Discounting the Future: on Climate Change, Ambiguity Aversion and Epstein-Zin preferences*

Stan Olijslagers† Sweder van Wijnbergen‡

May 23, 2022

Abstract

We show that deviations from standard expected time separable utility have a major impact on estimates of the willingness to pay to avoid future climate change risk. We propose a relatively standard integrated climate/economy model but add stochastic climate disasters. The model yields closed form solutions up to solving an integral, and therefore does not suffer from the curse of dimensionality of most numerical climate/economy models. We analyze the impact of substitution preferences, risk aversion (known probabilities), and specifically ambiguity aversion (unknown probabilities) on the social cost of carbon. Introducing ambiguity aversion leads to two offsetting effects on the social cost of carbon: a positive direct effect and a negative effect through discounting. Our numerical results show that for reasonable calibrations, the direct effect dominates and that ambiguity aversion gives substantially higher estimates of the SCC.

JEL codes: Q51, Q54, G12, G13
Keywords: Social Cost of Carbon, Ambiguity Aversion, Epstein-Zin preferences, Climate Change

*We thank Bob Pindyck, Nick Stern, Rick van der Ploeg, the late Martin Weitzman and seminar participants at the EEA meetings in 2020, DNB and the University of Amsterdam for helpful comments and pension administrator and wealth manager MN for financial support.

†University of Amsterdam, s.w.j.olijslagers@uva.nl

‡University of Amsterdam, Tinbergen Institute, CEPR, s.j.g.vanwijnbergen@uva.nl

1 Introduction

Climate change is one of the main risks the economy will face in the upcoming decades or possibly even centuries. But although climate scientists agree on the fact that climate change will most likely have dramatic negative consequences for the environment and economic growth, there is still much uncertainty surrounding the extent and timing of future damages induced by climate change (cf IPCC (2021) for a very recent assessment). Despite all the uncertainty about the timing and the exact structure and extent of the damages that climate change will cause, we do know that if they are to be avoided policies need to be implemented today. This should place the issue of how to discount future uncertain cost of climate change back towards today to allow comparison to the costs of today's policy interventions, at the center-stage of the climate change debate. And hence the subject of this paper, on climate change, risk, ambiguity aversion and Epstein-Zin preferences, and what it all implies for the Social Cost of Carbon (SCC).

Rather than arguing about specific numerical values for parameters such as time preference, we challenge the structure of preferences commonly assumed to derive the appropriate discounting procedures and discount rates.[1] Specifically, in this paper we model climate damages as disaster risk and assume that there is ambiguity about the arrival rate and size of future climate disasters. We show that implementing these extensions leads to estimates of the social cost of carbon that are substantially higher than have been derived so far using conventional approaches to time and risk discounting. Since so little is known about the exact distribution of uncertain climate shocks, we focus predominantly on Ambiguity Aversion under different assumptions about preferences, and its impact on the SCC.

The impact of climate change on the economy is most commonly modeled using combined economy/climate models called Integrated Assessment Models (IAMs). IAMs integrate the knowledge of different domains into one model. In the case of climate change, IAMs combine an economic model with a climate model. Three well-known IAMs are DICE (W. Nordhaus, 2014), PAGE (Hope, 2006) and FUND (Tol, 2002).[2] These models are, among others, used as policy tools for cost-benefit analyses. They provide a conceptual framework to better understand the complex problem of climate change by combining different fields and allowing for feedback effects between those fields.

But IAMs also have major drawbacks. To quote Pindyck (2017): "*IAM-based analyses of climate policy create a perception of knowledge and precision that is illusory ...*" His critique is that the models are (1) very sensitive to the choices of parameters and functional forms, especially the discount rate. Besides, we know very little about (2) climate sensitivity and (3) damage functions. Lastly, (4) IAMs don't incorporate tail risk. He recommends simplifying the problem by focusing on the

[1] For a very different (and strongly worded) view focusing on the social welfare aspects of the rate of time preference rather than on individual preferences, see Stern (2015) and Chichilnisky, Hammond, and Stern (2018) who look at a positive rate of time preference as discrimination between generations that happen to have been born at different moments in time.

[2] The references do not contain the most recent versions of the IAMs.

catastrophic outcomes of climate change, instead of modeling the underlying causes. In line with that view we focus on disaster risks by modeling them as tail risks (Poisson shocks). And our focus on ambiguity aversion naturally follows from his observation that we know very little about the precise stochastics of climate disasters.

The three main IAMs are deterministic, largely because stochastic models with many state variables are more difficult to solve than deterministic models. To nevertheless capture uncertainty, some authors perform an Initial Value Monte Carlo-like approach (IVMC) by analyzing several deterministic runs with different parameter values and then taking a weighted average of all runs (Dietz, 2011; W. D. Nordhaus, 2014). Such an analysis is useful if we are interested in the sensitivity of the models to different parameter values. However, it is conceptually different from explicitly using stochastic processes, since under this IVMC approach for each individual run all uncertainty is resolved at time 0. Crost and Traeger (2013) compare the IVMC approach to an explicitly stochastic model and find that the IVMC approach underestimates the impact of climate damages. And as we will discuss below, the timing of the resolution of uncertainty particularly matters a great deal under the structure of preferences we are analyzing.

We propose an analytically solvable IAM (Integrated Assessment Model) that addresses both the critiques of Pindyck (2017) and of Crost and Traeger (2013) on the use of deterministic IAMs. Since there is so little known about the damage functions, we investigate the impact of both attitudes towards well defined measurable risks and ambiguity aversion towards unmeasurable uncertainty on the willingness to pay for avoiding climate risk. Furthermore we model climate risk as disaster risk instead of assuming that temperature increases generate a certain amount of damage every year. The model is transparent due to the analytic solutions for the social cost of carbon. Where stochastic numerical IAMs commonly take hours or more to be solved, solving this model only requires numerical integration and is therefore solved within seconds, which makes it a useful tool for further analysis.

In the first part of the paper we provide analytical solutions. To make that possible we model the economy as a pure exchange economy with exogenous stochastic endowments. In this part of the paper we model emissions as an deterministic although time varying process, since linking stochastic emissions to the stochastics of output- and consumption processes precludes analytical solutions. We introduce more realistic stochastic emissions in the second part of the paper where we use numerical solution methods. We extend the general equilibrium Consumption-based Capital Asset Pricing Model (CCAPM), also known as Lucas-tree model, developed in Lucas Jr (1978) in several directions. In the literature, this model is widely used in conjunction with a lognormal distribution.[3] The diffusion component of the endowment captures fluctuations in consumption. But we take into account that the nature of climate risk is different from 'normal' economic risk as captured by a diffusion term. Climate disasters are events that occur rarely and take place abruptly (Goosse, 2015). To model this feature, we add a jump process to the endowment consumption stream to capture climate disaster risk.

[3] Although Lucas Jr (1978) doesn't assume a specific distribution for the endowment stream.

The intensity of the disasters is temperature-dependent. We model emissions, atmospheric carbon concentration and the temperature anomaly. The arrival rate of climate disasters is increasing in temperature. Furthermore we explicitly take into account that it is hard to estimate the probability that a disaster occurs and its expected impact by assuming that the agent does not know the exact probability distributions of the arrival rate of climate disasters and the size of the disasters: there is so called ambiguity about the characteristics of the jump risk component. And the agent is assumed to be averse to this ambiguity or Knightian uncertainty.

Finally we use the continuous time version of Epstein-Zin utility, which allows us to separate the intertemporal elasticity of substitution from the degree of risk aversion. In the widely used power utility specification risk aversion and elasticity of intertemporal substitution (EIS) are captured by one parameter, they are equal to each other's inverse. There is strong empirical evidence placing the relative degree of risk aversion in the range of 5 - 10 (Cochrane, 2009). Using such estimates in combination with power utility then results in implied estimates for the EIS much lower than direct empirical estimates suggest. But especially for long term problems such as climate change intertemporal choices play an important role and restricting parameters such as the EIS is a severe limitation. Epstein-Zin preferences make it possible to separate risk aversion and the elasticity of intertemporal substitution.

We can therefore disentangle risk aversion effects (known probabilities), ambiguity aversion effects (unknown probabilities) and substitution effects. The Epstein-Zin preferences also allow for the possibility that the agent has a preference for early resolution of risk, clearly of relevance in a discussion on climate risks. We show that the specification of the agent's preferences in combination with stochastic disaster risk has large effects on how much one is willing to pay to reduce climate risk.

The literature reports diverging results on the impact of ambiguity aversion on the SCC. Our comprehensive framework incorporating both Epstein-Zin preferences and Ambiguity Aversion proves its value added here: we show that assumptions on the structure of preferences have a major impact on estimates of the link between ambiguity aversion and the SCC made using these assumptions. We conclusively show that for empirically supported values of risk aversion and the EIS, Ambiguity Aversion has a substantial impact on the SCC.

In this paper we explicitly focus on the valuation of climate risk in the Business As Usual (BAU) scenario and do not analyse optimal abatement policies at this stage yet, optimal policy is integrated into the analysis in a companion paper (Olijslagers, van der Ploeg, and van Wijnbergen (2021)). The idea is that an analysis of the environmental costs of current policies (not current plans...) is useful in the climate policy debate. A commonly used measure for the cost of carbon emissions is the social cost of carbon (SCC), the long term discounted damage in dollar terms of emitting one ton of carbon today. The BAU scenario is also the default scenario to calculate the social cost of carbon in W. Nordhaus (2014). Note that the social cost of carbon in our model is not equal to the globally optimal Pigouvian carbon tax, since we do not consider abatement policy in this model. One of us compares the two in Olijslagers (2020). The social cost of carbon using a baseline scenario can be interpreted as the monetized welfare loss of emitting one additional unit of carbon today, given the current global carbon abatement policy scenario under the assumption that no measures will be taken in the future either. This seems to us an important first step to take for as long as effective international policies are not yet agreed upon and future agreement is not yet certain. In those circumstances the cost of doing nothing should surely be an important input in the debate.

Our base calibration yields a sizable social cost of carbon. Similar to the numerical IAMs, the SCC in our model is very sensitive to the choice of the input parameters. But in addition because we do have (mostly) analytical solutions, we can easily explore the implications of ambiguity aversion, preferences for early resolution of uncertainty and (related to that) a higher elasticity of intertemporal substitution (EIS). The preference structure we use (Epstein-Zin preferences) allows for variation of the EIS without corresponding variation in the degree of risk aversion; the two are inversely related under the more commonly used power utility assumption. In spite of incorporating all these generalizations we can still derive analytic expressions for the SCC, up to an integral, in our core model setup, making it transparent how ambiguity aversion and Epstein-Zin preferences influence the SCC. Our numerical example using best estimates of the various parameters indicates that introducing ambiguity aversion yields a SCC that is between 65% and 83% higher depending on the structure of climate risk. Moreover we highlight that the social cost of carbon is also sensitive to choices about time discounting, either via the pure rate of time preference, risk aversion or the elasticity of intertemporal substitution, and that all these parameters interact with the cost of ambiguity aversion in complicated ways. But the overall conclusion remains: insufficient attention to risk and ambiguity pricing leads to substantial underestimation of the SCC. These are non-trivial results because risk and ambiguity aversion also lead to higher risk premia. The impact of higher risk premia is more than offset however by the impact of risk aversion and ambiguity aversion on the certainty equivalence estimates that are subsequently discounted back to today. The net impact of higher aversion to risk and to ambiguity is to substantially raise the SCC, in particular for the realistic case of stochastic emissions which we analyse in the numerical solutions section..

In the first part of the paper we assumed that emissions are an independent and non-stochastic process. This is clearly an unrealistic assumption, but making it allows for analytical solutions. We introduce stochastic emissions in the section using numerical solution procedures and show that this does not materially change the results. And analytical solutions give considerably more insight and generality than can be obtained from numerical solutions, but assuming non-stochastic emissions obviously leaves out a part of reality that is important for the questions we ask. This shortcoming is remedied in the second part of Section 5.3, where emissions are modelled as an explicitly stochastic process correlated to output. There we show that the stochastic nature of emissions (and the correlation to output) adds additional sources of risk and leads to a slightly higher impact of risk aversion and ambiguity aversion on the Social Cost of Carbon than we already obtained in the first part of the paper.

The plan of the paper is as follows: After the introduction (Section 1) we discuss related literature in Section 2 and introduce the basic model in Section 3, focusing on the endowment process in Section 3.1, on modeling climate change and its economic impact in Section 3.2 and on preference structure in Section 3.3 where we introduce Epstein-Zin preferences. In Sections 3.4 we outline our approach to ambiguity Aversion and what that implies in the current model (Section 3.5). In Section 4 we present our analytical results on discount rates, the social cost of carbon and ambiguity aversion. Section 5 switches to the use of numerical solution methods; we first calibrate our model (Section 5.1), and use the calibrated version to illustrate our analytical results quantitatively, still assuming deterministic emissions (Section 5.2). In Section 5.3 we analyse the full model with stochastic emissions. Section 6 concludes.

2 Related literature

This paper is related to two strands in the literature. First, our methodology is related to consumption based asset pricing models with disaster risk and/or non-expected utility. And second and more important, the paper is related to research on the impact of climate change on the economy.

The model we develop is an extension of the Consumption based Capital Asset Pricing Model (CCAPM) by Lucas Jr (1978). Mehra and Prescott (1985) point out that for plausible parameter values, the CCAPM produces a way too low equity premium and correspondingly a too high risk-free rate. Jump risk or disaster risk has been proposed as a possible solution of these puzzles (Barro, 2006; Rietz, 1988). Extensions to the early disaster/jump risk models are the use of Epstein-Zin utility instead of power utility, and the introduction of time-varying disaster probabilities and multi-period (i.e. persistent) disasters (Barro, 2009; Tsai & Wachter, 2015; Wachter, 2013). We build on these extensions and take a similar approach, because climate change is widely thought to give rise to abrupt destructive changes in the Earth's environment (Goosse, 2015). We therefore define climate shocks as disasters whose occurrence has a small probability at any given moment of time but with possibly large negative effects on the economy once they do take place.

Ambiguity aversion, aversion of unmeasurable or Knightian uncertainty, is the second extension of the CCAPM we introduce to our climate model. This has been done before in the Finance literature: Liu, Pan, and Wang (2004) consider a general equilibrium model with rare disasters and ambiguity aversion in their analysis of option pricing and the well known failure of the Black-Scholes-Merton model to adequately reflect tail risks (vide the appearance of 'smirks' in implied volatility graphs). Their agent is only concerned about misspecification of the jump process and not of the diffusion terms, a logical choice that we follow, since the probability distribution of rare events is by their very nature (they do not occur regularly) much harder to estimate than the diffusion component.

Finally, since abrupt climate change is anticipated to take place far into the future, intertemporal choice plays an important role as well. Power utility is then an unsatisfactory framework since with that structure of preferences, risk aversion and elasticity of intertemporal substitution (EIS) cannot be varied independently, cf Epstein and Zin (1989) and a still growing literature that took off after their paper. We adopt the continuous time implementation of the Epstein-Zin framework introduced by Duffie and Epstein (1992a) and Duffie and Epstein (1992b). With Epstein-Zin preferences, the risk aversion parameter and the EIS are no longer restricted to be each other's inverse.

Furthermore, our paper is related to the literature on climate change economics, and more specifically the part that considers risk, ambiguity aversion and non-expected utility. These issues are not yet analyzed in the most well-known integrated assessment model, the DICE model (W. D. Nordhaus, 2017). This model is still deterministic and the representative agent is assumed to have power utility. Several papers have recently studied the impact of risk and more complex preference structures on the social cost of carbon. For instance Cai and Lontzek (2019), Hambel, Kraft, and Schwartz (2021) and Jensen and Traeger (2014) study integrated assessment models with Epstein-Zin preferences and different types of economic and climate risk. They show that Epstein-Zin preferences can have a substantial effect on the discount rate, for obvious reasons a very important parameter in climate models. Barro (2015) extends his disaster risk model with environmental disasters and focuses on discount rates and optimal environmental investment. He does not incorporate a climate model but simply assumes that the disaster probability is constant and that it can be reduced by environmental investment. Bansal, Kiku, and Ochoa (2016) propose a climate model based on the Long-Run-Risk (LRR) model of Bansal and Yaron (2004). In the LRR-model, the agent has Epstein-Zin preferences and consumption growth is subject to persistent shocks. Bansal et al. (2016) model climate disasters as a jump process that affects both consumption itself and the growth rate of consumption. They also show that their results are very sensitive to choices of the EIS. Karydas and Xepapadeas (2019) consider a dynamic asset pricing framework with both macroeconomic disasters and climate change related disasters and analyze the implications for portfolio allocation. We add to this part of the literature by including ambiguity aversion in our framework.

Most integrated assessment models are solved using numerical methods. The disadvantage of numerical solution approaches is that the choice of the input parameters in numerical solutions has a large influence on the results but not always in transparent ways. That is why analytical approaches such as adopted in the first part of our paper are useful. They can be used to show how exactly these parameters influence the outcomes. Of course we are not the first ones going for analytical solutions: Golosov, Hassler, Krusell, and Tsyvinski (2014) already obtain closed form solutions in a climate economy model. However, this required strict assumptions such as logarithmic utility and full depreciation of capital every decade. Bretscher and Vinogradova (2018) develop a stylized production-based model where the current carbon concentration directly enters the damage function and obtain closed form solutions for the optimal abatement policy. Van den Bremer and Van der Ploeg (2021) consider a stochastic production-based model with Epstein-Zin preferences, convex damages, uncertainty in state variables, correlated risks and skewed distributions to capture climate feedbacks. Since the model is too complex to obtain exact analytic solutions, they obtain closed form approximate solutions using perturbation methods. Lastly, Traeger (2021) extends the model of Golosov et al. (2014). The model allows for uncertainty in both the climate model and in climate damage impacts. Our contribution to this literature is that we obtain analytic results for a model with ambiguity about climate impacts.

There is a recent literature emerging on climate change and ambiguity aversion but this literature seems to come out on diverging views. Millner, Dietz, and Heal (2013) consider ambiguity about the climate sensitivity parameter within the DICE model and conclude that ambiguity aversion can lead to much higher optimal abatement policies. Barnett, Brock, and Hansen (2020) study ambiguity about both the climate sensitivity and damages within a climate economy model with a focus on the optimal carbon price. In their setup ambiguity aversion leads to a substantially higher optimal carbon price. On the other hand, Lemoine and Traeger (2016) model ambiguity around tipping points and claim it plays a minor role in the optimal carbon price.

Our analytic framework gives additional insights about the implication of ambiguity aversion in a climate model: we show that different outcomes can be traced back to particular assumptions on parameter values of the preference parameters. Our analytic results specifically allow for a decomposition of the effect of ambiguity aversion on the social cost of carbon. Introducing ambiguity aversion will lead to a higher social cost of carbon but also has an indirect effect on the discount rate. We are able to disentangle the direct and discounting effects that work in opposite directions. The size of both effects interacts with the other preference parameters: risk aversion and the EIS. These effects cannot be separated in a numerical application.

3 The Model

We extend a standard endowment economy by assuming that the stochastic endowment stream is subject to climate disasters, where the probability of a climate disaster depends on the temperature level. An endowment economy is in our view a suitable starting point given our focus on the social cost of carbon and the way it depends on uncertainty and ambiguity for given policies. In particular we analyse the SCC in Nordhaus' Business As Usual scenario. In a companion paper (Olijslagers et al., 2021) we endogenize abatement policy and analyse the price of carbon under optimal abatement policies and different objective functions.

3.1 The economy

The aggregate endowment process follows a geometric Brownian motion with an additional jump component that represents climate disasters:

$$dC_t = \mu C_t dt + \sigma C_t dZ_t + J_t C_{t-} dN_t.^4 \tag{1}$$

In equilibrium, aggregate consumption must equal the aggregate endowment and therefore we also refer to the process as the aggregate consumption process. The growth rate μ and the volatility σ are constant. Z_t is a standard Brownian motion

[4] C_{t-} denotes aggregate endowment just before a jump ($C_{t-} = \lim_{h \downarrow 0} C_{t-h}$).

that captures 'normal' uncertainty, i.e. non-climate uncertainty. N_t is a Poisson process which represents climate disasters. The arrival rate of a climate disaster equals λ_t, which we assume to be a function the temperature level T_t.

When a climate disaster strikes at time t, the size of the disaster is controlled by the random variable J_t. The distribution of the size of disasters is assumed to be the same for any t. We assume that J_t has the density $f(x) = \eta(1+x)^{\eta-1}$ where $-1 < x < 0$. J_t represents the percentage loss of aggregate consumption after a disaster. The expected disaster size then equals $E[J_t] = \frac{-1}{\eta+1}$ and the moments $E\Big[(1+J_t)^n\Big] = \frac{\eta}{\eta+n}$ can be easily calculated. In line with the subject of climate disasters, jumps can only be negative.

3.2 The climate model

The arrival rate of disasters is assumed to be temperature dependent. We assume that damages are linearly increasing in temperature: $\lambda_t = \lambda_T T_t$. However, all our derivations remain valid for non-linear specifications of the arrival rate. We discuss this assumption in the calibration section.

In the first part of the paper we make a number of simplifying assumptions to allow for analytic solution of the model. The main solvability requirement is that the state variables of the climate submodel are deterministic, and this in particular affects the way we model emissions. Carbon emissions are the product of the carbon intensity of aggregate output and aggregate output itself. Simply introducing emissions like this in the model would make emissions stochastic, which in turn precludes analytical solutions. So our main simplification in the analytics part of the paper is the assumption that aggregate emissions are driven by an independent deterministic process, which is an unavoidable simplification if one is to obtain analytical solutions. In section 5.3 we use numerical methods and introduce stochastic emissions correlated with output. Making emissions stochastic clearly adds realism and enriches the results, but the numerical models do show the benefits of the additional insights one can obtain from the analytical results obtained earlier.

So we assume for the first part of the paper that emissions E_t are exogenous. E_t is growing at a non-stochastic rate $g_{E,t}$. The growth rate itself moves gradually towards the long-run equilibrium $g_{E,\infty}$ at a rate δ_E. By assuming a high initial growth rate but a negative long run rate ($g_{E,\infty} < 0$), we have growing emissions today; but the growth rate starts declining immediately and eventually turns negative because of $g_{E,\infty} < 0$, so emissions will go to zero eventually. This is a plausible assumption since there is a point where the stock of fossil fuels will be depleted. All this leads to the following process for emissions:

$$\begin{aligned} dE_t &= g_{E,t} E_t dt, \\ dg_{E,t} &= \delta_E (g_{E,\infty} - g_{E,t}) dt. \end{aligned} \tag{2}$$

We calibrate this process to match the baseline scenario in W. D. Nordhaus (2017).

We use the climate model (carbon cycle and temperature model) discussed in Mattauch et al. (2018), which they call the IPCC AR5 impulse-response model. This model is in line with recent insights from the climate literature and is also used in IPCC (2013). Specifically, this climate model incorporates the fact that thermal inertia play a smaller role than commonly assumed in the climate modules in economic models. Climate modules commonly used in economic models tend to overstate the time it takes for the earth to warm in response to carbon emissions (cf Dietz, van der Ploeg, Rezai, and Venmans (2021)).

Define by M_t the carbon concentration with respect to pre-industrial emissions M_{pre}. In our model, M_t is the sum of four artificial carbon boxes: $M_t = \sum_{i=0}^{3} M_{i,t}$. This specification can capture that the decay of carbon has multiple time scales and that a fraction of emissions will stay in the atmosphere forever. The dynamics of carbon box i are given by:

$$dM_{i,t} = \nu_i \Big(E_t - \delta_{M,i} M_{i,t} \Big) dt. \tag{3}$$

ν_i is the fraction of emissions that ends up in carbon box i, which implies that $\sum_{i=0}^{3} \nu_i = 1$. $\delta_{M,i}$ controls the decay rate of carbon in box i. We assume that all carbon that ends up in box 0 will permanently stay in the atmosphere, such that $\delta_{M,0} = 0$. The other three boxes have a positive decay rate: $\delta_{M,i} > 0, i = \{1, 2, 3\}$.

The next step is to model the impact of carbon concentration on temperature. This requires modeling what is called radiative forcing: the difference between energy absorbed by the earth from sunlight and the energy that is radiated back to space. A higher atmospheric carbon concentration strengthens the greenhouse effect and therefore leads to higher radiative forcing. The relation between atmospheric carbon concentration and radiative forcing is logarithmic:

$$F_{M,t} = \alpha \frac{\upsilon}{\log(2)} \log \Big(\frac{M_t + M_{pre}}{M_{pre}} \Big). \tag{4}$$

α equals the climate sensitivity: the long-run change in temperature due to a doubling of the carbon concentration compared to the pre-industrial level. υ is a parameter that is also part of the temperature module and this parameter will be discussed later.

Finally we also include non-carbon related (exogenous) forcing $F_{E,t}$, which follows:

$$dF_{E,t} = \delta_F (F_{E,\infty} - F_{E,t}) dt. \tag{5}$$

Total radiative forcing is the sum of carbon-related radiative forcing and exogenous forcing: $F_t = F_{M,t} + F_{E,t}$.

The final step moves from F_t to the actual surface temperature T_t. T_t is the difference between the actual temperature compared to the pre-industrial temperature level. The change in surface temperature is a delayed response to radiative forcing. Call the heat capacity of the surface and the upper layers of the ocean τ while τ_{oc} equals the heat capacity of the deeper layers of the ocean. The parameter κ captures the speed of temperature transfer between the upper layers and the deep layers of the ocean. The dynamics of temperature are then given by:

$$\begin{aligned} dT_t &= \frac{1}{\tau} \Big(F_t - \upsilon T_t - \kappa (T_t - T_t^{oc}) \Big) dt, \\ dT_t^{oc} &= \frac{\kappa}{\tau_{oc}} (T_t - T_t^{oc}) dt. \end{aligned} \tag{6}$$

From this equation, one can derive a long run equilibrium temperature for a given level of radiative forcing F_t:

$$T_t^{eq} = \frac{F_t}{\upsilon} \tag{7}$$

The parameter υ controls the equilibrium temperature response to a given level of forcing. Note that equation (4) tells us that when $M_t = 2M_{pre}$, we get that $F_t = \alpha\upsilon + F_{E,t}$ and $T_t^{eq} = \alpha + \frac{F_{E,t}}{\upsilon}$. Therefore the parameter α can indeed be interpreted as the equilibrium temperature response to doubling of the carbon concentration.

Using equation(7), we can rewrite the first line of equation (6) as:

$$dT_t = \frac{1}{\tau}\Big(\upsilon(T_t^{eq} - T_t) - \kappa(T_t - T_t^{oc})\Big). \tag{8}$$

Written this way the equation is more intuitive, since it captures the fact that the temperature moves towards its equilibrium level at a rate proportional to $T_t^{eq} - T_t$. The second part shows that the oceans are delaying this convergence. It takes time for T_t^{oc} to adjust towards T_t and this will also delay the convergence of T_t towards the equilibrium level T_t^{eq}. As specified earlier, the arrival rate of climate disasters is a linear function of temperature T_t.

3.3 Preference specification

The representative agent maximizes utility of consumption over an infinite planning horizon. Because of the different roles played by intertemporal substitution and risk aversion in determining risk premia, the safe rate of interest, and therefore also the social cost of carbon, we use Epstein-Zin (EZ) preferences (Epstein & Zin, 1989); EZ preferences allow us to vary the elasticity of intertemporal substitution (EIS) and the coefficient of relative risk aversion independently. This is additionally important in our framework since both the elasticity of intertemporal substitution (through the discount rate) and risk aversion interact with ambiguity aversion. We use the continuous time version of of Epstein-Zin utility, a special case of stochastic differential utility introduced by (Duffie & Epstein, 1992b).

The agent's utility or value function is:

$$\begin{aligned} &V_t = E_t\Big[\int_t^\infty f(C_s, V_s)ds\Big] \\ &\text{where} \\ &f(C, V) = \frac{\beta}{1 - 1/\epsilon}\frac{C^{1-1/\epsilon} - \big((1-\gamma)V\big)^{\frac{1}{\zeta}}}{\big((1-\gamma)V\big)^{\frac{1}{\zeta}-1}} \quad \text{for } \epsilon \neq 1 \\ &\text{with} \;\; \zeta = \frac{1-\gamma}{1-1/\epsilon}. \end{aligned} \tag{9}$$

γ denotes risk-aversion, ϵ is the elasticity of intertemporal substitution and β equals the time preference parameter. We will focus on the more general case where $\epsilon \neq 1$.

For the case $\epsilon = 1$ one can take the limit $\epsilon \to 1$ or follow the same derivation but with $f(C, V) = \beta(1-\gamma)V\Big(\log C - \frac{1}{1-\gamma}\log\big((1-\gamma)V\big)\Big)$. Finally if $\gamma = \frac{1}{\epsilon}$, the utility specification reduces to standard power utility.

3.4 Ambiguity

There is much uncertainty regarding the arrival rate and magnitude of climate disasters. Pindyck (2017) already stresses that we know very little about the damage functions. And where consumption growth and volatility can be estimated accurately from historical data, the estimation of the climate disaster parameters will be much harder since climate disasters do not happen that often. It is fair to state that we simply do not know the exact distribution of climate damages. We should therefore account for the possibility that the 'best estimate' model is not the true model: there is ambiguity. We assume that the representative agent is ambiguity averse.

It is important to stress the difference between risk and ambiguity. When we are talking about risk, an agent knows the probabilities and possible outcomes of all events. When the agent has to deal with ambiguity, the probabilities attached to particular events are unknown. The distinction between risk and ambiguity is already extensively discussed in Knight (1921), which is why ambiguity is often referred to as Knightian uncertainty. Ellsberg (1961) shows using the Ellsberg Paradox that people are ambiguity averse, i.e. they prefer known probabilities over unknown probabilities.

We use the *recursive multiple priors utility* developed in continuous time by Chen and Epstein (2002) to model ambiguity aversion. This method selects a set of models that are relatively similar. The size of the set depends on the degree of ambiguity aversion. Given this set of reasonable models, the worst case model is selected. The decision maker thus makes a robust decision given the set of reasonable models.

To apply the approach of Chen and Epstein (2002) to modeling ambiguity aversion we begin by defining the 'best estimate' model or reference model as the agent's most reliable model, with probability measure $\mathbb{P}$. But the agent takes into account that his reference model may not be the true model and specifies a set of models $\mathcal{P}^\theta$ that he considers possible. The alternative models have measure $\mathbb{Q}^{a,b}$; the jump arrival rate becomes $\lambda_t^{\mathbb{Q}^{a,b}} = a\lambda_t$ and the jump size parameter becomes $\eta_t^{\mathbb{Q}^{a,b}} = b\eta$. Remember that the expected jump size equals $\frac{-1}{\eta+1}$, i.e. a low b leads to a more negative jump size. Given the set of models $\mathcal{P}^\theta$, Chen and Epstein (2002) then assume that the agent optimizes assuming the worst case, in line with the axiomatic Minimax approach advocated by (Gilboa & Schmeidler, 1989).

An alternative way to model ambiguity is to use the *smooth ambiguity* model (Klibanoff, Marinacci, & Mukerji, 2005). Assume again that the agent does not know the true values of λ and η. In this approach the agent first constructs a prior probability distribution that reflects his beliefs on λ and η. To incorporate ambiguity aversion, he then transforms this distribution to put more weight on the events that give him low utility and less weight on the events that give high utility. This transformation works in a similar way as risk aversion in a standard utility function but adds another layer to your utility specification. This may be a matter of taste, but we think that the assumption of probabilities attached to the different priors is in fact at variance with the basic assumption that ambiguity is about unmeasurable processes, i.e. we cannot map events to probability densities, or in this case priors to model probabilities. Additionally the recursive multiple priors approach is simpler and leads to more tractable results. We therefore chose to use the recursive multiple priors approach.

All models with a distance smaller than θ are in the set of admissible models, so the size of the set of models depends on the ambiguity aversion parameter θ; and θ can be interpreted as a measure of the extent of ambiguity. We measure distance between the reference model $\mathbb{P}$ and an alternative model $\mathbb{Q}^{a,b}$ using the concept of *relative entropy*, a common metric for the distance between two probability measures (see for example Hansen and Sargent (2008)). Relative entropy thus gives information about how similar two probability measures are. To obtain our distance measure, we scale relative entropy by the arrival rate λ_t.[5] Without this scaling, the optimal a^* and b^* would be time-varying. This would imply that the decision maker is continuously updating a^* and b^*. A constant a^* and b^* are both more intuitive and more tractable.

The distance between the reference and alternative model depends on the parameters a and b and can therefore be written as $d(a, b)$. The distance measure satisfies $d(a, b) \geq 0 \ \forall (a, b)$ and $d(1, 1) = 0$: the distance of the reference model to itself is by definition equal to 0. If θ is large, the agent is very ambiguity averse and thus considers a large set of models. The preferences of the agent then become:

$$\begin{aligned} & V_t = \min_{\mathbb{Q}^{a,b} \in \mathcal{P}^\theta} V_t^{\mathbb{Q}^{a,b}} \\ & \text{where} \ \ V_t^{\mathbb{Q}^{a,b}} = E_t^{\mathbb{Q}^{a,b}} \Big[\int_t^\infty f(C_s, V_s^{\mathbb{Q}^{a,b}}) ds \Big] \\ & \text{and} \ \ \mathcal{P}^\theta = \{ \mathbb{Q}^{a,b} : d(a,b) \leq \theta \ \forall t \}. \end{aligned} \tag{10}$$

Here $V_t^{\mathbb{Q}^{a,b}}$ is the value function assuming $\mathbb{Q}^{a,b}$ is the true probability measure. $\theta = 0$ implies that $\mathcal{P}^\theta = \{\mathbb{P}\}$ and the agent only considers one measure, namely the reference measure. Thus there is no ambiguity aversion when $\theta = 0$. Where the risk aversion parameter γ can be seen a parameter that is relevant for any risky bet, the parameter θ captures intrinsic ambiguity aversion (one person might be more ambiguity averse than another), but it is also source dependent. If there is a lot of information and data available about a process, θ will be smaller and the set of admissible models will be smaller compared to a process about which not much is known. But at the same time θ also captures aversion to ambiguity similar to risk aversion.

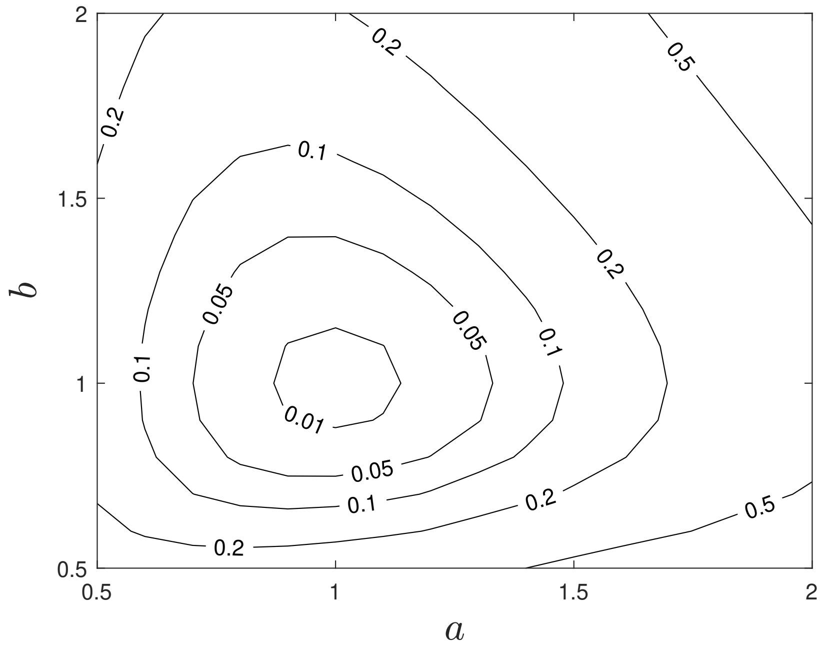
In appendix A we derive that the distance measure equals:

$$d(a, b) = (1 - a) + a\Big(\log(ab) + \frac{1}{b} - 1 \Big). \tag{11}$$

It is easy to verify that d(1,1) = 0, the distance between the reference distribution and itself is zero. When one or both of the two variables a and b deviate from the reference

[5] Liu et al. (2004) and Maenhout (2004) also use a normalisation factor to scale their distance measure in order to get tractable results.

Figure 1: Distance measure for different values of a and b.



model, $d(a, b)$ increases. Every contour in figure 1 gives a set of combinations (a, b) that yields the same distance. If for example $\theta = 0.1$, then all (a, b) combinations within that contour line are included in the set of admissible models. The worst case probability measure will be the probability measure for which either a is large (high arrival rate) and/or b is small, since the expectation of the jump size under the alternative measure is inversely related to b: $E^{\mathbb{Q}^{a,b}}[J_t] = \frac{-1}{b\eta+1}$.

From the current setup, it is hard to argue what a reasonable value for ambiguity aversion θ would be. In order to give more guidance about reasonable values for θ, we use the concept of *detection error probabilities* introduced by Anderson, Hansen, and Sargent (2003).[6] Consider the following thought experiment. Assume that the representative agent would be able to observe the process of consumption over the next N years, and after observing the process the agent has to choose which of the two models (the reference model or the worst-case model) is most likely. There are two types of errors in this case. The agent could choose the reference model while the process was actually generated by the worst-case model and he could also make the opposite error. The detection error probability is defined as the average of the probability of the two errors. Appendix B describes how the detection error probability is calculated.

The detection error probability depends on N, since when the agent observes the process for a longer period, the probability of a mistake will be smaller. The detection error probability also depends on the ambiguity aversion parameter: when θ is small, the reference and worst-case model are similar to each other and the probability of a

[6]See for example Maenhout (2006) for another application of detection error probabilities.

mistake is large. On the other hand, when the agent is extremely ambiguity averse (or there is a lot of ambiguity) the reference and worst-case models are very different and the detection error probability becomes small. The representative agent wants to make the set of models sufficiently large to make a robust decision, but on the other hand does not want to take into account implausible models. The detection error probability gives guidance about whether the set of admissible models is too small or too large. Since the detection error also depends on the other parameters of the model, we come back to the issue of calibrating the ambiguity aversion parameter in the calibration section.

3.5 Optimal a and b

As discussed before, the agent has the following utility function: $V_t = \min_{\mathbb{Q}^{a,b} \in \mathcal{P}^\theta} V_t^{\mathbb{Q}^{a,b}}$, where $\mathcal{P}^\theta$ is the set of all probability measures that satisfy the distance constraint. Since every probability measure $\mathbb{Q}^{a,b}$ is uniquely defined by the parameters a and b, minimizing over $\mathbb{Q}^{a,b}$ is equivalent to minimizing over the parameters a and b. In appendix C we show that this minimization problem boils down to minimizing the following expression:

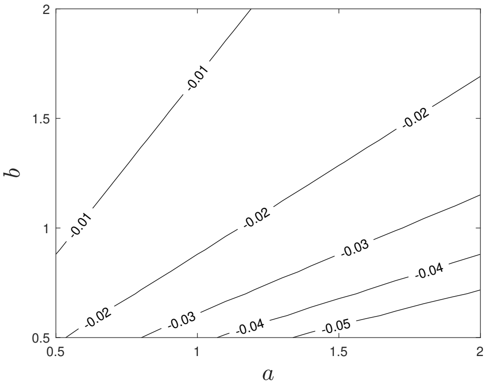
$$\min_{(a,b)} a\lambda_t \frac{-1}{b\eta + 1 - \gamma} \quad s.t. \quad d(a,b) \leq \theta. \tag{12}$$

We start with discussing this minimization problem assuming that the agent would be risk-neutral ($\gamma = 0$) but ambiguity averse ($\theta > 0$). At every time unit, the expected loss equals the probability of a disaster times the expected disaster size. Assuming that $\mathbb{Q}^{a,b}$ is the true measure, the arrival rate becomes $a\lambda_t$ and the expected disaster size equals $\frac{-1}{b\eta+1}$. The expected loss of a climate disaster therefore equals $a\lambda_t \frac{-1}{b\eta+1}$. The agent then chooses the combination of a and b that gives the most negative expected loss while still satisfying the distance restriction $d(a,b) \leq \theta$. A higher θ allows a larger range of values for a and b that satisfy the distance constraint.

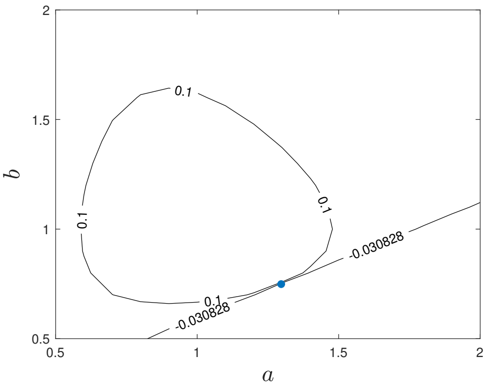
In our specification the agent is both risk averse and ambiguity averse. Instead of minimizing the expected loss, the agent minimizes the certainty equivalent of a climate disaster: $a\lambda_t \frac{-1}{b\eta+1-\gamma}$. The certainty equivalent is more negative than the expected loss since it contains a correction for risk aversion. The optimal parameters a^* and b^* are thus a function of ambiguity aversion θ, the jump size parameter η but also of risk aversion γ.

Figure 2 illustrates the optimization problem. Given an ambiguity budget θ, one can determine the feasible set of (a, b). Figure 1 shows the feasible sets for several budgets. A contour plot of the objective function for several (a, b) combinations is given in subfigure 2a. Clearly combinations in the bottom right corner (high a, low b) give the lowest value of the objective function. The optimization will thus lead to $a^* > 1$ and $b^* < 1$ since b^* and the disaster size are inversely related. The goal is to minimize this function, given the distance constraint. Subfigure 2b shows how the optimal combination (a^*, b^*) is determined. The point where objective function touches the feasible region is the optimal solution. From now on we use the following notation for the optimal arrival rate and jump parameter: $\lambda_t^* = a^*\lambda_t$ and $\eta^* = b^*\eta$.

Figure 2: Selection of the optimal a and b.



(a) Contour plot of the objective function of the constrained minimization problem for different values of a and b.



(b) Illustration of selection of optimal (a, b). The oval area shows all admissible values for a and b that are within the ambiguity budget of 0.1. The straight line is the objective function.

4 Discounting and the Social Cost of Carbon: Analytical Solutions

We are now ready to address the key questions raised in the introduction, how to discount future carbon damages and what that implies for the Social Cost of Carbon (SCC). We focus first on the appropriate discount rates.

4.1 On Discounting

Consider first the risk-free rate and the risk premium; we then derive the growth-adjusted consumption discount rate, the rate at which future consumption streams (or their decline) need to be discounted towards today, which is used to discount future damages when calculating the SCC.

In appendix D we derive the expressions for the interest rate, the risk premium and the consumption rate of interest from no arbitrage conditions for the valuation of respectively a safe asset B_t, aggregate wealth and a synthetic asset paying out aggregate consumption at a specified future time. We label latter the CDR_t; it equals the return on wealth $r + rp$ minus a correction for the growth in consumption.

Consider first the expression for the safe rate of interest, derived from the no arbitrage condition of a safe asset B_t (cf appendix D for the derivation details):

$$\begin{aligned} r_t = \beta + \frac{\mu}{\epsilon} - \Big(1 + \frac{1}{\epsilon}\Big)\frac{\gamma}{2}\sigma^2 - \Big(\gamma - \frac{1}{\epsilon}\Big)a^*\lambda_t \frac{-1}{b^*\eta + 1 - \gamma} \\ - a^*\lambda_t\Big(\frac{b^*\eta}{b^*\eta - \gamma} - 1\Big). \end{aligned} \tag{13}$$

We return to this expression below in the discussion. The relevant risk premium is the excess return on a claim on consumption, or, more precisely, a stock S_t paying out continuous dividends C_t. The value of the stock can also be interpreted as aggregate wealth, since total wealth of the representative agent is equal to the total claim on future consumption. Requiring once again the familiar no arbitrage condition gives the expression for the risk premium (again cf appendix D):

$$rp_t = \gamma\sigma^2 + a^*\lambda_t\Big(\frac{-1}{b^*\eta+1} - \frac{b^*\eta}{b^*\eta+1-\gamma} + \frac{b^*\eta}{b^*\eta-\gamma}\Big). \tag{14}$$

Without climate risk (the Poisson terms), the risk premium boils down to the well known expression: $\gamma\sigma^2$.

Finally we use the results for the safe rate of interest and the risk premium in the derivation of the expression for the growth-adjusted Consumption Discount Rate CDR_t. The consumption discount rate CDR_t is the relevant discount rate for discounting climate damages when calculating the social cost of carbon. In appendix D we derive the expression for this discount rate using the results we obtained so far for r_t and rp_t. The no-arbitrage condition is applied to a synthetic asset H_t paying out aggregate consumption at time $s > t$:

$$\begin{aligned} CDR_t &= \underbrace{r_t}_{I} + \underbrace{rp_t}_{II} - \underbrace{\Big(\mu + a^*\lambda_t\frac{-1}{b^*\eta+1}\Big)}_{III} \\ &= \beta + (1/\epsilon - 1)\Big(\mu - \frac{\gamma}{2}\sigma^2 + a^*\lambda_t\frac{-1}{b^*\eta+1-\gamma}\Big). \end{aligned} \tag{15}$$

CDR_t consists of three terms, labeled I, II and III. Part I is the risk-free rate. But future economic growth is uncertain, so we need to add a risk premium since damages are a fraction of the economy and thus have an impact on consumption: part II. Lastly, the discount rate should be corrected for the growth of the aggregate consumption process (part III). Future damages are larger because the future economy is larger, which is why we need to correct the discount rate for future growth. On average, consumption grows at a rate $\mu + a^*\lambda_t\frac{-1}{b^*\eta+1} < \mu$: the average growth rate is smaller than μ since climate disasters are expected to have a negative impact on consumption.

In the simplest case, without any risk at all, the risk premium is zero and the interest rate then reduces to the well-known Ramsey rule (Ramsey, 1928):

$$(\sigma, \lambda_T) = (0,0) \;\; => r_t = \beta + \frac{\mu}{\epsilon}, \tag{15a}$$

which implies a growth corrected discount rate $r_{n,t}$ for the case of $(\sigma, \lambda_T) = (0,0)$ equal to:

$$r_{n,t} = \beta + (1/\epsilon - 1)\mu. \tag{15b}$$

Clearly a higher value for ϵ implies a lower growth corrected discount rate: a higher willingness to substitute over time implies less discounting of the future. Adding diffusion risk ($\sigma > 0, \lambda_T = 0$) leads to well known results: this will both affect the safe interest rate, which falls due to a flight to safety effect, and the risk premium, which now becomes $\gamma\sigma^2$:

$$(\sigma > 0, \lambda_T = 0) => r_t = \beta + \frac{\mu}{\epsilon} - (1 + 1/\epsilon)\frac{\gamma}{2}\sigma^2, \tag{15c}$$

$$rp_t = \gamma\sigma^2. \tag{15d}$$

Adding the risk premium to the risk-free rate and again correcting for the growth rate gives the growth-adjusted discount rate, still assuming $\sigma > 0, \lambda_T = 0$:

$$r_{n,t} = \beta + (1/\epsilon - 1)\Big(\mu - \frac{\gamma}{2}\sigma^2\Big). \tag{15e}$$

So the impact on the safe rate and on the risk premium are in opposite directions, as is well known from the literature. For $\epsilon = 1$ the two effects cancel out, for $\epsilon > 1$ the risk premium impact dominates and the overall discount rate increases with risk. For $\epsilon < 1$ the opposite result obtains and discount rates will actually go down with higher risk as the flight to safety effect dominates the impact on the risk premium. While ϵ determines the relative importance of the interest rate and risk premium effects, risk aversion γ determines their magnitude. A high degree of risk aversion amplifies the effect of risk on the discount rate. Of course when the agent is risk neutral ($\gamma = 0$), risk has no effect on the discount rate.

We now introduce climate uncertainty in addition to diffusion risk, and ambiguity aversion, the main topic of this paper. Adding climate disaster risk to diffusion risk implies: $\sigma > 0$ and $\lambda_T > 0$. To set a benchmark we first analyse the case where there is no ambiguity aversion ($\theta = 0$). This corresponds to $a^* = 1, b^* = 1$, the optimal and the reference case actually coincide when $\theta = 0$. Equation (15) then shows that adding climate disaster risk has an effect on both the interest rate and the risk premium very much like changes in σ have. The climate risk term is premultiplied by $(1/\epsilon - 1)$ in equation (15), so when $\epsilon < 1$ the interest rate effect dominates and adding disasters leads to a lower discount rate. But when $\epsilon > 1$, the risk premium effect dominates and adding climate disasters actually leads to higher discount rates. Finally when $\epsilon = 1$, the two effects cancel.

In our no-ambiguity-aversion benchmark case $a^* = 1, b^* = 1$, the climate related term in equation (15) then becomes:

$$\lambda_t \frac{-1}{\eta + 1 - \gamma}. \tag{16}$$

$\frac{-1}{\eta+1-\gamma}$ equals the certainty equivalent of the climate shock. When $\gamma = 0$, the certainty equivalent is equal to the expected value $E_t[J_t] = \frac{-1}{\eta+1}$. The term scales with the arrival rate λ_t: more frequent disasters have a larger effect on discount rates. Finally a higher γ leads to a smaller certainty equivalent (i.e. a larger negative shock), since η is substantially larger than 1.

Now introduce ambiguity aversion. The climate term in equation (15) now equals:

$$a^*\lambda_t \frac{-1}{b^*\eta + 1 - \gamma}. \tag{17}$$

Including ambiguity aversion leads to a larger worst case arrival rate: $a^* > 1 => a^*\lambda_t > \lambda_t$ so one can see from equation (17) that ambiguity aversion leads to a larger worst case arrival rate. Thus ambiguity aversion amplifies the impact that the arrival rate of climate disasters has on discounting. Also, we can see from (17) that ambiguity aversion implies a more negative certainty equivalent term since $b^* < 1$; so once again we find that ambiguity aversion leads to a larger impact of climate risk on discounting. Therefore we can unambiguously conclude that there is AA amplification: ambiguity aversion amplifies the effect of climate risk on discounting, both through its impact on the worst case arrival state and on the worst case certainty equivalent conditional on arrival.

Whether AA amplification leads to a higher or lower discount rate depends on the value of ϵ, much like in the earlier discussion on the impact of (climate) risk on interest rates in the absence of ambiguity.

$\epsilon = 1$: the impact of AA amplification on the safe rate and on the risk premium cancel each other out and the discount rate simply becomes β irrespective of climate risk (or for that matter any other risk).

$\epsilon < 1$: the flight to safety effect of magnification dominates the impact of a higher risk premium, so AA amplification actually leads to a lower discount rate.

$\epsilon > 1$: we get the presumably more intuitive outcome, with $\epsilon > 1$ the risk premium effect dominates and AA magnification actually leads to a higher discount rate than obtained without AA magnification.

4.2 The social cost of carbon

With the machinery developed sofar and using the value function from equation (9) we can take the next step and calculate the Social Cost of Carbon (SCC). We define the SCC as the marginal cost in terms of reduced welfare of increasing carbon emissions by one ton carbon scaled by the marginal welfare effect of one additional unit of consumption. This gives us the social cost of carbon in terms of the price of time t consumption units terms (conventionally referred to as 'in dollar terms'). In appendix E we derive the following expression for the SCC based on this definition:

$$SCC_t = C_t \int_0^\infty \underbrace{\exp\left\{-\int_t^{t+u} CDR_s ds\right\}}_{I} \underbrace{\int_t^{t+u} a^*\lambda_T \frac{\partial T_s}{\partial M_t} ds}_{II} \underbrace{\frac{1}{b^*\eta + 1 - \gamma}}_{III} du \tag{18}$$

Equation (18) shows first of all that the social cost of carbon is proportional to C_t, the aggregate consumption level: when the current aggregate consumption level C_t doubles, the SCC doubles as well. For a given consumption level, the SCC depends on three terms, labeled I, II and III respectively in equation 18. The social cost of carbon. the marginal welfare loss due to emitting an additional unit of carbon today, is the discounted sum of all current and future damages done by emitting one ton of carbon today. The outer integral indicates that all future marginal damages are included in the SCC. Future damages are discounted with the (cumulative) consumption discount rate (term I). Term II is the change in the probability of a disaster between time t and time $t+u$ due to an additional unit of emissions today. This change in probability is a function of the derivative of future temperature levels with respect to current carbon emissions (which marginally change current carbon concentration M_t). Term III captures the damages when a disaster actually takes place. It can be interpreted as a certainty equivalent: the expected value is adjusted for risk and ambiguity preferences.

Consider first the impact of risk aversion as measured by γ. Term III is clearly increasing in risk aversion. But risk aversion also has an effect on the discount rate CDR_t. As discussed before, increasing risk aversion increases the discount rate when $\epsilon > 1$. So when $\epsilon > 1$ the discounting effect works in opposite direction of the effect on the certainty equivalent: for $\epsilon > 1$ the impact of γ on the SCC is therefore ambiguous in general and will depend on the specific parameter values chosen (cf the numerical analysis in Section 5).

Consider next the impact of ϵ. The elasticity of intertemporal substitution ϵ only plays a role in the discount rate. When ϵ increases, the willingness to substitute over time increases which leads to lower discount rates. So a higher ϵ unambiguously leads to a higher SCC.

The ambiguity aversion parameter θ does not directly show up in the formula for the SCC, but its effect works through the choice of a^* and b^*. When ambiguity aversion is present, i.e. $\theta > 0$, $a^* > 1$ (higher worst-case arrival rate) and $b^* < 1$ (more negative worst-case jump size). With $\theta > 0$, the increase in the probability of a disaster happening (term II) is larger because the worst case arrival rate of disasters $a^*\lambda_t$ is higher. And term III in expression 18, the certainty equivalent damage term conditional on a disaster happening, is also higher. So through these two channels ambiguity aversion leads to a higher social cost of carbon.

But ambiguity aversion also affects discount rates and the sign again depends on the elasticity of intertemporal subtitution ϵ. When $\epsilon < 1$, ambiguity aversion additionally leads to a lower discount rate and thus an even higher SCC. When $\epsilon = 1$, the discount rate is simply β and ambiguity has no effect on the discount rate so in that case the SCC increases with ambiguity also. Lastly, when $\epsilon > 1$, increasing θ leads to higher discount rates. Therefore increasing ambiguity aversion then has two offsetting effects in this case and the net sign of the impact of θ on the SCC is in principle ambiguous. Since ambiguity aversion always leads to a higher SCC when $\epsilon \leq 1$, we only focus on the case of $\epsilon > 1$ in the Numerical Solutions Section 5 below.

Summarizing, when considering the effect of ambiguity aversion on the social cost of carbon we can identify two effects. First, including ambiguity aversion leads to a higher arrival rate and a larger certainty equivalent of expected damages conditional on arrival, which unambiguously pushes the social cost of carbon up. We call this effect the *direct* effect of ambiguity aversion. Second, there is a more *indirect* general equilibrium effect through the impact of ambiguity aversion on discount rates. The discount rate that should be used to discount future climate disasters is the consumption discount rate. When $\epsilon \leq 1$, the discount rate goes down as ambiguity aversion increases but when $\epsilon > 1$ the elasticity of substitution is larger than 1, ambiguity aversion leads to a higher consumption discount rate. This is an intuitive result: if the representative agent is very ambiguity averse about climate disasters, he would rather like to consume today than to postpone consumption since the future consumption level is uncertain. Ambiguity aversion therefore increases the consumption discount rate when $\epsilon > 1$. When $\epsilon > 1$ it is ultimately a numerical issue which of the two effects dominates. We will highlight both effects separately in the numerical section and show that for our calibration the first effect dominates. Thus in our numerical analysis more ambiguity aversion leads to a higher SCC for all values of ϵ.

5 Climate change and the social cost of carbon: numerical results

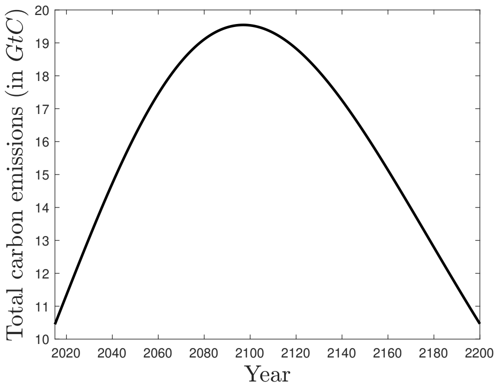
We first discuss our calibration choices (section 5.1). We then show numerical results for the model we analyzed analytically in the previous section, i.e. still the model version with non-stochastic emissions (Section 5.2). Finally in section 5.3 we introduce and analyse the full model where, more realistically, we assume emissions are also driven by a stochastic process correlated to output and consumption.

5.1 Calibration

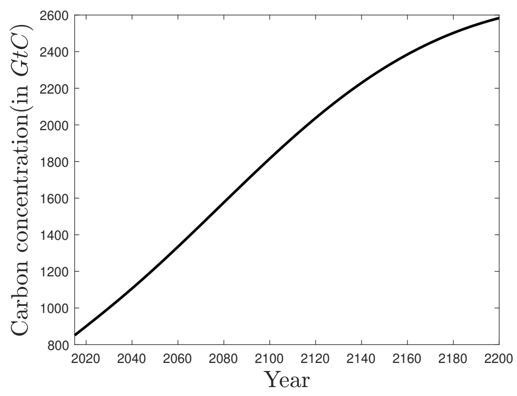
Appendix F gives the full details of the calibration of the climate model. Parameters for the growth rate of emissions and the initial level are chosen to match the baseline scenario of the DICE-2016 calibration (W. D. Nordhaus, 2017). The parameters of the carbon cycle and temperature model are taken from Mattauch et al. (2018). In addition, and different from Mattauch et al. (2018), we also include a base level of non-carbon related radiative forcing and calibrate it to match exogenous forcing in DICE-2016. Figure 3 shows the future path of the climate state variables using our emissions path and climate model. Under the Business-As-Usual scenario, emissions are projected to peak at the end of the century, and decline from then on. The surface temperature will then rise by almost 4 degrees in 2100. Note that the SCC is not very sensitive to assumptions about the emissions scenario, since it will only affect the SCC via the discount rate CDR_t and via the derivative of temperature with respect to the carbon concentration.

The calibration of the economic parameters is given in table 1. Since we consider an exogenous endowment economy, output and consumption are the same thing in our model. That leaves the question open whether we should calibrate the endowment to output or to consumption data. The focus of the paper is on the social cost of carbon. What ultimately matters for the social cost of carbon is consumption, since utility depends on consumption and not on output. To make our results more comparable to other models, we therefore calibrate endowment to consumption data. The next choice to be made is whether one should aggregate output or consumption data using

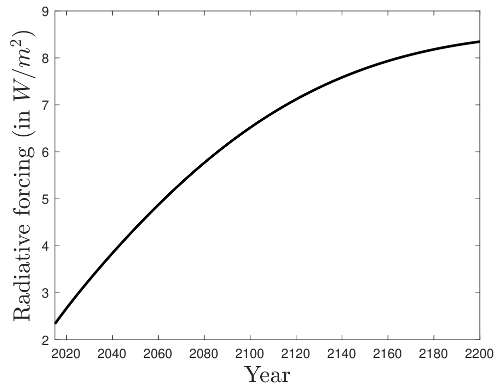
Figure 3: Future path of climate variables.



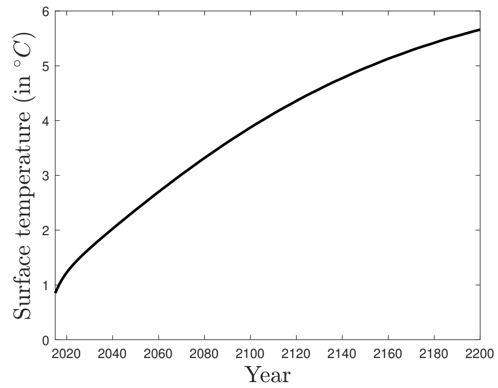
(a) Total emissions



(b) Carbon concentration



(c) Radiative forcing



(d) Surface temperature

market exchange rates or using purchasing power parities (PPP). In line with the DICE-2016 model we use purchasing power parity exchange rates. Consumption data is not directly available in PPP. To obtain a proxy for world consumption in PPP we first obtain output data in PPP. Then we determine the world consumption ratio using market exchange rates. Our proxy for world consumption in PPP is then output in PPP multiplied by the world consumption ratio. Real world GDP (PPP) in 2015 equals 114.137 trillion 2015$ (IMF World Economic Outlook October 2016). World consumption in 2015 using market exchange rates equals 55.167 (in trillion 2010 $), while world GDP using market exchange rates equals 75.803 (in trillion 2010 $) (Worldbank Database). This yields a consumption-output ratio of 72.78%. Applying this ratio to World GDP (PPP) then gives 83.065 (in trillion 2015 $) for aggregate consumption in PPP terms.

The next step is to calibrate the climate disaster distribution, and in particular the parameters λ_T and η. Our setup does allow for an arrival rate that is convex in temperature, but we do not consider this extension since it would give another free parameter to calibrate. Karydas and Xepapadeas (2019) also consider climate disasters and assume, based on natural disaster data, that for every degree warming

Table 1: Parameters for the economic model

Par.	Description	Value
C_t	Initial consumption level (PPP, in trillion 2015$)	83.07
λ_T	Arrival rate parameter	0.02 / 0.04
η	Disaster size parameter	30.25 / 61.5
$E[J]$	Expected disaster size	-0.032 / -0.016
γ	Risk aversion	5
θ	Ambiguity aversion parameter	0.1
a^*	Optimal ambiguity parameter	1.27 / 1.30
b^*	Optimal ambiguity parameter	0.74 / 0.75
ϵ	Elasticity of substitution	1.5
CDR_0	Consumption discount rate	1.5%

the arrival rate increases by 6%. The disaster size is calibrated to 1.6%. This implies that the expected growth loss due to climate change would be $6\% \times 1.6\% = 0.096\%$ per degree global warming. W. D. Nordhaus (2017) models the economic impact of climate change as the percentage loss of output as a function of temperature (level impact). Hambel et al. (2021) consider both a level and a growth impact of climate damages. They find that a loss of 0.026% per degree warming gives the same GDP loss in the year 2100 as the level impact of W. D. Nordhaus (2017). Setting the disaster size to 1.6% and calibrating λ_T such that on average climate disasters lead to a loss of 0.026% gives $\lambda_T = 1.63\%$, much lower than the arrival rate assumed in Karydas and Xepapadeas (2019). The latter obviously get much higher expected damages than the calibration of W. Nordhaus (2014) yields.

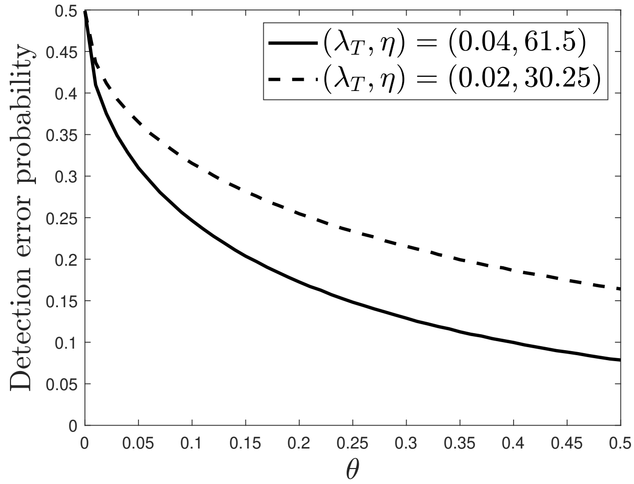
We decide to choose $\lambda_T = 4\%$, which is in between these two calibrations and set $\eta = 61.5$ which yields $E_t[J_t] = -1.6\%$, in line with Karydas and Xepapadeas (2019). Additionally, we consider a variant with less frequent but on average larger disasters: $\lambda_T = 2\%$, and a disaster size parameter $\eta = 30.25$ which gives $E_t[J_t] = -3.2\%$. While both calibrations have on average the same impact, their impact on risk premia is very different.

We now turn to the calibration of risk aversion and ambiguity aversion. We set risk aversion equal to 5. This level of risk aversion can be seen as conservative if we compare it to common values in the asset pricing literature.[7]

The level of ambiguity aversion is harder to calibrate. To get a feeling for reasonable values of ambiguity aversion, we use the concept of detection error probabilities. The ambiguity aversion parameter θ pins down the arrival rate and the expected jump size in the worst-case scenario. A higher θ leads to a higher worst-case arrival rate and a more negative worst-case expected jump size. The detection error probability is the probability of choosing the wrong model (so choosing the reference model $\mathbb{P}$ when the worst-case $\mathbb{Q}^{a,b}$ is true and vice-versa). When θ is higher, the two models are more different and the probability of making a mistake is therefore lower. When

[7]A coefficient of relative risk aversion between 5 and 10 is common in the asset pricing literature according to (Cochrane, 2009).

Figure 4: Detection error probabilities as a function of θ.



the detection error probability is close to 50%, the two models are very similar. This is an indication of a low ambiguity aversion parameter. On the other hand, when the detection error probability is close to 0, it is easy to distinguish the worst-case model from the reference model. This indicates that the worst-case model is extreme and the ambiguity aversion parameter very high.

We calculate the detection error probability assuming that the consumption process can be observed over a period of 100 years. The ambiguity aversion parameter θ is varied between 0 and 0.5. The results are given in figure 4. Detection error probabilities are decreasing in θ and are higher for a lower λ_T. This is intuitive, since a lower λ_T implies that there are less disasters over the observed time period and the probability of choosing the wrong model is therefore larger. We choose to set $\theta = 0.1$ in the base calibration, which gives a detection error probability of 24.6% for $(\lambda_T, \eta) = (0.02, 30.25)$ and 31.6% for $(\lambda_T, \eta) = (0.04, 61.5)$ (cf figure 4). This level of ambiguity aversion balances the trade-off between wanting to make a robust decision, but not taking into account too extreme models. The detection error probabilities for $\theta = 0.1$ are sufficiently far away from 50%, which implies that the reference model and the worst case model are not too similar. On the other hand, the detection error probabilities are also not close to 0, which would indicate an extreme amount of ambiguity aversion. However, since this parameter remains hard to calibrate, we do vary θ in robustness checks.

For the calibration $(\lambda_T, \eta) = (0.04, 61.5)$ the resulting optimal parameters with $\theta = 0.1$ are: $a^* = 1.30$ and $b^* = 0.75$. The arrival rate under the worst-case probability measure is 30% higher compared to the reference model. And the expected jump size becomes $\frac{-1}{b^*\eta+1} = -2.12\%$ compared to -1.6% in the reference model. The optimal parameters for the case $(\lambda_T, \eta) = (0.02, 30.25)$ are quite similar: $a^* = 1.27$ and $b^* = 0.74$.

The parameters that still have to be calibrated affect the social cost of carbon only indirectly, via the discount rate. Equation (15) shows that one can separate the expression for the Consumption Discount Rate (the relevant discount rate for the social cost of carbon) CDR_t in a time-independent part CDR_0 and a part that does depend on time as:

$$\begin{aligned} CDR_t &= CDR_0 + (1/\epsilon - 1)a^*\lambda_t \frac{-1}{b^*\eta + 1 - \gamma} \\ CDR_0 &= \beta + (1/\epsilon - 1)\Big(\mu - \frac{\gamma}{2}\sigma^2\Big). \end{aligned} \tag{19}$$

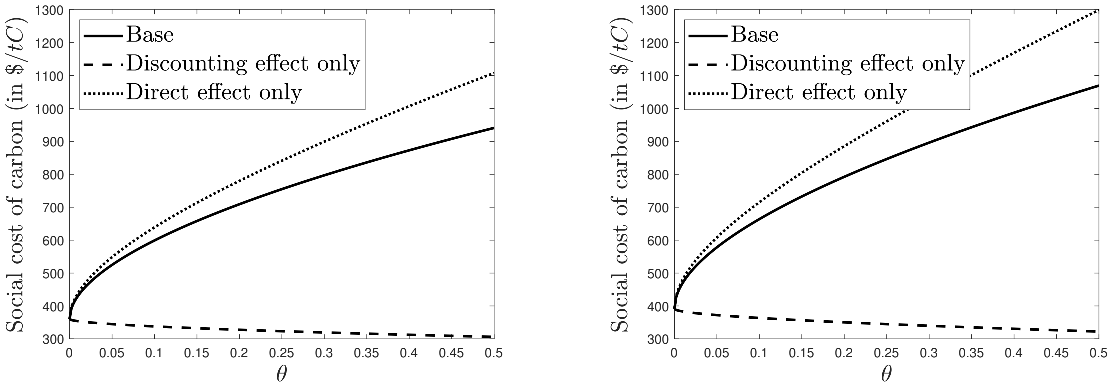
CDR_0 is the consumption discount rate in the absence of climate disasters. First, the value of the elasticity of intertemporal substitution ϵ determines whether additional risk increases or decreases the discount rate. Generally, there is strong empirical evidence of an EIS larger than one (Van Binsbergen, Fernández-Villaverde, Koijen, & Rubio-Ramírez, 2012; Vissing-Jørgensen & Attanasio, 2003). When $\epsilon > 1$, we are in the realistic situation that additional risk decreases asset prices. We choose $\epsilon = 1.5$, which is a common value in the literature on Epstein-Zin preferences. The growth rate μ, the volatility σ and the pure rate of time preference β only affect the social cost of carbon via CDR_0. The calibration of β has been widely discussed in the climate change literature. Additionally, we could calibrate σ from observed consumption volatility. However, as Mehra and Prescott (1985) point out, the model in that case would generate a way too low risk premium (the equity premium puzzle). A way to circumvent this is to calibrate σ to the volatility of stock prices, but this solution is also not very satisfactory. There have been several (partial) solutions proposed to the equity premium puzzle, for example including economic disaster risk. Solving the equity premium puzzle goes beyond the scope of this paper. Since both β and σ only affect the SCC via CDR_0, we choose to directly calibrate the consumption discount rate in the absence of climate risk. In our base calibration, we choose $CDR_0 = 1.5\%$, but we show our results for values of CDR_0 between 0.5% and 2.5%. The parameter combinations $(\beta, \mu, \sigma) = (2.25\%, 2.5\%, 3\%)$ and $(\beta, \mu, \sigma) = (1.5\%, 2.5\%, 10\%)$ for example yield a consumption discount rate $CDR_0 = 1.5\%$. Note that the actual consumption discount rate CDR_t is higher because of the impact of climate disasters on discounting.

5.2 The Social Cost of Carbon with non-stochastic emissions: the analytical model quantified

Our base calibration yields a social cost of carbon of \$599 per ton of carbon (\$163 per ton CO_2) with $(\lambda_T, \eta) = (0.04, 61.5)$ and \$664 per ton carbon (\$181 per ton CO_2) with $(\lambda_T, \eta) = (0.02, 30.25)$.[8] Comparing the two cases shows that it matters whether the disasters are frequent but small (large η) or more infrequent but larger (smaller η). The two sets of assumptions yield the same expected disaster shock, but in the low frequency/large-shock case risk aversion and ambiguity aversion play a larger role and the social cost of carbon is correspondingly higher.

[8] We express the social cost of carbon in the rest of this paper in dollars per ton carbon. To convert in dollars per ton CO_2, divide by 3.67.

Figure 5: Social cost of carbon as a function of θ.



(a) $(\lambda_T, \eta) = (0.04, 61.5)$ (b) $(\lambda_T, \eta) = (0.02, 30.25)$

This figure shows the social cost of carbon as a function of the ambiguity aversion parameter θ. The total effect of ambiguity aversion on the SCC is given by the solid line *(base)*. We additionally distinguish two special cases. In the *discounting effect only* case (dashed line) we assume that increasing θ does lead to an increase in the discount rate, but does not change the arrival rate and the certainty equivalent in the SCC formula. In the *direct effect only* case (dotted line) we look at the opposite case, where increasing θ is assumed to have an effect on the arrival rate and the certainty equivalent, but not on the consumption discount rate CDR_t.

Ambiguity aversion and the SCC

Figure 5 shows for each of the two sets of assumptions on the disaster risk parameters the social cost of carbon for different values of θ. Ambiguity aversion clearly leads to a substantially higher social cost of carbon. For the $(\lambda_T, \eta) = (0.04, 61.5)$ case, the SCC is 65% higher with $\theta = 0.1$ compared no the case without ambiguity aversion. The relative increase is even larger when we consider the $(\lambda_T, \eta) = (0.02, 30.25)$ case: the SCC is then 83% higher with ambiguity aversion. The intuition behind this difference is that risk aversion and ambiguity aversion have a larger effect with less frequent but larger disasters. Therefore the relative increase in the SCC due to ambiguity aversion is larger with the $\lambda_T = 0.02$ process than it is with $\lambda_T = 0.04$ setup.

We saw already from equation (19) that ambiguity aversion affects both the arrival rate and the certainty equivalent of climate disasters, but also the discount rate. In our calibration with $\epsilon = 1.5 > 1$, more ambiguity aversion leads to a higher discount rate which means the direct effect via the arrival rate and the certainly equivalent and the indirect effect via the discount rate have the opposite effect on the SCC. We show the two effects separately and combined in Figure 5. There we consider the indirect *discounting only* effect, in which we assume ambiguity aversion only affects the discount rate CDR_t (the dashed line); and the direct effect where we leave the consumption discount rate CDR_t unchanged, but take into account the direct effect of ambiguity aversion on the arrival rate and certainty equivalent of the climate disasters (dotted line) in figure 5. The two effects are combined in the case labeled "Base" (solid line). Figure 5 clearly indicates that ambiguity aversion increases the discount rate (remember we assume $\epsilon > 1$ in this set of simulations); but we also see that the direct effect on the SCC dominates, the solid line slopes upward. We conclude that even for $\epsilon > 1$ ambiguity aversion leads to a higher social cost of carbon, and in our

Table 2: Social cost of carbon as function of risk aversion and ambiguity aversion

Social Cost of Carbon	$(\lambda_T, \eta) = (0.04, 61.5)$	$(\lambda_T, \eta) = (0.02, 30.25)$
$\gamma = 0$, $\theta = 0$	363	363
$\gamma = 5$, $\theta = 0$	360	392
$\gamma = 5$, $\theta = 0.1$	599	664

calibration actually substantially so.

The elasticity of intertemporal substitution ϵ and the SCC

The sign of the discounting effect depends on the value of ϵ. When $\epsilon < 1$, additional risk, more risk aversion or more ambiguity aversion would lower discount rates and both the indirect discounting effect and the direct effect of ambiguity aversion would have the same sign. However, this leads to counter-intuitive effects. For example $\epsilon < 1$ implies that the consumption discount rate decreases when the volatility of consumption increases. For $\epsilon = 1$, the consumption discount rate CDR_t simply equals β and risk, risk aversion and ambiguity aversion do not affect discount rates.

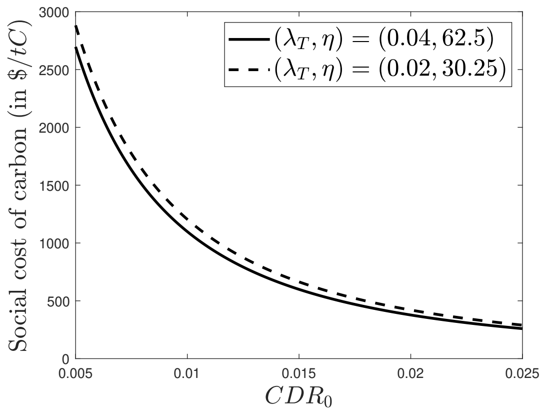
Risk aversion, ambiguity aversion and the SCC

In table 2 we compare the effect of risk aversion and of ambiguity aversion on the SCC numerically. By definition, the SCC is the same for both calibrations when risk aversion γ and ambiguity aversion θ are both 0. In that case the expected value of both calibrations is the same and since risk is not priced under those assumptions, the SCC is the same for both calibrations. Introducing risk aversion has a negligible effect on the SCC for the frequent disasters with low disaster size: for $(\lambda_T, \eta) = (0.04, 62.5)$ the direct impact of risk aversion on the certainty equivalent is small and is canceled out by the *discounting* effect: for this configuration the SCC is even slightly lower than what it is without risk aversion. This changes when damages are more infrequent but larger. In the alternative calibration with $(\lambda_T, \eta) = (0.02, 30.25)$, risk aversion does increase the social cost of carbon, from \$363 to \$392. Either way the impact of risk aversion on the SCC is not substantial. But the analysis shows very different results for ambiguity aversion. In both cases, introducing ambiguity aversion leads to a significantly higher value of the social cost of carbon, when θ goes from 0 to 0.1, the SCC increases by almost 70%. The table thus shows that that risk and ambiguity aversion have very different implications for the valuation of climate risk.

Discount rates and the SCC

Figure 6 shows the dependence of the SCC on the time-independent part of the consumption discount rate CDR_0, the core discount rate. Note that the actual discount rate that is used to discount future damages (CDR_t) is higher than CDR_0 due to the effect of climate disasters itself on discounting. When core discount rates are close to zero, the social cost of carbon becomes very high. With $CDR_0 = 0.5\%$, the SCC is even above \$2000, around four times higher than in the base calibration. On the other hand, setting $CDR_0 = 2.5\%$ gives a social cost of carbon that is less than half the value in the base calibration. This figure highlights the importance of the discount rate when analyzing climate change and in particular its impact on the social cost of carbon.

Figure 6: Social cost of carbon as a function of CDR_0.



5.3 The Full Model: the Social cost of carbon with stochastic emissions

So far we have made the obviously counterfactual assumption that emissions are a non-stochastic process, since assuming otherwise would preclude analytical solutions. In this section we remedy this shortcoming by modeling emissions as an explicitly stochastic process correlated to the process generating output. The short answer to the question what this brings about is that the main results are still true in this more realistic case. But stochastic emission processes add to risk and uncertainty, with as implication that higher risk aversion and more am case.

Thus assume now that emissions are the product of carbon intensity ψ_t and aggregate endowment C_t: $E_t = \psi_t C_t$. We calibrate the stochastic process for ψ_t such that in expectation emissions are similar to what they are in the non-stochastic emissions case. To bring this about we postulate that ψ_t declines at the rate $\psi_0 e^{-\alpha_\psi t} + \delta_\infty(1 - e^{-\alpha_\psi t})$ and set $\delta_0^\psi = -0.6\%$, $\delta_\infty^\psi = -6\%$ and $\alpha_\psi = 0.0045$. All other parameters are the same as in the exogenous case. The only difference is that future emissions are now also stochastic and correlated to output. The solution method is described in appendix G.

Table 3: Social cost of carbon as function of risk aversion and ambiguity aversion with stochastic emissions correlated to output

Social Cost of Carbon	$(\lambda_T, \eta) = (0.04, 61.5)$	$(\lambda_T, \eta) = (0.02, 30.25)$
$\gamma = 0$, $\theta = 0$	352	352
$\gamma = 5$, $\theta = 0$	368	399
$\gamma = 5$, $\theta = 0.1$	609	673

The results are given in Table 3. For zero risk aversion and in the absence of ambiguity aversion ($\gamma = 0$ and $\theta = 0$) the SCC is slightly smaller compared to the exogenous emissions case, although negligibly so: 352 $\$/tC$ instead of 363 $\$/tC$. But with endogenous and stochastic emissions, both risk and ambiguity aversion have larger effects on the social cost of carbon. For the high expected damages parametrization of the climate damages jump process $(\lambda_T, \eta) = (0.02, 30.25)$, column two in Table 3 shows that increasing γ from 0 to 5 leads to a 13 % increase in the SCC (go from the first to the second row in column two of table 3). Adding ambiguity aversion (go from the second to the third row in column two of table 3) leads to a further 69 % increase in the SCC. The combined impact of going from no risk/ambiguity aversion to our base case assumptions on the risk and ambiguity aversion parameters is a 91% increase in the SCC, up from 83% in the non-stochastic emissions case.

For the alternative low-expected-damages case $(\lambda_T, \eta) = (0.04, 61.5)$, column one in Table 3, the impact of increasing γ and θ is smaller although still sizable. Increasing γ from 0 to 5 leads to a 5 % increase in the SCC and subsequently increasing θ to 0.1 leads to a substantially larger increase in the SCC of 65%. The combined impact is in this case a somewhat smaller but still large: an increase in the SCC of 73%.

Overall Table 3 shows that our main results remain valid in the more realistic case where emissions are a function of aggregate endowment, while the impact of risk aversion and ambiguity aversion now is even a bit larger. These results should be intuitive since with endogenous emissions there is additional risk within the model: future emissions are now also stochastic. Especially risk aversion will therefore have a larger effect.

6 Conclusions

Climate change will beyond reasonable doubt have a large impact on economic growth in the future. However, because of the complex nature of the problem and the lack of data, it is not possible to accurately estimate the timing and extent of its impact. But one thing we do know is that potentially large and irreversible consequences are likely to take place unless mitigating policies are implemented in time. But these changes will happen possibly far into the future, while mitigating policies are (or should be) under consideration right now. That discrepancy should put the discussion on discounting at the center of the debate about the social cost of carbon and what we should do about climate change: to compare uncertain future damages with costs today, those future damages need to be discounted back towards today. The debate in the literature has largely zeroed in on the rate of time preference; the problem there is that to be consistent with capital market data, discount rates must be relatively high which in turn does not leave much once climate change consequences a century out are discounted back towards today (cf Weitzman (2007) for a very lucid overview of this debate). In this paper we also focus on the discounting question and its implication for the SCC, but we take a different approach. Rather than discussing numerical values of certain parameters, we explore alternative specifications of preferences, and show the implications for the social cost of carbon.

We focus on the effect of Epstein-Zin recursive preferences on outcomes of the model, on the impact of unmeasurable risk (ambiguity) and the interaction between those two. Both breaking the link between γ and the EIS (by introducing Epstein-Zin utility) and introducing ambiguity aversion are conceptually relevant in the climate change setting. The first extension is relevant because climate change problems have a very long horizon and therefore the elasticity of intertemporal substitution (EIS) unavoidably plays an important role. Arbitrarily restricting its value to $1/\gamma$ is then surely unsatisfactory. Second, conceptually ambiguity aversion is a logical extension, since we have no accurate estimation of climate damages nor in particular of their probability density function in the future. The assumption of unmeasurable risk ("Knightian uncertainty") then is a natural framework to use. Finally we highlight the sometimes complicated interactions between ambiguity aversion and intertemporal substitution elasticities for the value of the Social Cost of Carbon.

To do all this we set up an analytic IAM by extending a disaster risk model with a climate change model and a temperature dependent arrival rate. Furthermore, we model climate risk as tail risk instead of assuming that temperature increases generate a certain amount of damage every year. The model is transparent because we manage to derive closed form solutions for the social cost of carbon. Where stochastic numerical IAMs can take hours to be solved, solving our model only requires numerical integration and is therefore solved within seconds.

Our base calibration generates a substantial social cost of carbon, is between \$599 and \$664 per ton of carbon (\$163 - \$181 per ton CO_2) with non-stochastic emissions, and slightly more for the stochastic case (between \$609 and \$673 per ton of carbon). This is much higher than for example the estimate of \$114 per ton carbon that is obtained using the DICE-2016R model (W. D. Nordhaus, 2017), and also higher than current market prices in for example the European Emissions Trading System by the time of writing this paper. Most importantly, we use our model to highlight how ambiguity aversion changes the social cost of carbon.

Analysing the effect of ambiguity aversion on the SCC is not a trivial exercise since multiple potentially offsetting effects play a role: we show that ambiguity aversion has both a *direct* effect on the arrival rate and certainty equivalent of disasters for given discount rates (more ambiguity aversion leads to a higher certainty equivalent) and an *indirect* effect on the discounting component. The effect of ambiguity aversion on discounting depends on the intertemporal rate of substitution ϵ. When $\epsilon < 1$, increasing ambiguity aversion leads to a smaller effective discount rate on climate damages, making for a higher SCC since both the direct and the indirect effect work in the same direction. For the arguably interesting (because empirically supported) case $\epsilon > 1$, increasing ambiguity aversion has two offsetting effects on the SCC, the direct and indirect effects actually work in different direction. While the direct effect is qualitatively the same as in the $\epsilon < 1$ case, the indirect effect now has the opposite direction because with $\epsilon > 1$ discount rates actually go up with higher risk aversion and ambiguity aversion. However, we show that even then the direct effect dominates when evaluated numerically and therefore that the presence of ambiguity aversion leads to a (substantially) higher social cost of carbon.

Lastly, we also show the importance of fully considering the impact of the consumption discount rate on the social cost of carbon, not just the impact of the rate of time preference. It is of course well known that the social cost of carbon is very sensitive to changes in the discount rate, but we stress that analyzing the discount rate impact of climate change involves more than a discussion of the pure rate of time preference on the discount rate; a low discount rate can also be caused by a high elasticity of intertemporal substitution, and additionally the appropriate discount rate depends in elaborate ways on the growth rate of the economy, volatility, risk aversion, climate disaster risk and ambiguity aversion. Disentangling these various effects, their interactions and their impact on the SCC is the key contribution of this paper. One major theme emerges: proper risk pricing and incorporating ambiguity aversion leads to much higher estimates of the Social Cost of Carbon, literally by orders of magnitude. These findings are surely of more than just academic interest.

References

Anderson, E. W., Hansen, L. P., & Sargent, T. J. (2003). A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *Journal of the European Economic Association*, *1*(1), 68–123.

Bansal, R., Kiku, D., & Ochoa, M. (2016). *Price of long-run temperature shifts in capital markets* (Tech. Rep.). National Bureau of Economic Research.

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance*, *59*(4), 1481–1509.

Barnett, M., Brock, W., & Hansen, L. P. (2020). Pricing uncertainty induced by climate change. *The Review of Financial Studies*, *33*(3), 1024–1066.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, *121*(3), 823–866.

Barro, R. J. (2009). Rare disasters, asset prices, and welfare costs. *American Economic Review*, *99*(1), 243–64.

Barro, R. J. (2015). Environmental protection, rare disasters and discount rates. *Economica*, *82*(325), 1–23.

Bretscher, L., & Vinogradova, A. (2018). Escaping damocles' sword: Endogenous climate shocks in a growing economy. *Economics Working Paper Series*, *18*.

Cai, Y., & Lontzek, T. S. (2019). The social cost of carbon with economic and climate risks. *Journal of Political Economy*, *127*(6), 2684–2734.

Chen, Z., & Epstein, L. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica*, *70*(4), 1403–1443.

Chichilnisky, G., Hammond, P. J., & Stern, N. (2018). Should we discount the welfare of future generations? ramsey and suppes versus koopmans and arrow. *Warwick economics research papers series (WERPS)(1174), http://wrap. warwick. ac. uk/107726*.

Cochrane, J. H. (2009). *Asset pricing: Revised edition.* Princeton university press.

Crost, B., & Traeger, C. P. (2013). Optimal climate policy: uncertainty versus monte carlo. *Economics Letters*, *120*(3), 552–558.

Dietz, S. (2011). High impact, low probability? an empirical analysis of risk in the economics of climate change. *Climatic Change*, *108*(3), 519–541.

Dietz, S., van der Ploeg, F., Rezai, A., & Venmans, F. (2021). Are economists getting climate dynamics right and does it matter? *Journal of the Association of Environmental and Resource Economists*, *8*(5), 895–921.

Duffie, D., & Epstein, L. G. (1992a). Asset pricing with stochastic differential utility. *The Review of Financial Studies*, *5*(3), 411–436.

Duffie, D., & Epstein, L. G. (1992b). Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, 353–394.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, 643–669.

Epstein, L. G., & Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica: Journal of the Econometric Society*, 937–969.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, *18*(2), 141 - 153. Retrieved from `http://www.sciencedirect.com/science/article/pii/0304406889900189` doi: https://doi.org/10.1016/0304-4068(89)90018-9

Golosov, M., Hassler, J., Krusell, P., & Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, *82*(1), 41–88.

Goosse, H. (2015). *Climate system dynamics and modelling*. Cambridge University Press. Retrieved from `https://books.google.at/books?id=2-NBCgAAQBAJ`

Hambel, C., Kraft, H., & Schwartz, E. (2021). Optimal carbon abatement in a stochastic equilibrium model with climate change. *European Economic Review*, *132*, 103642.

Hansen, L. P., & Sargent, T. J. (2008). *Robustness*. Princeton university press.

Hope, C. (2006). The marginal impact of co2 from page2002: An integrated assessment model incorporating the ipcc's five reasons for concern. *Integrated assessment*, *6*(1).

IPCC. (2013). *Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change* [Book]. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Retrieved from `www.climatechange2013.org` doi: 10.1017/CBO9781107415324

IPCC. (2021). *Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change* [Book]. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Jensen, S., & Traeger, C. P. (2014). Optimal climate change mitigation under long-term growth uncertainty: Stochastic integrated assessment and analytic findings. *European Economic Review*, *69*, 104–125.

Karydas, C., & Xepapadeas, A. (2019). *Climate change risks: pricing and portfolio allocation* (Tech. Rep.). CER-ETH-Center of Economic Research (CER-ETH) at ETH Zurich.

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision

making under ambiguity. *Econometrica*, *73*(6), 1849–1892.
Knight, F. H. (1921). Risk, uncertainty and profit. *New York: Hart, Schaffner and Marx*.
Lemoine, D., & Traeger, C. P. (2016). Ambiguous tipping points. *Journal of Economic Behavior & Organization*, *132*, 5–18.
Liu, J., Pan, J., & Wang, T. (2004). An equilibrium model of rare-event premia and its implication for option smirks. *The Review of Financial Studies*, *18*(1), 131–164.
Lucas Jr, R. E. (1978). Asset prices in an exchange economy. *Econometrica: Journal of the Econometric Society*, 1429–1445.
Maenhout, P. J. (2004). Robust portfolio rules and asset pricing. *Review of financial studies*, *17*(4), 951–983.
Maenhout, P. J. (2006). Robust portfolio rules and detection-error probabilities for a mean-reverting risk premium. *Journal of Economic Theory*, *128*(1), 136–163.
Mattauch, L., Millar, R., van der Ploeg, R., Rezai, A., Schultes, A., Venmans, F., ... Teytelboym, A. (2018). Steering the climate system: an extended comment.
Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of monetary Economics*, *15*(2), 145–161.
Millner, A., Dietz, S., & Heal, G. (2013). Scientific ambiguity and climate policy. *Environmental and Resource Economics*, *55*(1), 21–46.
Munk, C. (2015). *Dynamic asset allocation*.
Nordhaus, W. (2014). Estimates of the social cost of carbon: concepts and results from the dice-2013r model and alternative approaches. *Journal of the Association of Environmental and Resource Economists*, *1*(1/2), 273–312.
Nordhaus, W. D. (2014). *A question of balance: Weighing the options on global warming policies*. Yale University Press.
Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 201609244.
Olijslagers, S. (2020). The social cost of carbon: Optimal policy versus business-as-usual.
Olijslagers, S. (2021). Solution methods for dsge models in continuous time: application to a climate-economy model. *University of Amsterdam*.
Olijslagers, S., van der Ploeg, R., & van Wijnbergen, S. (2021). On current and future carbon prices in a risky world.
Pindyck, R. S. (2017). The use and misuse of models for climate policy. *Review of Environmental Economics and Policy*, *11*(1), 100–114.
Ramsey, F. P. (1928). A mathematical theory of saving. *The economic journal*, *38*(152), 543–559.
Rietz, T. A. (1988). The equity risk premium a solution. *Journal of monetary Economics*, *22*(1), 117–131.
Stern, N. (2015). *Why are we waiting?: The logic, urgency, and promise of tackling climate change*. Mit Press.
Tol, R. S. (2002). Estimates of the damage costs of climate change. part 1: Benchmark estimates. *Environmental and resource Economics*, *21*(1), 47–73.
Traeger, C. P. (2021). Uncertainty in the analytic climate economy.

Tsai, J., & Wachter, J. A. (2015). Disaster risk and its implications for asset pricing. *Annual Review of Financial Economics*, *7*, 219–252.
Van Binsbergen, J. H., Fernández-Villaverde, J., Koijen, R. S., & Rubio-Ramírez, J. (2012). The term structure of interest rates in a dsge model with recursive preferences. *Journal of Monetary Economics*, *59*(7), 634–648.
Van den Bremer, T. S., & Van der Ploeg, F. (2021). The risk-adjusted carbon price. *American Economic Review*, *111*(9), 2782–2810.
Vissing-Jørgensen, A., & Attanasio, O. P. (2003). Stock-market participation, intertemporal substitution, and risk-aversion. *American Economic Review*, *93*(2), 383–391.
Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, *68*(3), 987–1035.
Weitzman, M. L. (2007). A review of the stern review on the economics of climate change. *Journal of economic literature*, *45*(3), 703–724.

A Relative Entropy and the Distance Measure d(a,b)

For each a and b we define the measure $\mathbb{Q}^{a,b}$ which is equivalent to $\mathbb{P}$ and has Radon-Nikodym derivative $\xi_t^{a,b}$ where $\xi_t^{a,b}$ follows:

$$d\xi_t^{a,b} = (\lambda_t - \lambda_t^{\mathbb{Q}^{a,b}})\xi_t^{a,b}dt + \Big(\frac{\lambda_t^{\mathbb{Q}^{a,b}} f^{\mathbb{Q}^{a,b}}(J_t)}{\lambda_t f(J_t)} - 1\Big)\xi_{t-}^{a,b}dN_t. \tag{20}$$

Under the alternative measure $\mathbb{Q}^{a,b}$ the arrival rate equals $\lambda_t^{\mathbb{Q}^{a,b}} = a\lambda_t$ and the jump size parameter equals $\eta^{\mathbb{Q}^{a,b}} = b\eta$. We can calculate in this case the fraction of the two probability distributions: $\frac{f^{\mathbb{Q}^{a,b}}(x)}{f(x)} = b(1+x)^{(b-1)\eta}$. Substituting this into (20) gives:

$$d\xi_t^{a,b} = (1-a)\lambda_t\xi_t^{a,b}dt + \Big(ab(1+J_t)^{(b-1)\eta} - 1\Big)\xi_{t-}^{a,b}dN_t. \tag{21}$$

The Radon-Nikodym derivative $\xi_t^{a,b}$ is the ratio between the alternative measure $\mathbb{Q}^{a,b}$ and the reference measure $\mathbb{P}$. The relative entropy between $\mathbb{Q}^{a,b}$ and $\mathbb{P}$ over time unit Δ is defined as $E_t^{\mathbb{Q}^{a,b}}\Big[\log(\frac{\xi_{t+\Delta}^{a,b}}{\xi_t^{a,b}})\Big]$. Here $E_t^{\mathbb{Q}^{a,b}}$ denotes the expectation with respect to the alternative measure $\mathbb{Q}^{a,b}$. Then divide by Δ and let $\Delta \to 0$ to obtain the instantaneous relative entropy: $RE(a,b) = \lim_{\Delta\to 0}\frac{1}{\Delta}E_t^{\mathbb{Q}^{a,b}}\Big[\log\Big(\frac{\xi_{t+\Delta}^{a,b}}{\xi_t^{a,b}}\Big)\Big]$.

Applying Itô's lemma for jump processes to $\xi_t^{a,b}$, we obtain the following dynamics for $\log(\xi_t^{a,b})$:

$$d\log(\xi_t^{a,b}) = (1-a)\lambda_t dt + \Big(\log(ab) + (b-1)\eta\log(1+J_t)\Big)dN_t. \tag{22}$$

Using integration by parts we can calculate that $E_t^{\mathbb{Q}^{a,b}}[\log(1+J_t)] = -\frac{1}{\eta^{\mathbb{Q}}}$. Therefore the (instantaneous) relative entropy at time t equals:

$$\begin{aligned} RE(a,b) = \lim_{\Delta\to 0}\frac{1}{\Delta}E_t^{\mathbb{Q}^{a,b}}\Big[\log\Big(\frac{\xi^{a,b}_{t+\Delta}}{\xi^{a,b}_t}\Big)\Big] = \\ (1-a)\lambda_t + a\lambda_t\Big(\log(ab)+\frac{1}{b}-1\Big). \end{aligned} \tag{23}$$

Scaling relative entropy by the arrival rate λ_t yields our distance measure:

$$d(a,b) = \frac{RE(a,b)}{\lambda_t} = (1-a) + a\Big(\log(ab)+\frac{1}{b}-1\Big). \tag{24}$$

B Calculating the detection error probability

After observing the process of consumption over a period N years, what is the probability of choosing the wrong model? Let us start with the case that the reference model $\mathbb{P}$ is the true model and the agent considers the alternative model $\mathbb{Q}^{a,b}$. Note that the Radon-Nikodym derivative informs us about the likelihood ratio of both models. When this derivative is larger than one after N years, the worst-case model $\mathbb{Q}^{a,b}$ is the most likely and the agent will choose the wrong model. The probability of making this error is equal to (see for example Maenhout (2006)):

$$P\Big(\xi^{a,b}_N > 1|\mathbb{P}\Big) = P\Big(\log(\xi^{a,b}_N) > 0|\mathbb{P}\Big). \tag{25}$$

We calculate this probability by simulating the process of $\log(\xi^{a,b}_t)$ forward. Simulation is done via a standard Euler method. Similarly, we can define the opposite mistake where the alternative model is actually true and the agent chooses the reference model. We now define the inverse Radon-Nikodym derivative: $\frac{d\mathbb{P}}{d\mathbb{Q}^{a,b}} = \tilde{\xi}_t^{\,a,b}$ where $\tilde{\xi}_t^{\,a,b}$ follows:

$$d\tilde{\xi}_t^{\,a,b} = (a-1)\lambda_t\tilde{\xi}_t^{\,a,b}dt + \Big(\frac{1}{ab}(1+J)^{(1-b)\eta}-1\Big)\tilde{\xi}_{t-}^{\,a,b}dN_t. \tag{26}$$

Applying Itô's lemma gives:

$$d\log(\tilde{\xi}^{a,b}_t) = (a-1)\lambda_t dt + \Big(-\log(ab)+(1-b)\eta\log(1+J_t)\Big)dN_t. \tag{27}$$

The probability of choosing the wrong model when actually the alternative model $\mathbb{Q}^{a,b}$ is true equals:

$$P\Big(\tilde{\xi_N}^{a,b} > 1|\mathbb{Q}^{a,b}\Big) = P\Big(\log(\tilde{\xi_N}^{a,b}) > 0|\mathbb{Q}^{a,b}\Big). \tag{28}$$

Again this probability can be calculated by simulating the process $\log(\tilde{\xi}_t)$ forward. The detection error probability is then defined as:

$$\frac{1}{2}P\Big(\log(\xi^{a,b}_N) > 0|\mathbb{P}\Big) + \frac{1}{2}P\Big(\log(\tilde{\xi_N}^{a,b}) > 0|\mathbb{Q}\Big). \tag{29}$$

C Solving the model

C.1 Hamilton-Jacobi-Bellman equation

We will first derive the Hamilton-Jacobi-Bellman equation for every measure $\mathbb{Q}^{a,b}$. In the next subsection of the appendix we introduce ambiguity.

The value function $V_t^{\mathbb{Q}^{a,b}} = V^{\mathbb{Q}^{a,b}}(C_t, X_t)$ is a function of aggregate consumption C_t and the vector of climate state variables X_t. Let $V_C^{\mathbb{Q}^{a,b}}$ denote the first derivative of the value function with respect to aggregate consumption, similar notation is used for the second derivative. For notational purposes, define the vector of climate state variables:

$$X_t = [g_{E,t}\ E_t\ M_{0,t}\ M_{1,t}\ M_{2,t}\ M_{3,t}\ F_{E,t}\ T_t\ T_t^{oc}]'. \tag{30}$$

The vector of state variables then follows: $dX_t = \mu_X(X_t)dt$. Denote by $V_X^{\mathbb{Q}^{a,b}}$ the row vector of partial derivatives of the value function $V_t^{\mathbb{Q}^{a,b}}$ with respect to the vector of state variables X_t: $V_X^{\mathbb{Q}^{a,b}} = \left[\frac{\partial V^{\mathbb{Q}^{a,b}}(C_t,X_t)}{\partial g_{E,t}} \ \ldots \ \frac{\partial V^{\mathbb{Q}^{a,b}}(C_t,X_t)}{\partial T_t^{oc}}\right]$.

Duffie and Epstein (1992b) show that the HJB-equation for stochastic differential utility equals:

$$0 = f(C_t, V_t^{\mathbb{Q}^{a,b}}) + \mathcal{DV}^{\mathbb{Q}^{a,b}}. \tag{31}$$

Here $\mathcal{DV}^{\mathbb{Q}^{a,b}}$ is the drift of the value function. In order to calculate the drift of the value function, we will apply Itô's lemma. By Itô's lemma for jump processes we have:

$$\begin{aligned} dV_t^{\mathbb{Q}^{a,b}} &= V_C^{\mathbb{Q}^{a,b}}\Big(\mu C_t dt + \sigma C_t dZ_t\Big) + V_X^{\mathbb{Q}^{a,b}}\mu_X(X_t)dt + \frac{1}{2}V_{CC}^{\mathbb{Q}^{a,b}}\sigma^2 C_t^2 dt \\ &+ \Big(V^{\mathbb{Q}^{a,b}}\big((1+J_t)C_{t-}, X_t\big) - V^{\mathbb{Q}^{a,b}}(C_{t-}, X_t)\Big)dN_t. \end{aligned} \tag{32}$$

Then the drift under $\mathbb{Q}^{a,b}$ equals:

$$\begin{aligned} \mathcal{DV}^{\mathbb{Q}^{a,b}} &= V_C^{\mathbb{Q}^{a,b}}\mu C_t + V_X^{\mathbb{Q}^{a,b}}\mu_X(X_t) + \frac{1}{2}V_{CC}^{\mathbb{Q}^{a,b}}\sigma^2 C_t^2 \\ &+ \lambda_t^{\mathbb{Q}^{a,b}} E^{\mathbb{Q}^{a,b}}\big[V^{\mathbb{Q}^{a,b}}\big((1+J_t)C_{t-}, X_t\big) - V^{\mathbb{Q}^{a,b}}(C_{t-}, X_t)\big]. \end{aligned} \tag{33}$$

This gives the following Hamilton-Jacobi-Bellman equation:

$$\begin{aligned} 0 &= f(C_t, V_t^{\mathbb{Q}^{a,b}}) + V_C^{\mathbb{Q}^{a,b}}\mu C_t + V_X^{\mathbb{Q}^{a,b}}\mu_X(X_t) + \frac{1}{2}V_{CC}^{\mathbb{Q}^{a,b}}\sigma^2 C_t^2 \\ &+ \lambda_t^{\mathbb{Q}^{a,b}} E^{\mathbb{Q}^{a,b}}\big[V^{\mathbb{Q}^{a,b}}\big((1+J_t)C_{t-}, X_t\big) - V^{\mathbb{Q}^{a,b}}(C_{t-}, X_t)\big]. \end{aligned} \tag{34}$$

We conjecture and verify that the value function is of the following form:

$$V^{\mathbb{Q}^{a,b}}(C_t) = g^{\mathbb{Q}^{a,b}}(X_t)\frac{C_t^{1-\gamma}}{1-\gamma}, \tag{35}$$

where $g^{\mathbb{Q}^{a,b}}(X_t)$ is some function of X_t. Substituting our conjecture $V^{\mathbb{Q}^{a,b}}(C_t, X_t) = \frac{g^{\mathbb{Q}^{a,b}}(X_t)C_t^{1-\gamma}}{1-\gamma}$ into $f(C_t, V_t)$ gives:

$$
\begin{aligned}
f(C_t, V^{\mathbb{Q}^{a,b}}(C_t, X_t)) &= \frac{\beta}{1-1/\epsilon} \frac{C_t^{1-1/\epsilon} - \left(g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma}\right)^{\frac{1}{\zeta}}}{\left(g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma}\right)^{\frac{1}{\zeta}-1}} \\
&= \frac{\beta}{1-1/\epsilon}\left(g^{\mathbb{Q}^{a,b}}(X_t)^{1-\frac{1}{\zeta}} C_t^{1-\gamma} - g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma}\right) \\
&= \beta\zeta\left(g^{\mathbb{Q}^{a,b}}(X_t)^{-\frac{1}{\zeta}} - 1\right) V^{\mathbb{Q}^{a,b}}(C_t, X_t).
\end{aligned} \tag{36}
$$

The partial derivatives of V are given by:

$$
\begin{aligned}
V_C^{\mathbb{Q}^{a,b}} &= g^{\mathbb{Q}^{a,b}}(X_t) C_t^{-\gamma}, \quad V_{CC}^{\mathbb{Q}^{a,b}} = -\gamma g^{\mathbb{Q}^{a,b}}(X_t) C_t^{-\gamma-1}, \\
V_X^{\mathbb{Q}^{a,b}} &= \frac{g_X^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma}}{1-\gamma}.
\end{aligned} \tag{37}
$$

Here $g_X^{\mathbb{Q}^{a,b}}$ denotes the row vector with partial derivatives to each of the state variables, similar to $V_X^{\mathbb{Q}^{a,b}}$. Additionally we can calculate the expected impact of a jump on the value function:

$$
\begin{aligned}
&E^{\mathbb{Q}^{a,b}}\left[V^{\mathbb{Q}^{a,b}}\big((1+J_t)C_{t-}, X_t\big) - V^{\mathbb{Q}^{a,b}}(C_{t-}, X_t)\right] = \frac{E^{\mathbb{Q}^{a,b}}\left[(1+J_t)^{1-\gamma}\right] - 1}{1-\gamma} g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma} \\
&= \frac{\frac{b_t\eta}{b_t\eta+1-\gamma} - 1}{1-\gamma} g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma} = \frac{-1}{b_t\eta + 1 - \gamma} g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma}.
\end{aligned} \tag{38}
$$

Substituting $f(C_t, V^{\mathbb{Q}^{a,b}}(C_t, X_t))$ together with the partial derivatives of $V_t^{\mathbb{Q}^{a,b}}$ and the expectation into (34) yields the following equation:

$$
\begin{aligned}
0 = &\frac{\beta}{1-1/\epsilon}\left(g^{\mathbb{Q}^{a,b}}(X_t)^{-\frac{1}{\zeta}} - 1\right) g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma} + \mu g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma} \\
&- \frac{\gamma}{2}\sigma^2 g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma} + \frac{g_X^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma}}{1-\gamma} \mu_X(X_t) + a_t\lambda_t \frac{-1}{b_t\eta+1-\gamma} g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma}.
\end{aligned} \tag{39}
$$

Dividing by $g^{\mathbb{Q}^{a,b}}(X_t) C_t^{1-\gamma}$ gives:

$$
\begin{aligned}
0 = &\frac{\beta}{1-1/\epsilon}\left(g^{\mathbb{Q}^{a,b}}(X_t)^{-\frac{1}{\zeta}} - 1\right) + \mu - \frac{\gamma}{2}\sigma^2 + \frac{g_X^{\mathbb{Q}^{a,b}}(X_t)}{g^{\mathbb{Q}^{a,b}}(X_t)(1-\gamma)} \mu_X(X_t) \\
&+ a_t\lambda_t \frac{-1}{b_t\eta + 1 - \gamma}.
\end{aligned} \tag{40}
$$

C.2 Optimal a and b

Given a probability measure $\mathbb{Q}^{a,b}$, we can solve equation (40) to find $g^{\mathbb{Q}^{a,b}}(X_t)$. Now let us return to the problem with ambiguity. We are not interested in the solution for every single measure $\mathbb{Q}^{a,b}$. The maxmin procedure advocated by Gilboa and

Schmeidler (1989) that we apply in this paper requires us to focus on the worst case distribution, which leads to the following minimization problem:

$$V_t = \min_{\mathbb{Q}^{a,b} \in \mathcal{P}^\theta} V_t^{\mathbb{Q}^{a,b}}. \tag{41}$$

And since every probability measure $\mathbb{Q}^{a,b}$ that we consider in our set $\mathcal{P}^\theta$ is uniquely defined by the parameters a and b, minimizing over $\mathbb{Q}^{a,b}$ is equivalent to minimizing over the parameters a and b. So we can replace the global minimization problem of equation (41) by an instantaneous optimization problem over a and b. The HJB-equation of the problem with ambiguity then becomes:

$$\begin{aligned} 0 = \min_{(a,b) \ s.t. \ d(a,b)\leq\theta} &\Big\{ \frac{\beta}{1-1/\epsilon}\Big(g^{\mathbb{Q}^{a,b}}(X_t)^{-\frac{1}{\zeta}} - 1\Big) + \mu - \frac{\gamma}{2}\sigma^2 \\ &+ \frac{g_X^{\mathbb{Q}^{a,b}}(X_t)}{g^{\mathbb{Q}^{a,b}}(X_t)(1-\gamma)}\mu_X(X_t) + a\lambda_t \frac{-1}{b\eta + 1 - \gamma}\Big\}. \end{aligned} \tag{42}$$

The optimal a and b can thus be found by solving:

$$\min_{(a,b)} a\lambda_t \frac{-1}{b\eta+1-\gamma} \quad s.t. \ \ d(a,b) \leq \theta. \tag{43}$$

We can drop λ_t and write this problem as a constrained optimization problem with Lagrangian:

$$L(a,b,l) = a\frac{-1}{b\eta+1-\gamma} - l\Big(d(a,b) - \theta\Big). \tag{44}$$

Here l is the Lagrange multiplier. a^* and b^* and the Lagrange-multiplier l are the solutions to the following first order conditions:

$$\begin{aligned} \frac{\partial}{\partial a}L(a,b,l) &= \frac{-1}{b\eta+1-\gamma} - l\Big(\log(ab) + \frac{1}{b} - 1\Big) = 0, \\ \frac{\partial}{\partial b}L(a,b,l) &= a\frac{\eta}{(b\eta+1-\gamma)^2} - la\frac{b-1}{b^2} = 0, \\ \frac{\partial}{\partial l_t}L(a,b,l) &= \theta - (1-a) - \Big(\log(ab) + \frac{1}{b} - 1\Big) = 0. \end{aligned} \tag{45}$$

From now on, we use the notation V_t for the optimal value function ($V_t = V_t^{\mathbb{Q}^{a^*,b^*}}$). Similar notation is used for g_t.

C.3 Solving the model

It is typically not possible to solve the partial differential equation of the problem with climate state variables unless one would make the highly restrictive assumption assumption of a unit EIS, which we choose not to do. However we are able to obtain exact solutions for the value function and the consumption-to-wealth ratio without making restrictive assumptions like $EIS = 1$, and the consumption-to-wealth ratio is what we need for assessing the SCC. We will now sketch our approach.

Duffie and Epstein (1992a) derive that the pricing kernel (or stochastic discount factor) with stochastic differential utility equals $\pi_t = \exp\left\{\int_0^t f_V(C_s, V_s)ds\right\} f_C(C_t, V_t)$. However, the pricing kernel has to be adjusted for the ambiguity aversion preferences. Chen and Epstein (2002) show that the pricing kernel in the ambiguity setting should be multiplied by the Radon-Nikodym derivative $\xi_t^{a^*,b^*}$ of the measure corresponding to the optimal a^* and b^*. $\xi_t^{a,b}$ is defined in (20). When calculating the pricing kernel, we obtain an expression that depends on the unknown function $g(X_t)$. But by substituting the HJB-equation into the pricing kernel we obtain an expression that only depends on known parameters.

As an intermediate step it is helpful to introduce the concept of consumption strips. A consumption strip is an asset that pays a unit of aggregate consumption C_s at time time $s > t$. Call its price at time t: $H(C_t, X_t, u)$, where u denotes the time to maturity; $u = s - t$. The price of a consumption strip paying out at time $s > t$ equals:

$$\begin{aligned} H_t &= H(C_t, X_t, u) \\ &= E_t\left[\frac{\pi_s}{\pi_t} C_s\right] = \exp\left\{-\int_t^{t+u} CDR_s ds\right\} C_t. \end{aligned} \tag{46}$$

We will refer to CDR_t as the consumption discount rate. Now define a stock S_t that gives a claim to consumption and therefore it pays a continuous stream of dividends C_t. The value of such a stock then obviously becomes:

$$S_t = \int_0^\infty H(C_t, X_t, u) du. \tag{47}$$

In equilibrium aggregate wealth must be equal to the value of the stock. The state-dependent consumption-wealth ratio therefore equals:

$$k(X_t) = \frac{C_t}{S_t} = \frac{C_t}{\int_0^\infty H(C_t, X_t, u) du} = \left(\int_0^\infty \exp\left\{-\int_t^{t+u} CDR_s ds\right\} du\right)^{-1}. \tag{48}$$

Using the expression for the consumption-wealth ratio, we can calculate the value function. At the optimum (see for example Munk (2015), Ch. 17), we have the envelope condition that $f_C = V_S$. Furthermore, we derived that $V(C_t, X_t) = \frac{g(X_t) C_t^{1-\gamma}}{1-\gamma}$. Using the chain rule we get:

$$V_S = V_C \frac{\partial C}{\partial S} = V_C k(X_t) = g(X_t) C_t^{-\gamma} k(X_t). \tag{49}$$

Also we have for the intertemporal aggregator:

$$f_C = \beta g(X_t)^{\frac{1/\epsilon - \gamma}{1-\gamma}} C_t^{-\gamma}. \tag{50}$$

Together this gives us:

$$g(X_t) = \left(\frac{k(X_t)}{\beta}\right)^{-\frac{1-\gamma}{1-1/\epsilon}}. \tag{51}$$

In appendix D we derive an expression for the consumption discount rate CDR_t. Given the consumption discount rate, we can solve for the consumption-wealth ratio and therefore we know the value function.

D Discount rates

D.1 The Pricing Kernel

Duffie and Epstein (1992a) derive that the pricing kernel with stochastic differential utility equals $\pi_t = \exp\Big\{ \int_0^t f_V(C_s, V_s)ds \Big\} f_C(C_t, V_t)$. However, the pricing kernel has to be adjusted for the ambiguity aversion preferences. Chen and Epstein (2002) show that the pricing kernel in the ambiguity setting should be multiplied by the Radon-Nikodym derivative $\xi_t^{a^*,b^*}$ of the measure corresponding to the optimal a^* and b^*. $\xi_t^{a,b}$ is defined in (20).

We will start with deriving the explicit stochastic differential equation of the pricing kernel. First we calculate the derivatives of $f(C_t, V_t)$ with respect to C_t and V_t:

$$\begin{aligned} f_C(C,V) &= \frac{\beta C^{-1/\epsilon}}{\big((1-\gamma)V\big)^{\frac{1}{\zeta}-1}}, \\ f_V(C,V) &= \beta\zeta\Big\{\Big(1-\frac{1}{\zeta}\Big)\Big((1-\gamma)V\Big)^{-\frac{1}{\zeta}} C^{1-1/\epsilon} - 1\Big\}. \end{aligned} \tag{52}$$

Substituting $V_t = g(X_t)\frac{C_t^{1-\gamma}}{1-\gamma}$ into $f_C(C_t, V_t)$ and $f_V(C_t, V_t)$ we obtain:

$$\begin{aligned} f_C(C_t,V_t) &= \beta g(X_t)^{1-\frac{1}{\zeta}} C_t^{-\gamma}, \\ f_V(C_t,V_t) &= \beta\zeta\Big\{g(X_t)^{-\frac{1}{\zeta}}\Big(1-\frac{1}{\zeta}\Big) - 1\Big\}. \end{aligned} \tag{53}$$

This gives:

$$\pi_t = \xi_t^{a^*,b^*} \exp\left(\int_0^t \beta\zeta\Big(g(X_s)^{-\frac{1}{\zeta}}\big(1-\frac{1}{\zeta}\big) - 1\Big)ds \right)\beta g(X_t)^{1-\frac{1}{\zeta}} C_t^{-\gamma}. \tag{54}$$

Take the logarithm and write as a differential equation:

$$\begin{aligned} d\log(\pi_t) = {} & \beta\zeta\Big(g(X_t)^{-\frac{1}{\zeta}}\big(1-\frac{1}{\zeta}\big) - 1\Big)dt - \gamma d\log(C_t) + d\log(\xi_t^{a^*,b^*}) \\ & + (1-\frac{1}{\zeta})d\log\big(g(X_t)\big). \end{aligned} \tag{55}$$

Apply Ito's lemma to $\log(C_t)$, $\log(\xi_t^{a^*,b^*})$ and $\log\big(g(X_t)\big)$ and substitute the results; this leads to the following differential equation:

$$\begin{aligned} d\log(\pi_t) = {} & \Big\{\beta\zeta\Big(g(X_t)^{-\frac{1}{\zeta}}\big(1-\frac{1}{\zeta}\big) - 1\Big) - \gamma\big(\mu - \frac{\sigma^2}{2}\big) + \lambda_t(1-a^*) \\ & + (1/\epsilon - \gamma)\frac{g_X(X_t)}{g(X_t)(1-\gamma)}\mu_X(X_t)\Big\}dt \\ & - \gamma\sigma dZ_t + \Big(\log(a^*b^*) + \big((b^*-1)\eta - \gamma\big)\log(1+J_t)\Big)dN_t. \end{aligned} \tag{56}$$

After applying Ito's lemma to $\log(\pi_t)$ we find:

$$
\begin{aligned}
d\pi_t =& \Big\{\beta\zeta\Big(g(X_t)^{-\frac{1}{\zeta}}(1-\frac{1}{\zeta})-1\Big)-\gamma\big(\mu-(\gamma+1)\frac{\sigma^2}{2}\big)+\lambda_t(1-a^*) \\
&+(1/\epsilon-\gamma)\frac{g_X(X_t)}{g(X_t)(1-\gamma)}\mu_X(X_t)\Big\}\pi_t dt + -\gamma\sigma\pi_t dZ_t \\
&+\Big(a^*b^*(1+J_t)^{(b^*-1)\eta-\gamma}-1\Big)\pi_{t-}dN_t.
\end{aligned} \tag{57}
$$

We can now substitute the HJB equation (40) into the pricing kernel. Several terms cancel out and we are left with:

$$
\begin{aligned}
d\pi_t =& \Big\{-\beta-\frac{\mu}{\epsilon}+\Big(1+\frac{1}{\epsilon}\Big)\frac{\gamma}{2}\sigma^2+\Big(\gamma-\frac{1}{\epsilon}\Big)\lambda_t^*\frac{-1}{b^*\eta+1-\gamma} \\
&+\lambda_t(1-a^*)\Big\}\pi_t dt-\gamma\sigma\pi_t dZ_t \\
&+\Big(a^*b^*(1+J_t)^{(b^*-1)\eta-\gamma}-1\Big)\pi_{t-}dN_t.
\end{aligned} \tag{58}
$$

D.2 The interest rate

Let B_t be the price of a risk-free asset with a return equal to the interest rate. By the no-arbitrage argument, the interest rate r_t should be such that $\pi_t B_t$ is a martingale, . Now write $d\pi_t = \mu_{\pi,t}\pi_t dt + \sigma_\pi \pi_t dZ_t + J_{\pi,t}\pi_{t-}dN_t$. The product with B_t then follows:

$$
d\pi_t B_t = (r_t+\mu_{\pi,t})\pi_t B_t dt + \sigma_\pi \pi_t B_t dZ_t + J_{\pi,t}\pi_{t-}B_t dN_t. \tag{59}
$$

This is a martingale if $r_t+\mu_\pi+\lambda_t E_t[J_{\pi,t}] = r_t+\mu_\pi+\lambda_t\Big(a^*\frac{b^*\eta}{b^*\eta-\gamma}-1\Big)=0$. Therefore the equilibrium interest rate equals:

$$
\begin{aligned}
r_t =& -\mu_\pi-\lambda_t\Big(a^*\frac{b^*\eta}{b^*\eta-\gamma}-1\Big) \\
=& \beta+\frac{\mu}{\epsilon}-\Big(1+\frac{1}{\epsilon}\Big)\frac{\gamma}{2}\sigma^2-\Big(\gamma-\frac{1}{\epsilon}\Big)a^*\lambda_t\frac{-1}{b^*\eta+1-\gamma} \\
&-a^*\lambda_t\Big(\frac{b^*\eta}{b^*\eta-\gamma}-1\Big).
\end{aligned} \tag{60}
$$

Substituting r_t into the pricing kernel gives:

$$
\begin{aligned}
d\pi_t =& \Big\{-r_t-\lambda_t\Big(a^*\frac{b^*\eta}{b^*\eta-\gamma}-1\Big)\Big\}\pi_t dt-\gamma\sigma\pi_t dZ_t \\
&+\Big(a^*b^*(1+J_t)^{(b^*-1)\eta-\gamma}-1\Big)\pi_{t-}dN_t.
\end{aligned} \tag{61}
$$

D.3 The equity premium

Consider a stock that pays continuous dividends at a rate C_t and has ex-dividend price S_t. We denote the cum-dividend stock price by S_t^d. We use the expression for the consumption-wealth ratio in combination with the HJB-equation to derive the risk premium. An alternative derivation is to apply the no arbitrage condition. Using equation (48) we can write $S_t = \frac{C_t}{k(X_t)}$. The stock price then follows:

$$\begin{aligned} dS_t^d = dS_t + C_t dt &= \frac{1}{k(X_t)} dC_t - \frac{C_t}{k(X_t)^2} dk(X_t) + k(X_t) S_t dt \\ &= \Big(\mu - \frac{k_X(X_t)}{k(X_t)} \mu_X(X_t) + k(X_t)\Big) S_t dt + \sigma S_t dZ_t + J_t S_{t-} dN_t. \end{aligned} \tag{62}$$

From equation (62), we know that the drift of the stock equals $\mu_{S,t} = \mu - \frac{k_X(X_t)}{k(X_t)}\mu_X(X_t) + k(X_t)$. From (51) we have: $k(X_t) = \beta g(X_t)^{-\frac{1-1/\epsilon}{1-\gamma}}$. This gives: $\frac{k_X(X_t)}{k(X_t)} = -\frac{1-1/\epsilon}{1-\gamma}\frac{g_X(X_t)}{g(X_t)}$. Rewriting the HJB equation (40) gives:

$$\begin{aligned} \frac{1-1/\epsilon}{1-\gamma}\frac{g_X(X_t)}{g(X_t)}\mu_X(X_t) + k(X_t) = \beta + (1/\epsilon - 1)\Big(\mu - \frac{\gamma}{2}\sigma^2 \\ + a^* \lambda_t \frac{-1}{b^*\eta + 1 - \gamma}\Big). \end{aligned} \tag{63}$$

Substituting this into $\mu_{S,t}$ gives:

$$\begin{aligned} \mu_{S,t} &= \mu - \frac{k_X(X_t)}{k(X_t)}\mu_X(X_t) + k(X_t) \\ &= \mu + \beta + (1/\epsilon - 1)\Big(\mu - \frac{\gamma}{2}\sigma^2 + a^*\lambda_t \frac{-1}{b^*\eta + 1 - \gamma}\Big). \end{aligned} \tag{64}$$

The risk premium is then equal to the excess return of the stock over the interest rate:

$$\begin{aligned} rp_t &= \mu_{S,t} + a^*\lambda_t \frac{-1}{b^*\eta + 1} - r_t \\ &= \gamma\sigma^2 + a^*\lambda_t\Big(\frac{-1}{b^*\eta + 1} - \frac{b^*\eta}{b^*\eta + 1 - \gamma} + \frac{b^*\eta}{b^*\eta - \gamma}\Big). \end{aligned} \tag{65}$$

D.4 Consumption strips

Let $H_t = H(C_t, X_t, s-t) = E_t\Big[\frac{\pi_s}{\pi_t} C_s\Big]$ be the price of an asset that pays out the aggregate consumption at time s. H_t is also called a consumption strip. Conjecture that $H(C_t, X_t, u) = \exp\Big\{-\int_t^{t+u} CDR_s ds\Big\} C_t$. u denotes the time to maturity of the consumption strip. Clearly, $H(C_t, X_t, 0) = C_t$. Applying Ito's lemma to H_t gives:

$$\begin{aligned} dH_t = H_C dC_t + H_X dX_t - \frac{\partial H_t}{\partial u} dt = \frac{1}{C_t} H_t dC_t \\ - \frac{\partial}{\partial X_t}\Big(\int_t^{t+u} CDR_s ds\Big)\mu_X(X_t) H_t dt \\ + \frac{\partial}{\partial u}\Big(\int_t^{t+u} CDR_s ds\Big) H_t dt. \end{aligned} \tag{66}$$

We can calculate both derivatives:

$$
\begin{aligned}
&\frac{\partial}{\partial X_t}\Big(\int_t^{t+u} CDR_s ds\Big)\mu_X(X_t) = \frac{\partial}{\partial t}\Big(\int_t^{t+u} CDR_s ds\Big)\frac{\partial t}{\partial X_t}\mu_X(X_t) \\
&= \frac{\partial}{\partial t}\Big(\int_t^{t+u} CDR_s ds\Big) = CDR_{t+u} - CDR_t,
\end{aligned} \tag{67}
$$

$$
\frac{\partial}{\partial u}\Big(\int_t^{t+u} CDR_s ds\Big) = CDR_{t+u}. \tag{68}
$$

Therefore dH_t becomes:

$$
dH_t = \Big(\mu + CDR_t\Big)H_t dt + \sigma H_t dZ_t + J_t H_{t-} dN_t. \tag{69}
$$

Now define $dH_t = \mu_{H,t} H_t dt + \sigma H_t dZ_t + J_t H_{t-} dN_t$. By the no arbitrage condition, $\pi_t H_t$ must be a martingale:

$$
\begin{aligned}
d\pi_t H_t = &(\mu_{\pi,t} + \mu_H + \sigma\sigma_\pi)\pi_t H_t dt + (\sigma + \sigma_\pi)\pi_t H_t dZ_t \\
&+ \Big((1 + J_t)(1 + J_{\pi,t}) - 1\Big)\pi_{t-} H_{t-} dN_t.
\end{aligned} \tag{70}
$$

We can calculate the expectation of the jump term:

$$
\begin{aligned}
&E_t[(1 + J_t)(1 + J_{\pi,t}) - 1] = E_t[a^* b^* (1 + J_t)^{(b^*-1)\eta+1-\gamma} - 1] \\
&= a^* \frac{b^*\eta}{b^*\eta + 1 - \gamma} - 1.
\end{aligned} \tag{71}
$$

Therefore $\pi_t H_t$ is a martingale if:

$$
0 = \mu_\pi + \mu_H + \sigma\sigma_\pi + \lambda_t\Big(a^* \frac{b^*\eta}{b^*\eta + 1 - \gamma} - 1\Big). \tag{72}
$$

Substituting μ_π, μ_H and $\sigma\sigma_\pi = -\gamma\sigma^2$ gives:

$$
\begin{aligned}
0 = &\mu + CDR_t - r_t - \lambda_t\Big(a^* \frac{b^*\eta}{b^*\eta - \gamma} - 1\Big) - \gamma\sigma^2 \\
&+ \lambda_t\Big(a^* \frac{b^*\eta}{b^*\eta + 1 - \gamma} - 1\Big).
\end{aligned} \tag{73}
$$

Note that this implies that: $CDR_t = r_t + rp_t - (\mu + a^*\lambda_t \frac{-1}{b^*\eta+1})$. Lastly, we can substitute r_t and rp_t, which yields:

$$
CDR_t = \beta + (1/\epsilon - 1)\Big(\mu - \frac{\gamma}{2}\sigma^2 + a^*\lambda_t \frac{-1}{b^*\eta + 1 - \gamma}\Big). \tag{74}
$$

E The Social Cost of Carbon

The Social Cost of Carbon is calculated as the derivative of the value function with respect to carbon emissions, scaled by instantaneous marginal utility. With a single carbon box, the marginal cost of increasing carbon emissions by one unit is the derivative of the value function with respect to the carbon concentration M_t: $\frac{\partial V_t}{\partial M_t}$. However, with multiple carbon boxes, emitting one unit of carbon leads to an increase of ν_i units in box i, $i = 0, 1, 2, 3$. We slightly abuse notation and define $\frac{\partial}{\partial M_t} \equiv \nu_0 \frac{\partial}{\partial M_{0,t}} + \nu_1 \frac{\partial}{\partial M_{1,t}} + \nu_2 \frac{\partial}{\partial M_{2,t}} + \nu_3 \frac{\partial}{\partial M_{3,t}}$. Differentiation of the value function gives:

$$\begin{aligned} SCC_t &= -\frac{\partial V_t / \partial M_t}{f_C(C_t, V_t)} = -\frac{\frac{\partial}{\partial M_t} g(X_t)}{(1-\gamma) g(X_t) k(X_t)} C_t = -\frac{\frac{\partial}{\partial M_t} \left(\frac{k(X_t)}{\beta}\right)^{-\frac{1-\gamma}{1-1/\epsilon}}}{(1-\gamma)\left(\frac{k(X_t)}{\beta}\right)^{-\frac{1-\gamma}{1-1/\epsilon}} k(X_t)} C_t \\ &= -\frac{C_t}{1/\epsilon - 1} \frac{\frac{\partial}{\partial M_t} k(X_t)}{k(X_t)^2} = \frac{C_t}{1/\epsilon - 1} \frac{\partial}{\partial M_t} \int_0^\infty \exp\Big\{ -\int_t^{t+u} CDR_s ds \Big\} du \\ &= C_t \int_0^\infty \exp\Big\{ -\int_t^{t+u} CDR_s ds \Big\} \int_t^{t+u} a^* \lambda_T \frac{\partial T_s}{\partial M_t} ds \frac{1}{b^* \eta + 1 - \gamma} du \end{aligned} \tag{75}$$

F Calibration of Climate model

Table 4: Parameters for the Climate model

Par.	Description	Value
E_0	Initial level of total emissions (in GtC, 2015)	10.45
g_0^E	Initial growth rate of emissions (2015)	0.017
g_∞^E	Long-run growth rate of emissions	-0.02
δ_{g^E}	Speed of convergence of growth rate of emissions	0.0075
M_0	Initial carbon concentration compared to pre-industrial (in GtC, 2015)	263
M_{pre}	Pre-industrial atmospheric carbon concentration (in GtC)	588
$M_{0,0}$	Initial carbon concentration box 0 (in GtC, 2015)	139
$M_{1,0}$	Initial carbon concentration box 1 (in GtC, 2015)	90
$M_{2,0}$	Initial carbon concentration box 2 (in GtC, 2015)	29
$M_{3,0}$	Initial carbon concentration box 3 (in GtC, 2015)	4
$\delta_{M,0}$	Decay rate of carbon box 0	0
$\delta_{M,1}$	Decay rate of carbon box 1	0.0025
$\delta_{M,2}$	Decay rate of carbon box 2	0.027
$\delta_{M,3}$	Decay rate of carbon box 3	0.23
ν_0	Fraction of emissions carbon box 0	0.217
ν_1	Fraction of emissions carbon box 1	0.224
ν_2	Fraction of emissions carbon box 2	0.282
ν_3	Fraction of emissions carbon box 3	0.276
F_0^E	Initial level of exogenous forcing (in W/m^2, 2015)	0.5
F_∞^E	Long-run level of exogenous forcing (in W/m^2)	1
δ_F	Speed of convergence exogenous forcing	0.02
T_0	Initial surface temperature compared to pre-industrial (in $°C$, 2015)	0.85
T_0^{oc}	Initial ocean temperature compared to pre-industrial (in $°C$, 2015)	0.0068
κ	Speed of temperature transfer between upper and deep ocean	0.73
υ	Equilibrium temperature response to radiative forcing	1.13
α	Equilibrium temperature impact of CO_2 doubling (in $°C$)	3.05
τ	Heat capacity of the surface	7.34
τ_{oc}	Heat capacity of the oceans	105.5

G Stochastic Emissions

The HJB-equation for this problem becomes:

$$
\begin{aligned}
0 = \min_{(a,b)\ s.t.\ d(a,b)\leq\theta} & \Big\{ f(C_t, V_t^{\mathbb{Q}}) + V_C^{\mathbb{Q}} \mu C_t dt + \frac{1}{2} V_{CC}^{\mathbb{Q}} \sigma^2 C_t^2 + V_X^{\mathbb{Q}} \mu_X(X_t, C_t) \\
& + \lambda_t^{\mathbb{Q}} E_t^{\mathbb{Q}} \big[V^{\mathbb{Q}}\big((1+J_t)C_{t-}, X_t\big) - V^{\mathbb{Q}}(C_{t-}, X_t) \big] \Big\}.
\end{aligned} \tag{76}
$$

The main difference with the HJB-equation without stochastic emissions is that now, the drift of the state variables μ_X also depends on aggregate endowment C_t. It is therefore not possible anymore to substitute out the variable C_t. We thus have to solve a seven dimensional model numerically. We use the stochastic grid method to numerically solve the model, as described in Olijslagers (2021). Similar to value function iteration, the time step is discretized and the problem is solved backwards. The stochastic grid method simulates random grid points every time period and uses regressions with basis functions to approximate the value function. The main advantage is that this method can handle high-dimensional problems while avoiding the curse of dimensionality (computing time growing exponentially along with dimensionality) and that derivatives of the value function can be calculated easily. This is useful to calculate the social cost of carbon, and also to solve the first order conditions.

The first order conditions for optimal consumption are:

$$
\begin{aligned}
& E_t^{\mathbb{Q}} \big[V^{\mathbb{Q}}\big((1+J_t)C_{t-}, X_t\big) - V^{\mathbb{Q}}(C_{t-}, X_t) \big] - l_t \Big(\log(ab) + \frac{1}{b} - 1 \Big) = 0, \\
& a \frac{\partial E_t^{\mathbb{Q}} \big[V^{\mathbb{Q}}\big((1+J_t)C_{t-}, X_t\big) \big]}{\partial b_t} - l_t a \frac{b-1}{b^2} = 0, \\
& \theta - (1-a) - a \Big(\log(ab) + \frac{1}{b} - 1 \Big) = 0.
\end{aligned} \tag{77}
$$