in the room. However, to reach 0.9 abatement a lot more expensive measures should be installed like getting your energy from a more expensive but greener power plant. Also the abatement costs will increase linearly with the emissions more emissions means that you have to abate more CO_2, in general it is more expensive for a car firm to achieve an abatement of 0.1 than for a bakery, because the car industry will have more emissions. So finally the abatement costs are given by:

$$c_\mu = l_{\mu_{old}} E_{gross} \mu^{2.15} \tag{2.2.11}$$

As discussed in section 2.2.3 the policymaker gives some subsidy(σ_μ) for abatement this subsidy is maximal half of the abatement cost or the entirety of the fund:

$$\sigma_\mu(c_\mu, \tau_\mu) = \min(0.5c_\mu, f(t)) \tag{2.2.12}$$

Then there will also be costs for the firm due to the tax of the policymaker, therefore the total cost of the firm is given by:

$$c = c_\mu(l_{\mu_{old}}, \mu, E_{gross}) - \sigma_\mu(0.5c_\mu, f) + E_{gross}(1 - \mu(t))\tau(t) \tag{2.2.13}$$

Firms are assumed to be selfish i.e. the firms only do abatement if it is financially beneficial for them. This means that the firms will do a certain abatement that minimizes their costs:

$$\mu = \min_\mu(c_{(\mu, E_{gross}, c_\mu, \sigma_\mu)}) \tag{2.2.14}$$

This is how μ is found in the Hot Small World Model.

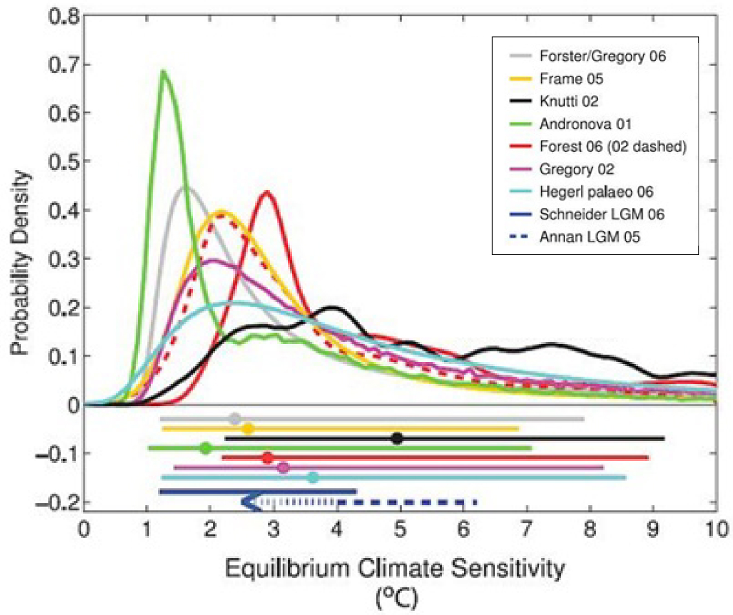


Figure 3: esitmates of the climate sensitivity over the years. The different colored curves represent different estimates of the climate sensitivity. As can be seen there is a lot of difference over the years in the estimates of the climate sensitivity.

2.3 Handling Uncertainty

The policy maker can regulate the climate by taxing the emissions, a higher tax and a higher fraction of this tax that goes to subsidy for abatement will encourage or even obligate firms to do abatement. There are other options to make abatement more attractive for firms, but only tax is discussed here. In principle the policymaker has two decisions that he can make. He can tax the emissions and he can decide what to do with this tax. For the latter it is assumed that the policymaker can put a certain factor of the money into subsidy the leftovers will be used to pay for the damages. The policymaker has a relatively easy job when he knows exactly how the global temperature will respond to emissions, or in other words when it knows the climate sensitivity exactly. Then the policy maker could just maximize the wealth with respect to the τ. A constant climate sensitivity however is not realistic, the climate sensitivity is not known exactly. The true value of the climate sensitivity is still highly uncertain, and estimates have hardly narrowed down over the last few decades. The latest estimates for the climate sensitivity are shown in figure 3. When the climate sensitivity is high the policy maker would tax the emissions more than when the climate sensitivity is low, because a high climate sensitivity would result in more damage per emitted Gt (gigatonne) CO_2. As of now there are only probabilities that some climate sensitivity will occur as can be seen in figure 3. there is a small but real chance that the climate sensitivity is around 10, this would probably mean that the policymaker has to take this into consideration and will probably want to tax more than just the optimal path for the mean climate sensitivity of the pdf(probability density function). In principle, we can measure climate sensitivity by emitting sufficient CO2 and determining the temperature change. Currently, the global warming signal is still to small (compared to natural variability) as to yield a sufficiently precise value. However, in the future, the policy maker can "learn" the true value by taking temperature measurements. This process can be formalized by the method of Bayesian learning. With Bayesian learning the policymaker could guess the climate sensitivity by the mean of a certain pdf that is constructed by temperature measurements over time. As time evolves the guess of the policymaker will be more accurate and reach the actual climate sensitivity. It is assumed that the CO_2 concentration is well known and that the only uncertainty is due to natural variability of the surface temperature. Since temperature variability on timescales of years to decades are largely due to fluctuations in the heat exchange between the surface and deeper ocean, we mimic this process by adding a noise to the temperature flux between the mixed and the second highest ocean layer, this is done by adding a noise in equations 2.1.14 and 2.1.16. The policymaker however does not know this and assumes that the noise is just a Gaussian noise in the surface temperature:

$$T = T_\lambda + \delta T \tag{2.3.1}$$

Where T_λ is the temperature increase due to emissions and δT is some added noise on the temperature, the policymaker measures T, but wants to know λ, if the policymaker would be able to measure T_λ it is simple to find λ via equation 2.1.11, 2.1.12 and 2.1.13.
The policymaker has an initial guess of the climate sensitivity in the form of a pdf, he then updates this guess given the measurements. Inspired by the paper by Baker et all [12] , κ_{out} can be taken as Gaussian, this yields a skewed pdf for λ. A first guess will resemble the current pdf of λ, a Gaussian for the outgoing radiation is in general given by:

$$p(\kappa_{out}) = \frac{1}{2\pi w_\kappa} e^{-\frac{(\kappa_{out}-m\kappa)^2}{2w_\kappa^2}} \tag{2.3.2}$$

Taking $m_\kappa = 1.23$ W/(m^{2}K) and $w_\kappa = 0.3$ W/(m^{2}K) yields a good pdf of the current estimates for λ. You can switch between pdf's of λ and κ_{out} by using $p(x(y)) = p(y)\frac{dy}{dx}$, or in this case

$$p(\lambda(\kappa_{out})) = p(\kappa_{out})\frac{d\kappa_{out}}{d\lambda} \tag{2.3.3}$$

Naturally the policymaker knows that there is noise, he can also predict the temperature $T_P(\kappa_{out})$ based on his guess of κ_{out} and the current CO_2 concentration. The policymaker puts a normal distribution around his predicted temperature, this normal distribution is given by:

$$p(T|\kappa_{out}) = \frac{1}{2\pi\sigma T} e^{-\frac{(T-T_P(\kappa_{out}))^2}{2\sigma_T^2}} \tag{2.3.4}$$

The noise that he assumes σT does not have to be equal to the initial noise guess δT. Basically $p(T|\kappa_{out})$ is the policy maker's estimate of how likely it is to measure temperature T if a certain value for κ_{out} holds. Now Bayesian learning can be used to actually learn something about $p(\kappa_{out}|T)$,this is the how likely it is to have a outgoing radiation κ_{out} when a certain T is measured. Bayes' rule is given by:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

in this case, E the "evidence" is the temperature that is measured and H the "hypothesis" is a value of κ_{out}. This means that Bayes' rule in this case can be written as:

$$P(\kappa_{out}|T) = \frac{P(T|\kappa_{out})P(\kappa_{out})}{P(T)} \tag{2.3.5}$$

Equation 2.3.2 and 2.3.4 give the two pdf's in the numerator of this rule. Now only $P(T)$ is unknown, this actually is not a problem because the pdf's are properly normalized $\int P(\kappa_{out}|T)d\kappa_{out} = 1$, the pdf for $P(T)$ can be taken as the normalization constant. Equation 2.3.5 makes it possible for the policy maker to update $P(\kappa_{out})$ after every temperature measure. Therefore the estimate of the climate sensitivity goes towards the actual climate sensitivity as more measurements are done. This means that the policymaker will make tax based on his current believe of the climate sensitivity and his uncertainty in this value. This means that it is assumed that the policymaker will tax in the following way:

$$\tau = h(E[\lambda](t), \sigma\lambda(t), t) \tag{2.3.6}$$

where h is an arbitrary function, $E[\lambda]$ is the mean climate sensitivity drawn from the pdf and $\sigma\lambda$ is the standard deviation in the climate sensitivity also drawn from the pdf. To guess this function the social cost of carbon is used. The social cost of carbon can be calculated given a certain climate sensitivity and is defined as:

$$\text{SCC} = \frac{\frac{\partial W}{\partial E}}{\frac{\partial W}{\partial C}} \tag{2.3.7}$$

With C the consumption, this gives the social cost of carbon the unit $\frac{\text{euro}}{\text{Gt}}$, i.e. euros per gigatonne CO_2. This can of course be any unit of money and any unit of carbon emissions. So the social cost of carbon measures how much consumption will increase the welfare as much as extra emissions will lowers it. For example, if the Social cost of carbon is 100 euros per tonne CO_2 it means that one tonne CO_2 is worth 100 euros. This means that it is better for the welfare if a company just emits this tonne than pay more than 100 euros. Hence this is exactly what the optimal tax should be. This means that the policy and hence the tax 2.3.6 can be written as:

$$\tau = h(\text{SCC}(E[\lambda](t), t), \sigma\lambda(t), t) \tag{2.3.8}$$

3 Methodology

The goal is to find $\tau(t)$ that maximizes the welfare as stated in equation 2.2.6 for current estimates of the climate sensitivity. First the policymaker would have to make a guess of the climate sensitivity, then the Social cost of carbon has to be estimated. Finally different policies are tried and are evaluated by calculating an estimated welfare. In figure 1. the schematic drawing of the Hot Small World Model is shown. It is important to keep this in the back of your mind. It only displays which important factors influence each other. The green lines in the chart mean positive feedback, while the red lines means negative feedback. So for example increasing the abatement lowers the output directly, but it also increases the output indirectly because it lowers the damage. In that case this climate cycle that works indirectly on the output will be a lot slower than the direct payment for abatement. In other words increasing the abatement is can be an investment for later.

3.1 Guessing The Climate Sensitivity

For the stochastic model it is important to remember that the underlying climate sensitivity still has a true value, this value determines how the climate will develop even though the policymaker does not know this value. The policymaker can only guess the climate sensitivity. He will update his current guess the climate sensitivity according to his observations of the temperature. Every time-step he measures the temperature and updates his guess for the outgoing radiation using Bayes' rule equation 2.3.5 and using equation 2.3.4 he has a guess. He will use this guess to adjust his policy, note that his guess is a probability density function so the policymaker can extract $E[\lambda]$ and $\sigma\lambda$ from the pdf. Assume that you have a pdf given by $p(\lambda)$ then the expectation value and the uncertainty can be calculated by

$$E[\lambda] = \sum_i \lambda_i p(\lambda) \tag{3.1.1}$$

$$\sigma\lambda = \sqrt{\sum_i [\lambda_i - E(\lambda)]^2 p(\lambda)} \tag{3.1.2}$$

Where λ_i are elements of a vector in which all possible values of λ are stored, $\lambda_i \in \{\lambda_0, \lambda_{\text{cutoff}}\}$. Here λ_0 is taken to be around 1.2 K and λ_{cutoff} is taken to be around 18.6 K. These are just the values that are allowed for the climate sensitivity, since we know quite certain that the climate sensitivity has a higher value than 1.5 K and a lower value than 10 K see figure 3. So guess for the climate sensitivity that has $\lambda_0 < 1.5$ and $\lambda_{cutoff} > 10$ will be a suitable. In figure 4. the Bayesian learning is shown, as can be seen the initial guess of the policymaker is wrong, and when time increases the guess of the policymaker goes towards the real climate sensitivity with more accuracy.

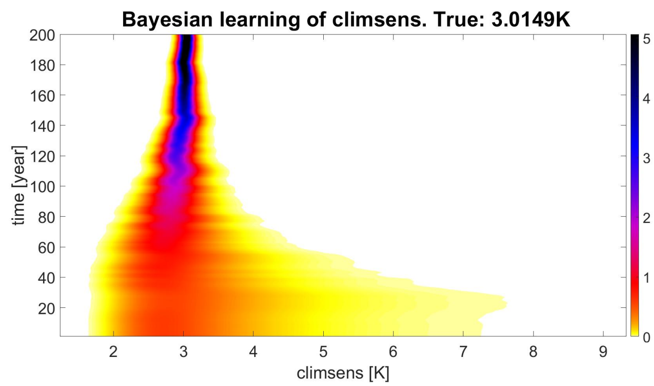


Figure 4: Bayesian learning the initial guess of the climate sensitivity is 3.14 K. The actual climate sensitivity is 3.0149 K. The color on the graph represents the probability density of that specific climate sensitivity at that specific time.

3.2 Social Cost of Carbon

The Welfare is an important factor for the calculation of the social cost of carbon. The Welfare is computed in the model in the following way:

$$W = \sum_{t_i}\sum_{j} U_j(t_i)\Big(\frac{1}{1+\rho(t_i\Delta t)}\Big)^{t_i\Delta t} \tag{3.2.1}$$

Where t_i is the number of iteration of the time-step and $\Delta t = 1$ is the size of one time-step. Then the sum over j means the sum over all firms. The utility of each firm at ti'th time-step is naturally given by equation 2.2.5 for $t = t_i$ The consumption per person at time-step t_i is given by 2.2.4 for $t = t_i$

The saving rate is taken as constant in this model, because it would only complicate the model and it will not have a significant impact on the Bayesian learning results. In this model the population is modeled in the following way to fit the RCP85 scenario[15] $N(t) = \frac{2}{1+e^{-0.03t}}$, t is the time at time-step t_i hence $t = t_i\Delta t$. For simplicity it is assumed that there is only one firm, one could see this firm as the world economy. The elasticity of the market is taken the same as in DICE [14], so $\chi_{elas} = 1.45$. Recall from section 2.3, the policy maker uses an estimate of the social cost of carbon (SCC) to set the carbon tax. Now looking at equation 2.3.7 it involves derivatives. It is not possible to calculate these derivatives analytically. This means that the derivative is approximated by:

$$\text{SCC} = \frac{\frac{W(Q,E+\Delta E)-W(Q,E)}{\Delta E}}{\frac{W(Q+\Delta Q,E)-W(Q,E)}{\Delta Q}} \tag{3.2.2}$$

When estimating the SCC, it is assumed that we have a reference scenario $W(Q, E)$ for the entire future. Even though one could argue that the policy maker, when estimating the SCC (and hence the carbon tax) does not exactly know the future, and thus the reference scenario does not need to be exactly the future, it is at least needed to have a reasonable guess. Note that $W(Q, E)$ in general depends on the climate sensitivity. It is assumed that the policy maker, when estimating the SCC, uses his current best guess of the climate sensitivity, $E[\lambda]$. Now ΔE is chosen to be 1/10 the emissions at that time plus 1/2 a gigatonne and ΔQ is taken to be 1/10 of the output at that specific time. This is different than equation 2.3.7, because the output is increased instead of the consumption, but since the two are related and have the same unit it does not matter if you increase the calculate the social cost of carbon using the output or the consumption. Now there is a problem for the policymaker who estimates the social cost of carbon it is not known what the abatement will be in the future. Several methods were tested to obtain the reference scenario for the abatement. The simplest way is to let the policy maker assume that abatement will remain constant in future. This is a pessimistic worldview, but one could argue that for climate change you can better be save than sorry. In this scenario the abatement for the social cost of carbon $\mu_{\mathrm{SCC}}(t)$ is given by:

$$\mu_{\mathrm{SCC}}(t) = \mu(t_0) \tag{3.2.3}$$

A slightly more sophisticated method would be to use Taylor expansion, a taylor expansion upto second order is used so this will result in the following $\mu_{\mathrm{SCC}}(t)$:

$$\mu_{\mathrm{SCC}}(t) = \mu(t_0) + \frac{\partial \mu}{\partial t}\Big|_{t_0} t + \frac{1}{2}\frac{\partial^2 \mu}{\partial t^2}\Big|_{t_0} t^2 \tag{3.2.4}$$

In equation 3.2.4 the partial derivatives are calculated with respect to the policy time-steps, because only every time the policy changes the tax the abatement will change. Since the policy time-step (dt_{pol} is not necessarily equal to the actual time-step dt this derivative will be calculated with respect to dt_{pol}. In the Hot Small World Model dt is set to one year, while dt_{pol} can be chosen as you like, in the DICE model it is assumed $dt_{\mathrm{pol}} = 5$ years. The derivatives from equation 3.2.4 are then calculated in the following way:

$$\begin{aligned}\frac{\partial \mu}{\partial t} &\approx \frac{\mu(t) - \mu(t - dt_{\mathrm{pol}})}{dt_{\mathrm{pol}}} \\ \frac{\partial^2 \mu}{\partial t^2} &\approx \frac{\mu(t) - 2\mu(t - dt_{\mathrm{pol}}) + \mu(t - 2dt_{\mathrm{pol}})}{dt_{\mathrm{pol}}^2}\end{aligned}$$

Only the backward difference method for calculating derivatives can be used, because the abatement in the future is not known.

It is important to understand the drawback of this method; because it is assumed that the tax is a function of the social cost of carbon the tax can only have indirect time dependence. In theory the carbon tax can be a function of the estimated climate sensitivity, the uncertainty in the estimation of the climate sensitivity and the time as shown in equation 2.3.6. If you look at equation 2.3.8 this is still valid, only in the results obtained it was assumed $\tau = \mathrm{SCC}(E[\lambda](t), \mu(t))$, this is of course a simplification. There is no time dependence and moreover the social cost of carbon is not a function of $\mu(t)$ it is a function of time and the optimal abatement as a function of time and the climate sensitivity is well known.This gives rise to a circular argument, you have to know the optimal deterministic policy to find the social cost of carbon. Imagine you are a policymaker, you do not know the optimal policy as a function of time and climate sensitivity, how would you be able to estimate the optimal tax? You could calculate the social cost of carbon, but then you had to estimate the abatement over the years. The only possible ways of doing that is looking back into the past and look at how the abatement had changed over the years. From this information you could estimate the social cost of carbon. This is a pessimistic simplification, you will probably have a higher carbon tax than optimal, however in the long run this will be a better policy than not taxing at all. In general a tax that is assumed to be $\tau = \text{SCC}$, will start taxing a lot in the beginning. This probably hurts the welfare a lot.

Therefore there is a need for a better estimation of μ_{SCC}; the best method would be to make an optimization of a deterministic version of the Hot Small World Model with various climate sensitivities, and let the policy maker assume that future abatement will follow this optimal trajectory. Since unfortunately the deterministic version of the Hot Small World Model could not be optimized numerically, optimized trajectories of the fairly similar DICE model to generate the policy maker's guesses of the reference scenario were used. The idea is to guess $\mu_{\text{optimal}}(\lambda, t)$ and use this as $\mu_{SCC}(t)$. This is done by assuming that the optimal policy in DICE follows the function:

$$\mu_{\text{optimal}}(\lambda, t) = \xi(\lambda) e^{\beta(\lambda) t} \tag{3.2.5}$$

By optimizing the Dice model for different λ $\xi(\lambda)$ and $\beta(\lambda)$ can be found. For a certain λ ξ can be found by using the following relations:

$$\begin{aligned} \xi(\lambda) &= \mu_{\text{optimal}}(\lambda, 0) \\ \beta(\lambda) &= \frac{\log\left(\frac{0.5}{\xi(\lambda)}\right)}{t_{\frac{1}{2}}} \end{aligned}$$

Where $t_{\frac{1}{2}}$ is the time at which $\mu_{\text{optimal}} = 0.5$. Then the final abatement that is used in the calculation of the social cost of carbon is:

$$\mu_{\text{SCC}}(t, E[\lambda]) = \xi(E[\lambda]) e^{\beta(E[\lambda]) t} \tag{3.2.6}$$

3.3 Calculating the Estimated Welfare

As discussed above it is now possible to calculate the welfare for different climate sensitivities, also the true climate sensitivity can be learned using Bayesian learning. Using the social cost of carbon the policy can be estimated, now the most important thing is to combine all these possibilities to calculate the estimated welfare for different taxes under the current estimate of the climate sensitivity. First it is important to model the current estimates of the climate sensitivity. As we know from Baker et all, the guess of the climate sensitivity is not normally distributed, but if the guess of the outgoing radiation is normally distributed the curve of the guess of the climate sensitivity can be replicated.[12] So it is assumed that the outgoing radiation is normally distributed, if the outgoing radiation is assumed to have a mean of $E(\kappa_{\text{out}}) = 1.23$ W/m^2 and a standard deviation of $\sigma\kappa_{\text{out}} = 0.3$ W/m^2, the curve for the climate sensitivity looks like the current estimates figure 3. From this distribution we can take a n_{mem} members, the probability of every member can be estimated using the distribution from figure 5. The n_{mem} members are linearly spaced between 1.5 K and 10 K. Because a climate sensitivity below 1.5 K and above 10 K are highly unlikely. So for

$$\lambda_i \in \mathbf{\Lambda},\ \lambda_i \in \{1.5, 10\}$$

With $\mathbf{\Lambda}$ a linearly spaced vector with n_{mem} elements over the range $\{1.5, 10\}$. For every λ_i a probability of this λ_i can be calculated, this will result in a weight ω_i. The weight is given

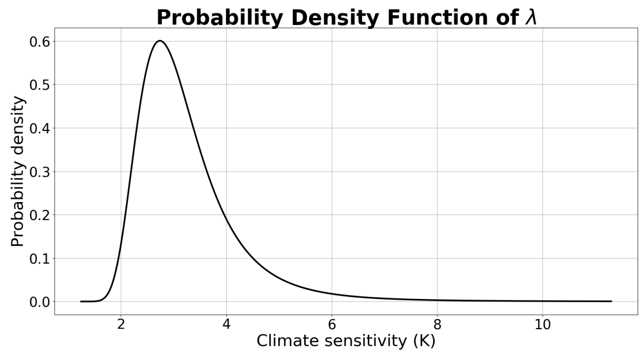


Figure 5: The probability density function (pdf) of the climate sensitivity that is assumed in the model.

by the distribution in figure 5., then for the weight the following relation holds:

$$\omega_i = p(\lambda_i) \tag{3.3.1}$$

With $p(\lambda_i)$ the distribution shown in 5. All these elements will result in a in a weight vector such that:

$$\mathbf{\Omega} = p(\mathbf{\Lambda}) \tag{3.3.2}$$

For every λ_i the model can be run and computes the welfare for a given λ_i. So the Hot Small World Model acts as a function that computes the welfare for a given climate sensitivity. This is of course all still dependent on the model parameters and model functions. But overly simplifying it can be seen that:

$$w_i = \text{HSM}(\lambda_i) \tag{3.3.3}$$

With w_i the welfare for λ_i and HSM the Hot Small World Model. All these elements of course form a welfare vector for a given $\mathbf{\Lambda}$:

$$\mathbf{W} = \text{HSM}(\mathbf{\Lambda}) \tag{3.3.4}$$

Then the expected welfare is defined as:

$$E[W] = \mathbf{\Omega} \cdot \mathbf{W} \tag{3.3.5}$$

This expected welfare depends on the policy, if equation 3.2.6 is close enough to the optimal tax, $\tau = SCC$ would be a good policy. however, estimating SCC via this approach seems to be too pessimistic (see results for details), leading to overly high taxes. therefore a scaling factor a is introduced then the tax is set to:

$$\tau = a\text{SCC} \tag{3.3.6}$$

Most DICE-like models ignore learning of the climate sensitivity. Therefore it is important to check if ignoring learning causes big deviations or whether the mistake of ignoring learning is small. This is done by running the model with Bayesian learning and without Bayesian learning and looking at the difference in estimated welfare between the two. Motivated by C. Wieners et all. [18] the performance of a policy (π) is defined; to say something about the significance in the expected welfare difference with and without Bayesian learning. The difference in performance between two policies is defined as:

$$\zeta(\pi_1, \pi_2) = 100\% \times \left(\frac{E[W]_{\pi_1} - E[W]_{\pi_2}}{E[W]_{\pi_2} - E[W]_0} + 1 \right) \tag{3.3.7}$$

Where $E[W]_{\pi_{1,2}}$ is the performance of policy $\pi_{1,2}$ and $E[W]_0$ is the expected welfare when there is no policy. In this specific case π_1 is the policy with learning and π_0 is the policy without learning. The performance is used to distinguish the effect two policies have on the welfare. It is the relative difference between two policies with respect to the business as usual scenario. If both policies are the same it means that both policies are equally effective hence $\zeta = 0\%$. So when policy π_1 is three times as effective as policy π_2 the performance is equal to 300%.

4 Results

In this section the results are given, first some general results of the Bayesian learning implementation are given. Then the results for the optimal abatement as a function of the time and climate sensitivity are given as in equation 3.2.6. Then some general results of the model are given to show how the model relates to the DICE model. Then the results of the optimal policies are given for different damage functions and different μ_{SCC}. After some results of this optimal policy are given. Lastly results for the estimate welfare with and without Bayesian learning are given.

4.1 Bayesian Learning

in figure 6 (next page) the results of the implemented Bayesian learning are shown. The true climate sensitivities are 1.5, 5.75 and 10 respectively. The tax in this simulation was set to equal the social cost of carbon. The colors in the plot shows the probability for a certain climate sensitivity that is on the x-axis, while time is on the y-axis. As can be seen the policy-maker alters his guess of the climate sensitivity and every time-step and learns the true climate sensitivity. The policymaker clearly learns the climate sensitivity in roughly 200 years except for the extreme case when the climate sensitivity is 10 K, then the policymaker does not have enough time to finish the learning because the actual value deviates a lot from the initial guess of the policymaker (3.14 K). The policymaker's uncertainty of the climate sensitivity also decreases overtime.

As can be seen in figure 6., the Bayesian learning seems to work but for high true climate sensitivity, the policy seems to struggle with guessing the climate sensitivity, this is because of the abatement. If the abatement maxes out, the only change in temperature he sees is due to the noise. This makes it hard to guess the climate sensitivity, because the temperature change is dominated by this noise and therefore if the assumed noise is not the actual noise, this is impossible to filter out. If the abatement is taken out the climate sensitivity guess will be a lot more accurate as shown in figures 7. and 8.

It seems to be hard to guess the climate sensitivity when it is so far away from the initial assumed climate sensitivity, but after 400 years the guess for the climate sensitivity is 9.8 K, this seems about right. While with policy and thus abatement the guess of the climate sensitivity is 9.1 K after 400 years. So the abatement makes a big difference. However this is not such a big problem, because if the abatement is already at 1 the guess of the climate sensitivity does not really matter, because the guess of the climate sensitivity is only used to make a policy. So however it is harder to guess the policy whenever the climate sensitivity becomes higher it is not a problem for the policymaker as abatement 1 is reached long before then. In general the Bayesian learning in the model yield reliable results.

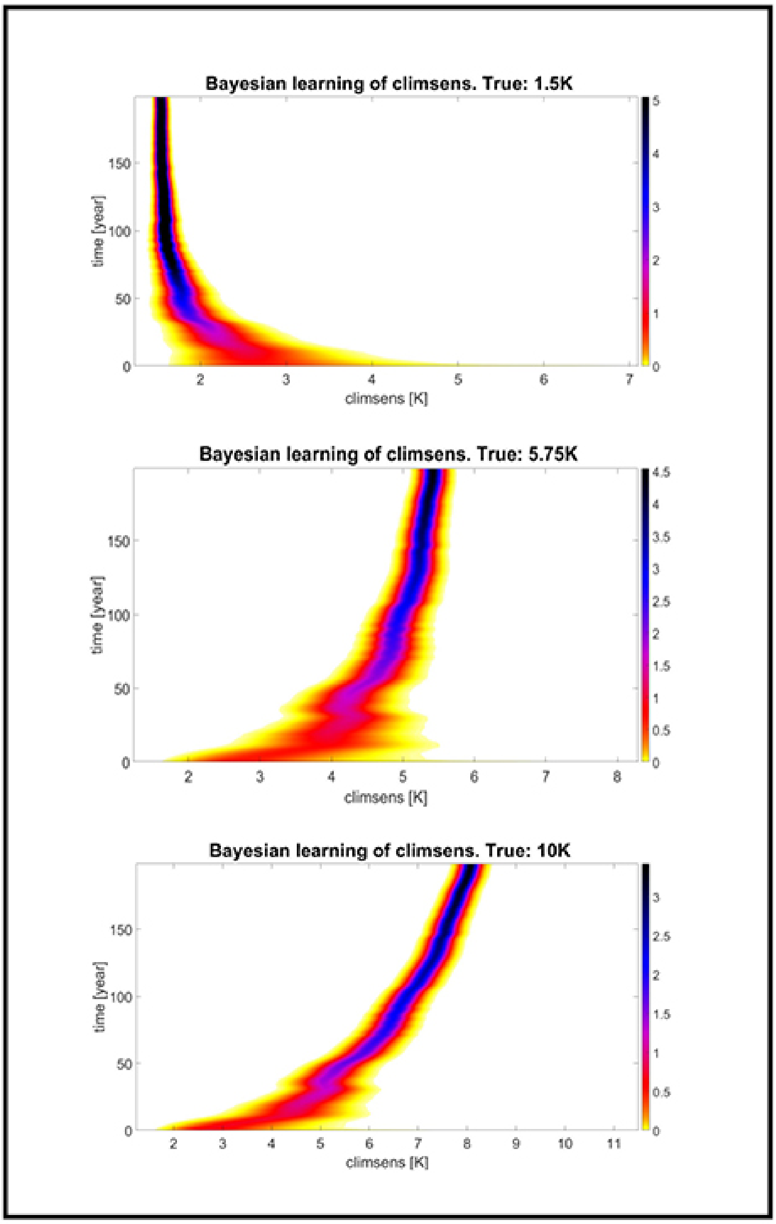


Figure 6: Bayesian learning the initial guess of the climate sensitivity is 3.14 K. The actual climate sensitivity is 1.5 K in the first plot, 5.75 K in the second plot and 10 K in the last plot. The color on the graph represents the probability density of that specific climate sensitivity at that specific time.

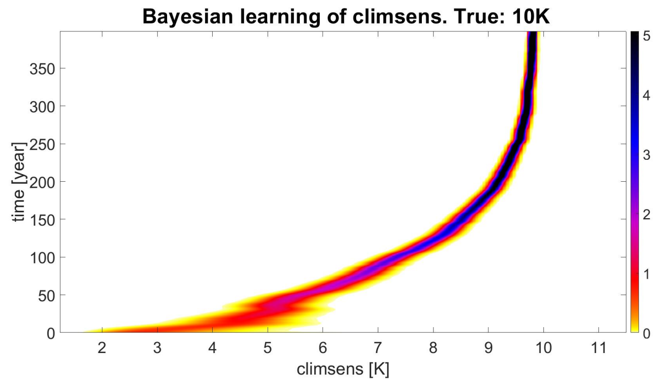


Figure 7: Bayesian learning with policy, the initial guess of the climate sensitivity is 3.14 K. The actual climate sensitivity is 10 K. The color on the graph represents the probability density of that specific climate sensitivity at that specific time. The simulation runs for 400 years

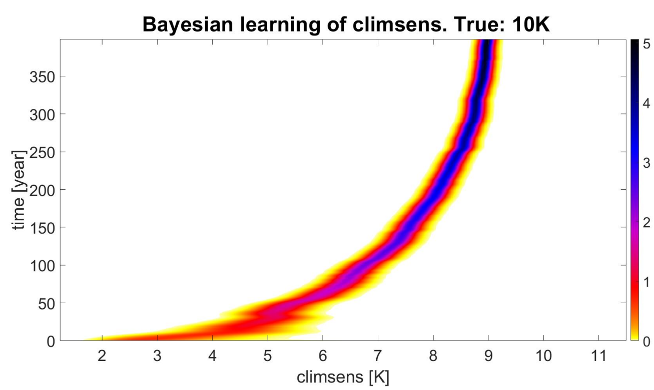


Figure 8: Bayesian learning without policy, the initial guess of the climate sensitivity is 3.14 K. The actual climate sensitivity is 10 K. The color on the graph represents the probability density of that specific climate sensitivity at that specific time. The simulation runs for 400 years

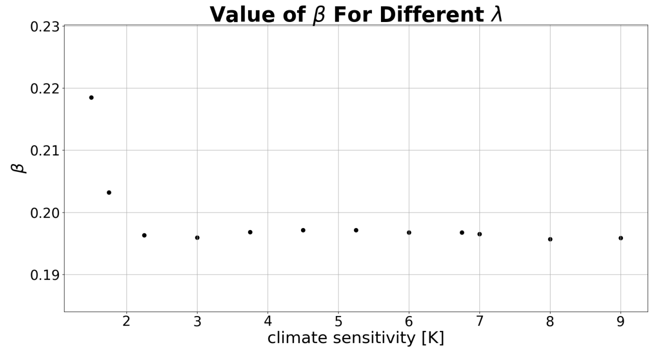


Figure 9: The β parameter for different values of the climate sensitivity. These results were obtained using equation 3.2.5.

4.2 Optimal Abatement DICE Results

By optimizing DICE and finding the value for ξ and β for different climate sensitivities a certain curve can be fitted through the results. The results for $\beta(\lambda)$ are given in figure 9. The results for $\xi(\lambda)$ are given in figure 10. As can be seen in figure 9. for small climate sensitivity β deviates from its average. For small climate sensitivity the effect of β on the optimal tax is negligible. Therefore $\beta(\lambda)$ is taken to be constant for its value $\lambda > 2$. This yields a value $\beta = 0.1965$. In figure 10. a curve is fitted through the data, also a generic data point is added namely the point $(0, 0)$ this is because it is known that the optimal abatement is 0 when the climate sensitivity is 0. This helps the fitting procedure, the fitted curve is given by:

$$\xi(\lambda) = 0.0137\text{erf}(0.6223(\lambda - 2.2868) + 1)\lambda^{0.7379}$$

Where $\text{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t} dt$ and the float numbers are fitted parameters. This is just done by trial and error. Having these results for $\beta(\lambda)$ and $\xi(\lambda)$ those can be used to calculate $\mu_{\text{SCC}}(t, \lambda)$ using equation 3.2.6:

$$\mu_{\text{SCC}} = 0.0137\text{erf}(0.6223(E[\lambda] - 2.2868) + 1)E[\lambda]^{0.7379} e^{\beta t} \tag{4.2.1}$$

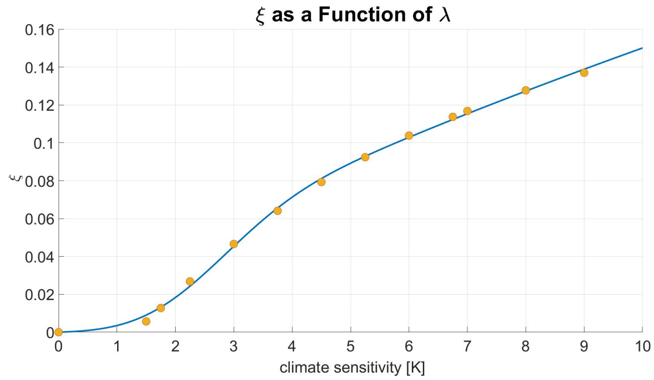


Figure 10: The ξ parameter for different values of the climate sensitivity. These results were obtained using equation 3.2.5. The line through the model is the fitted equation which is given in equation 4.2.1

4.3 General Results

In figure 11. and 12. the temperature and the emission control in DICE2013[9] are shown. Emission control is the same as abatement, the base scenario in figure 12. does not stay zero, because the abatement cost lower over time and some abatement becomes free. In figure 13. and 14. the temperature and the abatement are shown for the Hot Small World Model with no learning effect on the abatement like in equation 2.2.10, but abatement will get cheaper over time like in DICE. The tax in these simulations is set equal to the social cost of carbon. The abatement used in the calculation in the social cost of carbon is calculated using the results of equation 4.2.1. As can be seen from the figures the Hot Small World Model follows roughly the DICE mode. The biggest difference is that the Hot Small World Model naturally abates more than the DICE model. This is because of the pessimistic estimated social costs of carbon of the Hot Small World Model. Despite the differences the Hot Small World Model does not behave weirdly this is an indication of the validity of the Hot Small World Model.

The fact that the SCC policy differs from the optimal DICE policy is probably because the social cost of carbon is not the actual social cost of carbon, but it is just an estimation of the social cost of carbon. This will be discussed in a lot more detail in section 4.4. The estimation of the social cost of carbon pessimistic than the actual social cost of carbon, in theory the actual social cost of carbon should give a policy like DICE's optimal policy. Because the SCC policy is pessimistic the social cost of carbon calculated in the Hot Small World Model will be a lot higher than the actual social cost of carbon. This therefore will result in a higher tax than the theory would give us and therefore the abatement are not the same.

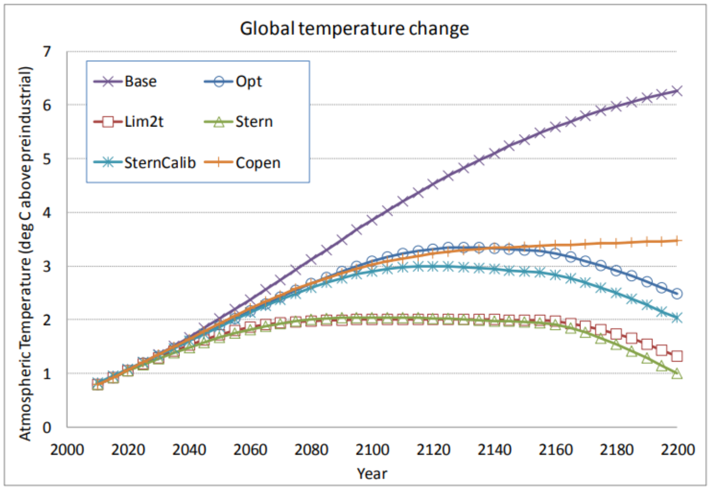


Figure 11: Evolution of the temperature for different scenarios in DICE2013[14], the interest lays primarily on the optimal case this should be roughly compare-able to figure 13.

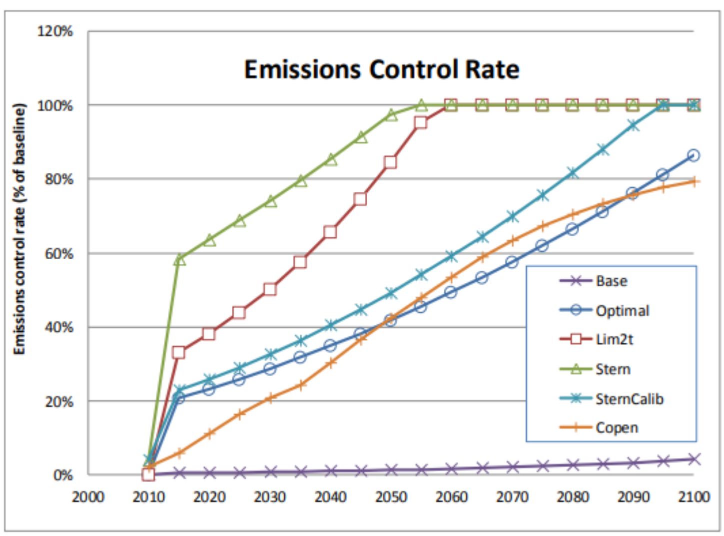


Figure 12: Abatement for different scenarios in DICE2013[14], the interest lays primarily on the optimal case this should be roughly compare-able to figure 14

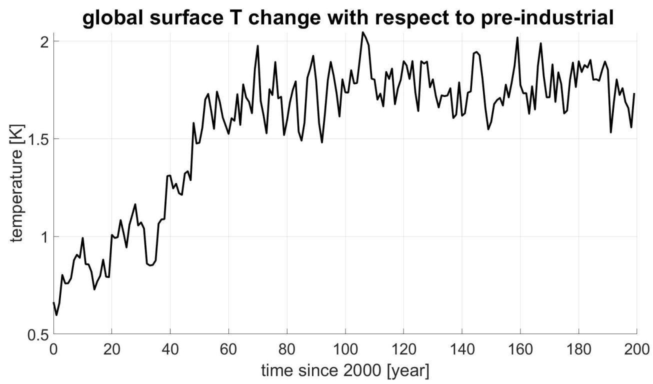


Figure 13: Temperature of the Hot Small World Model with abatement learning like in DICE, so no learning by doing. The tax is equal to the estimated social cost of carbon. Note that the temperature is stochastic.

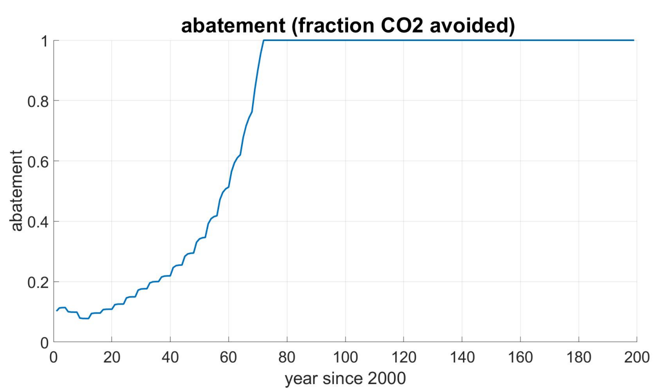


Figure 14: abatement of the Hot Small World Model with abatement learning like in DICE, so no learning by doing. The tax is equal to the estimated social cost of carbon. Note that $dt_{pol} = 4$.

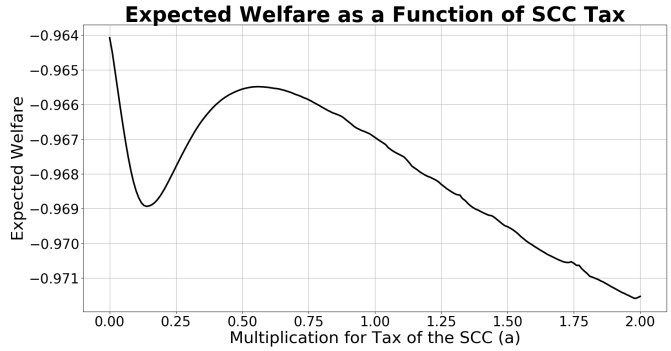


Figure 15: The expected welfare as a function of a from equation 3.3.6, DICE damage is assumed. The simulation runs for 200 years. μ_{SCC} is calculated using equation 3.2.4.

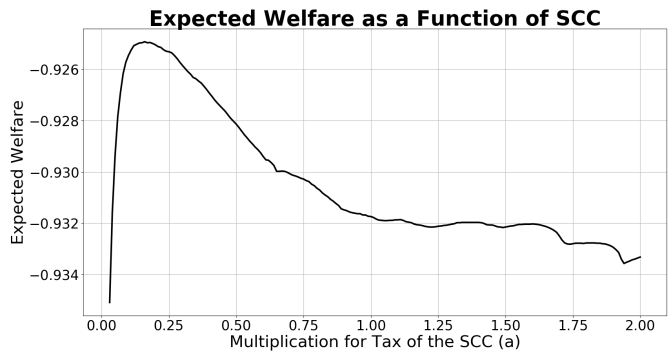


Figure 16: The expected welfare as a function of a from equation 3.3.6, Weitzman damage is assumed. The simulation runs for 200 years. μ_{SCC} is calculated using equation 3.2.4.

4.4 finding the best SCC estimate

Now the performance of different ways by which the policy maker can estimate the SCC are compared (See section 3.2). Firs the Taylor method see equation 3.2.4 is discusseed.This is done for two damage functions: the one from DICE see equation 2.2.8 and the one by Weitzmann see equation 2.2.9. As explained in section 3.3, the estimate for the SCC performs well if a in equation 3.3.6 equals 1, i.e. if it is optimal to set the tax exactly equal to the estimated SCC. However, in figure 15. it is shown that for the DICE damage function it is optimal to do no abatement and in figure 16. it is shown see that for Weitzman damage the optimal tax is equal to 0.16SCC. So clearly the Taylor method yields bad estimates for the SCC, why Taylor yields bad results was discussed in section 3.2. However most probably for a longer simulation for example 400 years or longer the optimal tax will most probably go towards the sub-optimal policy that is now around $a = 0.52$, because the damage for no policy will increase a lot as time goes on. For no policy the expected global temperature increase will only be around 6 K after 200 years as shown in figure 18., in 200 years the damage will be relatively low because of this. The welfare decrease due to the high abatement at the beginning by assuming $\tau = a\text{SCC}$ apparently is higher than the decrease in welfare due to the damage. In figure 16. it can be seen that $a_{\max} = 0.16$. This is lower than

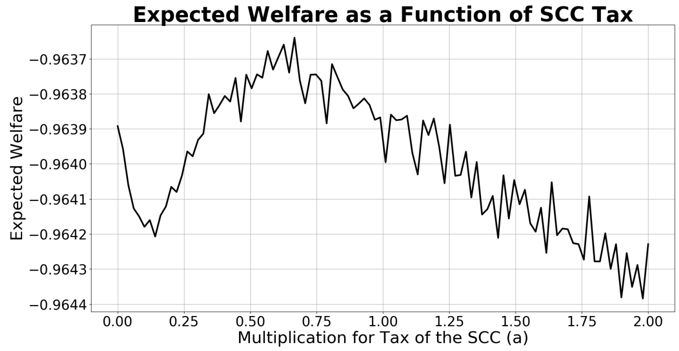


Figure 17: The expected welfare as a function of a from equation 3.3.6, $dt_{\text{pol}} = 4$, DICE damage is assumed. The simulation runs for 200 years. μ_{SCC} is calculated using 4.2.1.

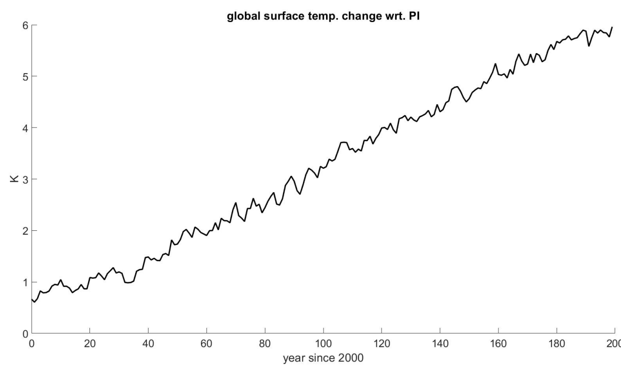


Figure 18: Temperature in the Hot Small World Model without policy.

expected, but this is due to the reason discussed above. At least for a more drastic damage function the model gives that the tax should be non zero. Again an important fact is that an oversimplification of the tax has been made and this tax is not the optimal policy for a policymaker. Therefore the estimate based on the optimal policy for DICE is used see section 3.3. for only the DICE damage function because the model was optimized using this damage. As can be seen in figure 17. there is an optimal value for the tax at $a_{max} = 0.667$. This is clearly the superior estimation of the social cost of carbon. It is still not an optimal path, this can be due to the assumption that μ_{optimal} for DICE has an exponential shape. Another thing is that the Hot Small World model is not exactly DICE so it will not behave exactly the same so taking μ_{SCC} from optimizing DICE is sub-optimal. This means that using DICE to estimate μ_{SCC} performs better than the Taylor method, this estimation of calculating the SCC is used to obtain all following results. In these simulations the total time was 200 years and the time over which the policymaker makes his assumptions for the SCC is also 200 years. Now in the remainder of this sections overall results are given of the model using the optimized DICE method to calculate the SCC. In figures 19., 20., 21., 22., 23. and 24. the Temperature, abatement, loss, emissions, output and expected welfare are shown for the tax $\tau = 0.667$SCC and the social cost of carbon is calculated using the optimization of DICE using equation 4.2.1. In these figures the climate sensitivity is set to 3.24 K because that is the expected climate sensitivity from the pdf in figure 5. Overall the results look promising, the temperature does not increase too much and it is a bit saver than DICE 2013. The abatement increases more, this is because the climate model is a bit less optimistic in the Hot Small World model than in DICE. Also abatement is probably a bit cheaper than in DICE because the Hot Small World model uses learning by doing on the abatement, so more abatement means that you know more about abatement so your abatement becomes cheaper. Therefore the abatement will rise quicker than in DICE.

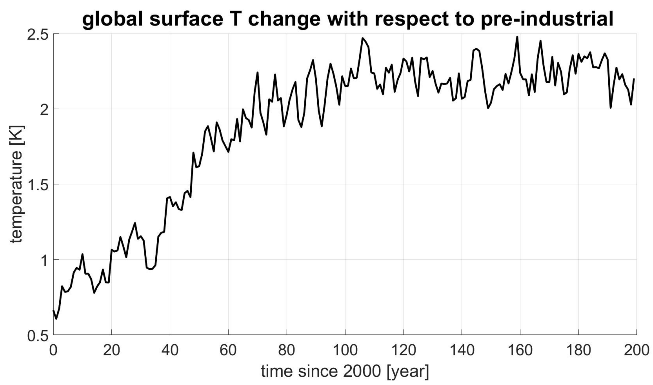


Figure 19: Global temperature increase with respect to pre-industrial with DICE damage for the tax equal to 0.667SCC. With climate sensitivity 3.24 K. The μ_{SCC} is calculated using 4.2.1.

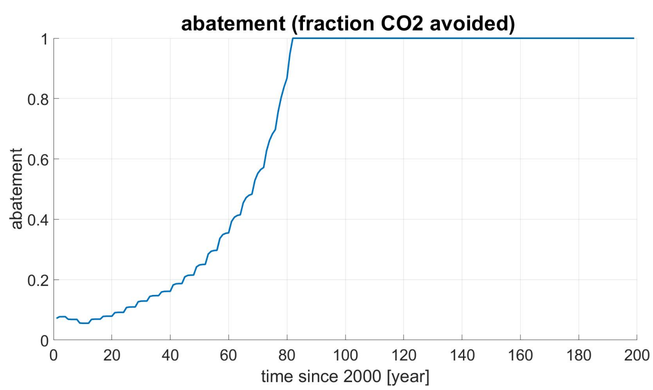


Figure 20: abatement done by firms over the years with DICE damage for the tax equal to 0.667SCC. With climate sensitivity 3.24 K. The μ_{SCC} is calculated using 4.2.1.

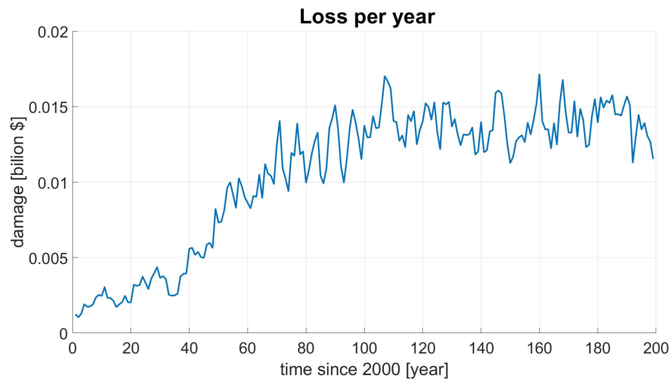


Figure 21: Loss of the global economy so $n_{firm} = 1$ with DICE damage for the tax equal to 0.667SCC. With climate sensitivity 3.24 K. The μ_{SCC} is calculated using 4.2.1.

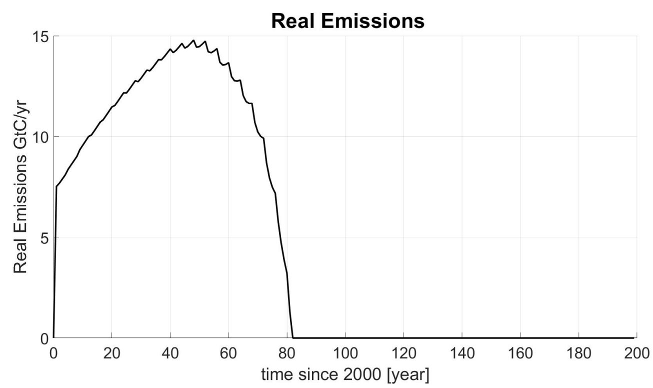


Figure 22: Total emissions with $n_{firm} = 1$ with DICE damage for the tax equal to 0.667SCC. With climate sensitivity 3.24 K. The μ_{SCC} is calculated using 4.2.1.

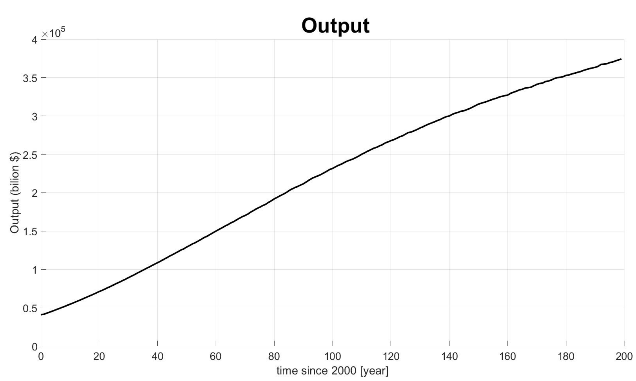


Figure 23: Output of the global economy so $n_{firm} = 1$ with DICE damage for the tax equal to 0.667SCC. With climate sensitivity 3.24 K. The μ_{SCC} is calculated using 4.2.1.

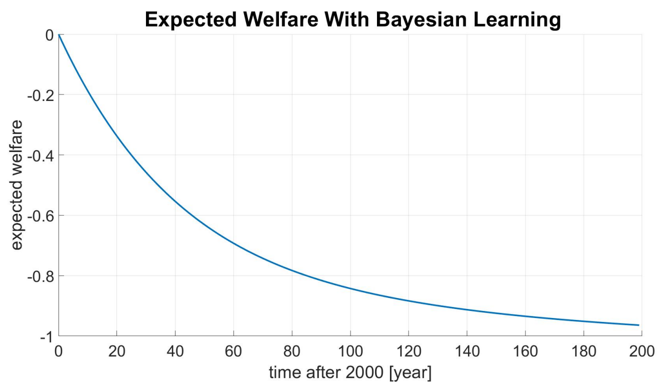


Figure 24: Expected welfare over time with $n_{firm} = 1$ with DICE damage for the tax equal to 0.667SCC. With climate sensitivity 3.24 K and with Bayesian learning. The μ_{SCC} is calculated using 4.2.1.

4.5 Estimated Welfare With and Without Bayesian Learning

Now the main question is investigated, namely whether Bayesian learning increases the welfare - compared to a world in which the policy maker does not learn but always assumes that the climate sensitivity equals 3.24K, i.e. the mean of the distribution. Figure 24. shows the expected welfare as a function of time when the policy learns about the climate sensitivity. In figure 25 the expected welfare without Bayesian learning is shown. Now an interesting thing is if that actually makes a difference from not learning. In figure 26. the difference in the expected welfare with and without Bayesian learning is displayed. A positive difference means that Bayesian learning has a positive impact and a negative difference means that Bayesian learning has a negative effect. As can be seen Bayesian learning has a positive effect on the expected welfare.

As can be seen in figure 26. there is a difference between Bayesian learning and no Bayesian learning. The estimated welfare when you use Bayesian learning is higher than when you don't. As time increases this difference gets bigger because the impact of not knowing your climate sensitivity gets higher, however at a certain point this difference will stay constant because in both cases the abatement already reached 1. It is interesting to see that Bayesian learning still has a positive impact after 200 years because the slope is still going up. This shows that the Damage that is caused by doing too less abatement for the climate sensitivity actually still hits after 200 years, this is why it is important to keep doing research in this area. The difference of the expected welfare is in the order of 10^{-4}, this does not seem much, but changes in the welfare are usually not big. So it is a relatively big difference and cannot be disregarded.

To dig deeper into this the performance of the learning effect can be calculated using 3.3.7. The difference in expected welfare of the policies at the end of the simulation is: $1.5160e-4$. The difference between no learning and no policy (the numerator of equation 3.3.7) is equal to $4.1284e-5$. This means that the performance of the learning effect is given by:

$$\zeta = 467\%$$

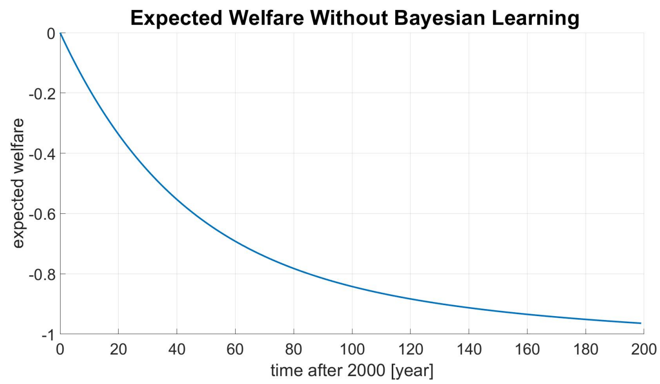


Figure 25: Expected welfare over time with $n_{firm} = 1$ with DICE damage for the tax equal to 0.667SCC. With climate sensitivity 3.24 and no Bayesian learning. The μ_{SCC} is calculated using 4.2.1.

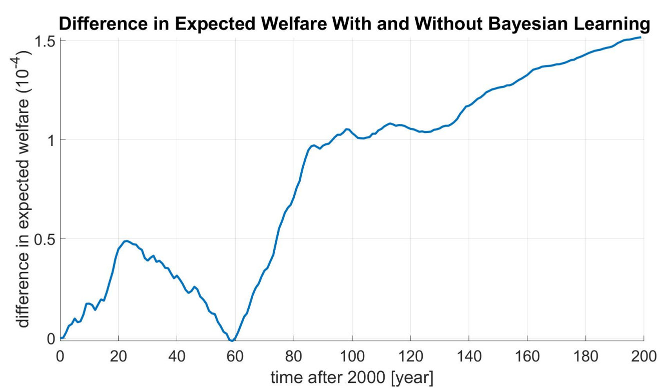


Figure 26: Difference in expected welfare with and without Bayesian learning over time with $n_{firm} = 1$ with DICE damage for the tax equal to 0.667SCC. With climate sensitivity 3.24 K. The μ_{SCC} is calculated using 4.2.1. With climate sensitivity 3.24 K

5 Discussion

5.1 Possible Problems With the Model

The first problem that arrises with the Hot Small World Model is that the the SCC is calculated using the estimated optimal abatement from DICE (see section 3.2). This is because the numeric optimization of the deterministic version of the Hot Small World Model did not work. One hypothesis for its vailer is that indeed all tax values above a certain threshold yield the same optimal abatement, namely $\mu = 1$. To be more precise, assume that for a tax of x euro per gigatonne CO_2 the abatement is equal to one. Then of course:

$$\tau = x + y \text{ where } y \geq 0 \quad \Rightarrow \quad \mu = 1$$

So at later times, where abatement should be =1, the optimization cannot tell which tax (above the threshold) it should take; this might cause the optimization to fail. In DICE the problem can be easily fixed by requiring an upper bound abatement $\mu \neq 1$. Here the upper bound for the tax is not known a priori.

An additional difference between our model and DICE is that climate policy always extracts money from the economy. After all, if there is a carbon tax, then firms pay both if they do abate they pay abatement costs and if they don't abate they pay taxes. Of course, the tax money is not entirely lost to the economy, as it may be payed back in the form of subsidies or damage compensation, but this can happen at a time lag, so the tax funds effectively lose value due to discounting and the rate of impatience $\rho > 0$.

This phenomenon may make it disadvantageous to prescribe modest values of the tax and hence abatement, as can also be seen from the local minimum at $a \approx 0.17$ in figure 15. and $a \approx 0.15$ in figure 17. In those cases, the tax is high enough to extract significant money from the firms, but too low to induce significant abatement, so it is better to either put the tax to zero to avoid tax costs for the firms or make it higher, so that the tax at least induces strong abatement. Thus we may have multiple local maximums of the carbon tax, hampering the optimization process.

Furthermore some of the problems of DICE also were inherited by the Hot Small World Model. The DICE model has been criticized for being overly simple [19]. In particular, it employs a very simple damage function for assessing the material and immaterial cost of climate change [7], which ignores the fact that damages can be irreversible or delayed (for example, slow melt of the ice on the artics) in later model versions it received minder updates, despite new studies on the subject[15]. Neither does it include climate adaptation. In addition, DICE has an overly simplified energy sector with exogenous costs for CO2 reduction and does not include negative emission techniques. Finally, assuming only one policymaker, it disregards the possibility of competitions or collaboration between policymakers.[20]

Lastly the implementation of learning-by-doing on the abatement is not very well looked at. There is no experimental proof that the learning-by-doing effect described by equation 2.2.10 is the actual learning effect on the abatement. In addition, we did not add limitations on the jumps of the abatement. It is not realistic that the abatement will jump by much, especially it will not be realistic that the abatement will go down a lot after reaching 1. In principle it is possible for the abatement to jump from 1 to 0 in one dt_{pol} and then return back to 1 again. If this actually would happen it would bring a lot of costs with it, you would have to destruct all the infrastructure and technology that made the abatement happen only to build it up again in the next time step.

5.2 Further Research

A lot of interesting research can be done on this model. First of all it would be interesting to uptimize DICE using Weitzman damage function see equation 2.2.9. This owuld give a better feel of the impact of the damage function on the social cost of carbon. As can be seen in figures 15. and 16. Weitzman gives a better estimation of the social cost of carbon when using the Taylor method (see equation 3.2.4).

It would also be interesting to make the damage stochastic to allow for tipping points. This would be more realistic, for further reading on how this implementation would work see C. Wieners et all. [20]

Furthermore it would be interesting to look at the optimization of the fraction of the tax that is used for subsidy in abatement. As of now the Hot Small World Model uses a fund and first asks the firms to do abatement before compensating damages as discussed in sections 2.2.3 and 2.2.4. Will optimizing this factor give significantly different results? Is it better to use the tax only to compensate damages or maybe use it to only subsidize abatement?

In addition as for now we did not look at the impact of different firms, in theory it is possible to run the model for different firms and look at the difference that might make. Especially how competition between firms would change the outcome. Competition still needs to be implemented, this can be done using a matrix in which the connections between firms are described, a positive matrix element would mean that firms cooperate and a negative matrix element means that there is competition between the two firms.

An other interesting thing is implement the energy sector in a better way. Ideally you would want to have a clean energy sector and a dirty energy sector. This means that you do not assume abatement costs but you let the firms choose between green en grey energy like in F. Lamperti et all [8]. This however would be hard to implement on short timescale. Something that is in principle possible is to run the model all over again and assume the abatement costs of DICE[14] to see if the effect is of any importance for the Hot Small World Model.

Finally another thing to study is using the carbon tax as a way to compete with different countries. This means that you will have n_{firms} representing countries, but every country could choose its own tax. A low tax would make it more attractive for firms to settle in that country. This would simulate the real world economy better because it is not realistic that the entire world would agree on setting a carbon tax. It would however be hard to optimize for all countries. Something that would be possible however is to let the tax increase in all but one country and look if it is optimal for that one country to introduce a carbon tax.

6 Conclusion

the aim of this study was to assess whether future Bayesian learning of the climate sensitivity improves the welfare. This question is investigated using a simple model. The results show that Bayesian learning improves the performance of the policy by $\zeta = 476\%$, this means that learning will make the policy perform 4 times better.

To conclude the learning process of the policy maker is important, and omitting this process might cause significant errors. We believe that it would be useful to incorporate Bayesian learning into more sophisticated models. This might both validate our results and potentially make these other models more realistic.

References

[1] Arrhenius, S. (1896). On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground, Philosophical Magazine and Journal of Science, 41 (5), 237-276.

[2] NASA, (2019, May 23), Scientific Consensus: Earth's Climate is Warming, https://climate.nasa.gov/scientific-consensus/. Retrieved : 2019, July 9

[3] Saad, L. (2019, March 25), Americans as Concerned as Ever About Global Warming, https://news.gallup.com/poll/248027/americans-concerned-ever-global-warming.aspx. Retrieved: 2019, July 9

[4] Aengenheyster, M. Feng, Q. van der Ploeg, F. Dijkstra, H. A. (30-08-2018). The point of no return for climate action: effects of climate uncertainty and risk tolerance, Earth System Dynamics, 9 (3), 1085-1095.

[5] Viscide, S. Graham, N. (2019, January, 4), Brazil Was a Global Leader on Climate Change. Now It's a Threat., Foreign Policy.

[6] Goodland, R. Anhang, R. (2009). Livestock and Climate Change What if the key actors in climate change are... cows, pigs, and chickens? Washington D.C., USA: World Watch

[7] Nordhaus, W. D. Boyer, J. (2000). Warming the World Economic Models of Global Warming. London, England: The MIT Press.

[8] Lamperti, F. Dosi, F. Napoletano, M. Roventini, A. Sapio, A. (2017 April). Faraway, so Close: Coupled Climate and Economic Dynamics in an Agent-Based Integrated Assessment Model.

[9] Watts, A. (2012, March 19), Lower Climate Sensitivity Estimates: New Good News, https://wattsupwiththat.com/2012/03/19/lower-climate-sensitivity-estimates-new-good-news/. Retrieved: 2019, July 9

[10] Weitzman, M.L. (2009), On Modelling and Interpreting the Economics of Catastrophic Climate Change, Review of Economics and Statistics, 91, (1), 1-19.

[11] Traeger, C. (2015 April), Analytic Integrated Assessment and Uncertainty, *Working paper.*

[12] Lemoine, D. Rosenthal, M. Rudik, (2018, February 6) I. General Bayesian Learning in Dynamic Stochastic Models: Estimating the Value of Science Policy, *working paper.*

[13] Climate Interactive (2019), C-ROADS, https://www.climateinteractive.org/tools/c-roads/. Retrieved: 2019, July 9

[14] Nordhaus, W. D. Sztorc, P. (2013). DICE 2013R: Introduction and Users Manual.

[15] IPCC AR5, (2014). Climate Change 2014 Synthesis Report.

[16] Weitzman, M. L. (2010), What is The Damages Function For Global Warming And What Difference Might it Make?, Climate Change Economics, 1 (1), 5769.

[17] Roe, G. H. Baker, M. B. (2007). Why Is Climate Sensitivity So Unpredictable?, Science, 318, 629-632.

[18] Helwegen, K. G. Wieners, C. E. Frank, J. E. and Dijkstra, H. A. (30 January 2019). Complementing CO_2 emission reduction by Geoengineering might strongly enhance future welfare, *working paper*.

[19] Pindyck, R. S. (2017). The Use and Misuse of Models for Climate Policy, *working paper*.

[20] Helwegen, K. G. Wieners, C. E. Frank, J. E. and Dijkstra, H. A. (30 January 2019). Complementing CO_2 emission reduction by Solar Radiation Management might strongly enhance future welfare, *working paper*.

7 Appendix

Table with all variables and parameters if the value is open it means that the symbol represents a variable.

Symbol	Explanation	Equation	Value
P_N	net production of plants	2.1.1, 2.1.2 and 2.1.4	
P_{N_0}	initial net production	2.1.1	85.1771 Gt/year
p_{fert}	CO_2 fertilization for plants	2.1.1	0.42 Gt^{-1}
C_{at}	atmospheric CO_2 concentration	2.1.1, 2.1.4, 2.1.8, 2.1.9 and 2.1.12	
C_{at_0}	initial atmospheric CO_2 concentration	2.1.1, 2.1.8, 2.1.9 and 2.1.12	769.3061 Gt
p_{heat}	heat stress in plants	2.1.1	-0.01 K^{-1}
ΔT	temperature change	2.1.1 and 2.1.8	
C_{bio}	CO_2 in biosphere	2.1.2, 2.1.3 and 2.1.4	
τ_{bio}	residence time for biosphere	2.1.2, 2.1.3 and 2.1.4	10.6 year
C_{hum}	CO_2 in the humus	2.1.3 and 2.1.4	
p_{hum}	humification fraction	2.1.3 and 2.1.4	0.428
τ_{hum}	residence time for humus	2.1.3 and 2.1.4	27.8 year
E	Carbon emissions	2.1.4, 2.1.10, 2.2.7 and 3.2.2	
OCEAN	carbon in the atmosphere due to ocean processes	2.1.4	
$M_{k,k+1}$	carbon flux between layer k and k+1	2.1.5, 2.1.6 and 2.1.7	
χ_{eddy}	diffusivity of eddies in ocean	2.1.5 and 2.1.14	4400 m^2/year
$C_{on}(k)$	CO_2 in ocean layer k	2.1.5, 2.1.6, 2.1.7 and 2.1.8	
$d_{lay}(k)$	thickness of ocean layer k	2.1.5, 2.1.14 and 2.1.16	
$C_{mix,ref}$	reference CO_2 in the mixing layer	2.1.8	1047.3 Gt
R_T	T dependence of CO_2 in mixed layer	2.1.8	0.003 K^{-1}
R_0	Revelle factor	2.1.9	9.7
R_C	CO_2 dependent Revelle factor	2.1.9	3.92

Symbol	Explanation	Equation	Value
λ	climate sensitivity	2.1.11, 2.3.3, 2.3.6, ??, 2.3.8, 3.1.1, 3.1.2 3.3.1 and 3.3.3	
f_{CO_2}	scale of radiative forcing	2.1.11 and 2.1.12	5.35 W/m^2
κ_{out}	outgoing radiation	2.1.11, 2.1.13, 2.3.2, 2.3.4 and 2.3.5	
F_{CO_2}	radiative forcing	2.1.12	
F_{out}	forcing due to outgoing radiation	2.1.13	
$\mathcal{H}_{k,k+1}$	heat flux from layer k to layer $k+1$	2.1.14 and 2.1.15	
$H_{on}(k)$	heat in layer k	2.1.14, 2.1.15 and 2.1.16	
F	Energy Forcing into ocean	2.1.15	
α_{yr}	seconds to year transformation	2.1.15	31557600 s/year
A_{sea}	area of the sea surface	2.1.15	361.1508e6 km^2
$T_{on}(k)$	temperature in layer k	2.1.16	
c	heat capacity of ocean per volume	2.1.16	4.23e6 J/m^3
Q_{gross}	gross output	2.2.1, 2.2.8 and 2.2.9	
A_0	constant that determines initial gross output	2.2.1	1359.7e9 dollar
K	capital	2.2.1 and 2.2.3	
L	labor force	2.2.1	
γ	fraction that determines how much capital and labor force contribute to the output	2.2.1	0.3
K_0	initial capital	2.2.3	85.34e12 dollar
ΔK	capital depreciation	2.2.3	5.526e12 dollar/year
s	saving rate	2.2.3	0.23
Q	net output	2.2.3 and 2.2.2	
μ	abatement	2.2.7, 2.2.14, 2.2.13, 2.2.12, 2.2.11, 2.2.10, 3.2.3 and 3.2.4	
E_0	initial emissions	2.2.7	8 Gt
α	technological advancement	2.2.7	0.0057 year^{-1}
D	Damage	2.2.8, 2.2.9 and 2.2.2	
ψ_0	damage parameter DICE	2.2.8	0.0028388
ψ_1	damage parameter Weitzman	2.2.9	0.00245
ψ_2	damage parameter Weitzman	2.2.9	0.0000198

Symbol	Explanation	Equation	Value
W	welfare	2.2.6, 3.2.1, 3.3.3, 3.3.4 and 3.3.5	
U	utility	2.2.6 and 3.2.1	
ρ	impatience rate	2.2.6 and 3.2.1	
χ_{elas}	elasticity of the market	2.2.5	1.45
C_{pp}	consumption per person	2.2.4 and 2.2.5	
ρ_0	initial rate of impatience		0.03
g_ρ	depreciation of rate of impatience		0.00257 year^{-1}
E_{gross}	gross emissions	2.2.14, 2.2.13 and 2.2.11	
$c(\mu, E_{gross}, c_\mu, \sigma_\mu)$	cost of a firm	2.2.13 and 2.2.2	
c_μ	cost of abatement	2.2.13 and 2.2.11	
l_μ	learning effect of abatement	2.2.13, 2.2.11 and 2.2.10	
σ_μ	subsidy for abatement	2.2.13 and 2.2.12	
τ	tax	2.2.13, 2.3.6, 2.3.8, 2.2.12 and 3.3.6	
f	fund of tax that was collected	2.2.12 and 2.2.13	
g_0	costs when no abatement is done	2.2.10	$205e9$ dollar
g_l	learning effect of the abatement	2.2.10	$\frac{1}{3}$
T	measured surface temperature	2.3.1, 2.3.4 and 2.3.5	
T_λ	non stochastic temperature increase	2.3.1	
δT	noise on surface	2.3.1	0.3 K
SCC	social cost of carbon	2.3.7, 3.2.2, 2.3.8, and 3.3.6	