

TI 2019-030/VI  
Tinbergen Institute Discussion Paper

# Discounting the Future: on Climate Change, Ambiguity Aversion and Epstein-Zin Preferences

*Stan Olijslagers<sup>1</sup>*  
*Sweder van Wijnbergen<sup>1,2</sup>*

<sup>1</sup> UvA

<sup>2</sup> CDPR, TI

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: [discussionpapers@tinbergen.nl](mailto:discussionpapers@tinbergen.nl)

More TI discussion papers can be downloaded at <https://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam  
Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900

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

Stan Olijslagers<sup>†</sup>      Sweder van Wijnbergen<sup>‡</sup>

April 30, 2019

## Abstract

We focus on the effect of preference specifications on the current day valuation of future outcomes. Specifically, we analyze the effect of risk aversion, ambiguity aversion and the elasticity of intertemporal substitution on the willingness to pay to avoid climate change risk. The first part of the paper analyzes a general disaster (jump) risk model with a constant arrival rate of disasters. This provides useful intuition in how preferences influence valuation of long-term risk. The second part of the paper extends this model with a climate model and a temperature dependent arrival rate. Since the model yields closed form solutions up to solving an integral, our model does not suffer from the curse of dimensionality of numerical IAMs with several state variables. Introducing Epstein-Zin preferences with an elasticity of substitution higher than one and ambiguity aversion leads to much larger estimates of the social cost of carbon than obtained under power utility. The dominant parameters are the risk aversion coefficient and the elasticity of intertemporal substitution. Ambiguity aversion is of second order importance.

JEL codes: Q51, Q54, G12, G13

Keywords: Social Cost of Carbon, Ambiguity Aversion, Epstein-Zin preferences, Stochastic Differential Utility, Climate Change

---

\*We thank Bob Pindyck, Nick Stern and Martin Weitzman for useful comments. We are also grateful for financial support from pension administrator and wealth manager MN.

<sup>†</sup>University of Amsterdam, [s.w.j.olijslagers@uva.nl](mailto:s.w.j.olijslagers@uva.nl)

<sup>‡</sup>University of Amsterdam, Tinbergen Institute, CEPR, [s.j.g.vanwijnbergen@uva.nl](mailto:s.j.g.vanwijnbergen@uva.nl)

## Executive summary

The consequences of climate risk will take place far in the future while if they are to be avoided policies need to be implemented today. This has made the issue of discounting future costs of climate change back towards today arguably the most important element of the climate change debate and that is the subject of this paper.

Climate disasters are events that have a large impact on the economy, occur very rarely and take place abruptly (Goosse, 2015). To model this feature, we add a jump (Poisson) process to the endowment consumption stream in an otherwise standard Lucas tree General Equilibrium asset pricing model. In the second part of this paper we extend this model by adding an explicit climate dynamics model relating the jump risk to temperature rises coming out of the climate model. We model the climate dynamics similar to the DICE model (W. Nordhaus, 2014).

In the first part of this paper (Sections 4, 5 and 6) we study an asset pricing model with constant disaster risk to bring the link between risk and the structure of preferences in sharp focus. The economy is modeled as a pure exchange economy with exogenous endowments. When a disaster arrives, a random variable determines the damage of the climate disaster. We focus explicitly on risk until the last section of the second paper where we add expected losses. But before that section we assume that climate disasters do not decrease economic growth on average so as to completely focus on the risk side of the problem.

Because these jumps have not actually occurred already, it is hard to estimate the probability that a disaster occurs or its expected impact. Therefore we assume that the agent does not know the exact probability distributions of the arrival rate of climate shocks and the size of the shocks. There is so called ambiguity (or Knightian uncertainty) about the characteristics of the jump risk component while the agent is assumed to be averse to this ambiguity. Moreover we use the continuous time version of Epstein-Zin utility (Stochastic Differential Utility), which allows us to separate the intertemporal elasticity of substitution from the degree of risk aversion so we can disentangle risk aversion effects (known probabilities), ambiguity aversion effects (unknown probabilities) and substitution effects.

Solving the model yields closed form solutions for the interest rate, risk premium and consumption-wealth ratio. These formulas give very useful insights in how preferences influence the key characteristics of the model. It is especially interesting to see how jump risk in combination with non-expected utility yields much more realistic interest rates in a general equilibrium model. The interest rate plays an important role in discounting the future. We are also able to calculate in closed form how much time 0 consumption an agent is willing to give up to reduce climate risk. This amount is very sensitive to both ambiguity aversion and the elasticity of intertemporal substitution (EIS).

In the second part of the paper we extend the model by introducing a climate model and a temperature-dependent disaster arrival rate. In this model we can actually calculate the social cost of carbon. The social cost of carbon measures the cumulative future damage in dollar terms of emitting one ton of carbon today. We derive an analytic expression for the social cost of carbon (SCC) in the climate risk

extended variant where jump risks depend on temperature (Sections 7 and 8). The SCC is represented by an integral over all future expected damages. The integrand consists of two parts: a weighting function that determines how much weight one puts on the future vs. today and a certainty equivalent. We show that risk aversion and the elasticity of intertemporal substitution are the dominant preference parameters in determining the SCC. Risk aversion increases the SCC through the certainty equivalent: a more risk-averse agent dislikes risk more and therefore has a lower certainty equivalent. The elasticity of intertemporal substitution affects the weighting function: it puts more weight on the (far) future and these are exactly the dates where expected climate damages are large. Therefore a larger EIS yields a larger SCC.

Ambiguity aversion has an effect on both the certainty equivalent and the weighting function. Increasing ambiguity aversion always leads to a lower certainty equivalent, which would in turn leads to a higher SCC. But the effect of ambiguity aversion on the weighting function depends on the elasticity of intertemporal substitution. When the EIS is smaller than one, income effects dominate and increasing ambiguity aversion leads to more weight on future dates. The alternative is true for values of the EIS above one. In this case, a high ambiguity aversion leads to more weight on the short-term. The certainty equivalent effect and the weighting effect work in opposite directions when the EIS is larger than one. These results demonstrate the insights to be gained from deriving analytical solutions to the model, the literature has so far largely used computer simulations of special cases or analytical solutions using logarithmic utility ( $EIS = 1$ ).

Our numerical example using best estimates of the various parameters, ambiguity aversion and especially preferences for early resolution of uncertainty yields a SCC that is an order of magnitude larger than conventional approaches yields, namely a SCC of \$1256 per ton of carbon (\$342 per ton of  $CO_2$ ). This result is mainly driven by a higher elasticity of intertemporal substitution (1.5) compared to power utility, and by the fact that climate disasters are stochastic and therefore risk aversion (5) plays an important role. Ambiguity aversion does play a role but the effect is small.

As mentioned before, the paper focuses specifically on risk. We have assumed that on average, climate risk has no impact on economic growth and risk-neutrality would imply a SCC of zero. In reality, climate change increases risk but also decreases economic growth. Relaxing the assumption that climate change has on average no effect on economic growth and including the expected loss, we obtain a much higher estimate of the SCC. With the same calibration as the previous estimate, the SCC almost doubles to \$2178 per ton of carbon (\$593 per ton of  $CO_2$ ).

# 1 Introduction

Climate change is one of the main risks the economy will face in the upcoming decades or possibly even centuries. However, there is still much uncertainty about climate change. While (almost) all scientists agree on the fact that climate change will have possibly dramatic negative consequences for the environment and economic growth, we are still not able to accurately estimate the extent and timing of future damages induced by climate change. But one thing we do know is that such consequences will take place far in the future while if they are to be avoided policies need to be implemented today. This has made the issue of discounting future costs of climate change back towards today arguably the most important element of the climate change debate and that is the subject of this paper. But 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<sup>1</sup>. Another extension in comparison to the traditional approaches is that we model climate risk explicitly as disaster risk. We show that implementing empirically supported deviations from time separable power utility and modeling climate damages as rare disasters leads to estimates of the Social Cost of Carbon that are orders of magnitude larger than have been derived so far using conventional approaches to time discounting.

The impact of climate risk on the economy is often modeled using Integrated Assessment Models (IAMs). IAMs integrate the knowledge of different domains into one model. In the case of climate change, the models combine economic factors (economic growth, technological improvement, fossil fuel use), demographic changes (population growth) and climate factors (atmospheric carbon concentration, temperature, ecosystem)<sup>2</sup>. The three main IAMs are DICE (W. Nordhaus, 2014), PAGE (Hope, 2006) and FUND (Tol, 2002)<sup>3</sup>. 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, and can fool policy-makers into thinking that the forecasts the models generate have some kind of scientific legitimacy.*" 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 catastrophic outcomes of climate change, instead of modeling the underlying causes.

---

<sup>1</sup>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 Chichilnisky, Hammond, and Stern (2018); Stern (2015) who look at a positive rate of time preference as discrimination between generations that happen to have been born at different moments in time.

<sup>2</sup>This list is not exhaustive, but is used to give an idea of the nature of an IAM.

<sup>3</sup>The references do not contain the most recent versions of the IAMs.

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 a Monte Carlo-like approach by analyzing several deterministic runs with different parameter values and then taking a weighted average of all runs (W. D. Nordhaus (2014); Dietz (2011)). 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 variables, since for each run all uncertainty is resolved at time 0. Crost and Traeger (2013) compare the Monte Carlo approach to a model that actually uses random variables and find that the Monte Carlo approach underestimates the impact of climate damages.

We propose an analytically solvable IAM (Integrated Assessment Model) that addresses the critiques of Pindyck (2017) and Crost and Traeger (2013) on the use of deterministic IAMs. This IAM is able to capture the climate dynamics of the numerical IAMs very well. 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 tail risk instead of assuming that temperature increases generate a certain amount of damage every year. The model is transparent due to the closed form solutions for the social cost of carbon. Where stochastic numerical IAMs can take hours to be solved, solving this model only requires numerical integration and is therefore solved within seconds.

The paper is divided into two parts. The first part studies an asset pricing model with constant disaster risk. The economy is modeled as a pure exchange economy with exogenous endowments. We extend the Consumption-based Capital Asset Pricing Model (CCAPM) developed in Lucas Jr (1978). In the literature, this model is widely used in conjunction with a lognormal distribution<sup>4</sup>. The diffusion component of the endowment captures fluctuations in consumption. 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 have a large impact on the economy, occur very rarely and take place abruptly (Goosse, 2015). To model this feature, we add a jump process to the endowment consumption stream to capture the disaster risk, explicitly modeling climate risk as tail risk.

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 shocks and the size of the shocks. 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, also called stochastic differential utility (SDU), 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 in-

---

<sup>4</sup>Although Lucas Jr (1978) doesn't assume a specific distribution for the endowment stream.

verse. 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 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. The SDU 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 risk.

The asset pricing model with constant disaster risk gives a very useful insight in how our preference and risk specifications affect prices, interest rates and valuation of risks. In the second part of the paper we extend this asset pricing model with a climate model. We model carbon emissions, atmospheric carbon concentration and temperature increases. The arrival rate of climate disasters is now increasing in temperature instead of constant. Our analytic model is able to match the climate dynamics of numerical IAMs and specifically matches the DICE model's dynamics very closely. Furthermore we explicitly focus on the valuation of climate risk and we therefore do not analyse optimal abatement policies at this stage yet. A commonly used measure for the cost of carbon emissions is the social cost of carbon (SCC), the long term damage in dollar terms of emitting one ton of carbon today. Our base calibration yields a sizable social cost of carbon. Similar to the numerical IAMs, the SCC in our model is also very sensitive to the choice of the input parameters but in addition we can explore the implications of ambiguity aversion, preferences for early resolution of uncertainty and (related to that) a higher EIS than implied by commonly accepted values for the degree of risk aversion. 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. In our numerical example using best estimates of the various parameters, ambiguity aversion and especially preferences for early resolution of uncertainty yield a SCC that is an order of magnitude larger than conventional approaches yield.

## 2 Related literature

This paper is related to two strands of literature. First, our methodology is related to portfolio choice and consumption based asset pricing models with jump risk and/or non-expected utility. And second, 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 Capital Asset Pricing Model (CCAPM) by Lucas Jr (1978). Breeden (1979) derives a similar model in



continuous time. [Naik and Lee \(1990\)](#) consider the pricing of options in a general equilibrium consumption-based framework where asset prices follow a jump-diffusion process. [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. These puzzles are called respectively the Equity Premium Puzzle and the Risk-Free Rate Puzzle. Jump risk or disaster risk has been proposed as a possible solution of these puzzles ([Rietz \(1988\)](#); [Barro \(2006\)](#); [Pindyck and Wang \(2013\)](#)). Extensions to the early disaster/jump risk models are the use of Stochastic Differential Utility (SDU) instead of power utility, and the introduction of time-varying disaster probabilities and multi-period (i.e. persistent) disasters ([Barro \(2009\)](#); [Wachter \(2013\)](#); [Tsai and Wachter \(2015\)](#)). Climate change induced disasters fit in the rare disaster literature since climate change is increasingly thought to give rise to large and abrupt destructive changes in the Earth’s environment ([Goosse, 2015](#)) whose occurrence has a small probability at any given moment of time but with possibly large negative effects on the economy.

Ambiguity aversion, aversion of unmeasurable or Knightian uncertainty, is the second extension of the CCAPM we introduce to our climate model. [Brock and Hansen \(2017\)](#) stress the importance of taking into account risk, ambiguity and model misspecification in climate economics. [Liu, Pan, and Wang \(2004\)](#) consider a general equilibrium model with rare disasters and ambiguity aversion<sup>5</sup>. The agent is only concerned about misspecification of the jump process, a logical choice that we follow, since the probability distribution of rare events is by their very nature much harder to estimate compared to the diffusion component. Other papers on partial equilibrium portfolio choice models with ambiguity aversion and jump-diffusion processes are [Jin, Luo, and Zeng \(2017\)](#) and [Branger and Larsen \(2013\)](#). The first paper uses a non-parametric approach to model ambiguity aversion in the jump component, the second paper focuses on utility losses if ambiguity is ignored.

Our stochastic and preference structure combines the model of [Liu et al. \(2004\)](#) with ambiguity aversion with respect to the parameters of the jump-risk component, with the i.i.d. model of [Tsai and Wachter \(2015\)](#). This implies that the agent has SDU preferences, the stochastic consumption process follows a jump diffusion process and the representative agent is ambiguity averse with respect to the arrival rate and the jump probability of the jump-risk component.

Risk aversion and ambiguity aversion are obviously important in a climate change setting, but since abrupt climate change is anticipated to take place far into the future, intertemporal choices play an important role as well. Power utility is therefore an unsatisfactory framework since with that structure of preferences, risk aversion and IES cannot be varied independently. This is the reason why we adopt the Stochastic Differential Utility framework introduced by [Duffie and Epstein \(1992b\)](#) since with SDU the risk aversion parameter and the elasticity of intertemporal substitution (EIS) are no longer restricted to be each other’s inverse. Furthermore, we go beyond the setting of [Tsai and Wachter \(2015\)](#) by introducing ambiguity aversion. This extension

---

<sup>5</sup>[Liu et al. \(2004\)](#) looked at the implication of disaster risk for option pricing and show how the model can generate option smirks that are observed in the data.

can also be relevant in the general disaster model, but is especially relevant in the climate disaster model since there is no clear history of events on which we can base our estimates of the damages.

The second strand in the literature our paper is related to is the literature on climate change economics and especially to the part that considers non-expected utility and the part that looks at analytic approaches to solve their models. [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 the consumption growth contains 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 show that the outcomes of their model are very sensitive to choices of the EIS. Non-expected utility can have a substantial effect on the discount rate, which is obviously a very important parameter in climate models since climate change is expected far into the future but countermeasures may well have to be taken today. Of course it is not just the structure of preferences but also the pure rate of time preference that has an impact on the discount rate. There is a large debate on the choice of the pure rate of time preference in climate models ([Stern et al. \(2006\)](#); [Chichilnisky et al. \(2018\)](#); [W. D. Nordhaus \(2007\)](#); [Weitzman \(2007\)](#)). [Traeger \(2014\)](#) explicitly considers the effect of Epstein-Zin preferences and ambiguity aversion on interest and discount rates. He models ambiguity aversion using the smooth recursive preferences approach by [Klibanoff, Marinacci, and Mukerji \(2005\)](#); we choose a different approach to modeling ambiguity aversion, which we discuss below in section [3.2](#).

The three IAMs that were mentioned in the introduction are all solved using numerical methods. However, since it has become clear that the choice of the input parameters has a large influence on the results, we think it is useful to know how these parameters exactly influence the outcomes and therefore opt for models allowing for analytical solutions. There are a few very recent papers that also focused on obtaining analytic solutions. [Golosov, Hassler, Krusell, and Tsyvinski \(2014\)](#) were the first to obtain closed form solutions in an IAM. However, this required quite strict assumptions such as logarithmic utility and full depreciation of capital every decade. [Bretscher and Vinogradova \(2018\)](#) develop a stylized production-based model where emissions directly enter the damage function and obtain closed form solutions for the optimal abatement policy. [Bremer and van der Ploeg \(2018\)](#) consider a rich 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.

The Analytic Climate Economy (ACE) model of [Traeger \(2018\)](#) is closest to our setup. He extends the model of [Golosov et al. \(2014\)](#). Where in other models either carbon emissions ([Bretscher & Vinogradova, 2018](#)) or atmospheric carbon concentration ([Golosov et al., 2014](#)) directly enter the damage function, ACE explicitly models the effect of carbon concentration on temperature while the damages are induced by an increasing temperature. The model is able to match the climate dynamics of the numerical IAM models. [Traeger \(2018\)](#) also considers the effect of uncertainty in

state variables. Our approach is closely related to ACE, we model the dynamics of our state variables in a similar way. However, we consider an endowment economy in continuous time instead of a production economy in discrete time because continuous time allows us to sharply distinguish abrupt jumps from smooth diffusion stochastics, a difference that is blurred in discrete time frameworks. Furthermore, in our framework damages enter the consumption process using a Poisson process (disaster risk), in line with the climate scientists' discussion of abrupt changes, where in ACE, damages enter by decreasing the output every time period, equivalent to through the drift term in a continuous time setting. But our main extension is that we obtain solutions with SDU preferences and ambiguity aversion, where ACE only solves the model for a specific case of Epstein-Zin utility, namely with a unit EIS. When the EIS is equal to one, the problem vastly simplifies since income and substitution effects cancel out. But that simplification comes at a cost: there is empirical evidence that the EIS is well above one (Van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramírez (2012); Vissing-Jørgensen and Attanasio (2003)), which as we will demonstrate, has a major impact on the SCC. Also in the finance literature (Bansal et al. (2016); Barro (2009)) and in climate change models with recursive preferences (Jensen and Traeger (2014); Cai, Judd, and Lontzek (2015)) values of the EIS between 1.5 and 2 are often used. Given the large impact of changes in the EIS given risk aversion on intertemporal choice it seems to us important to study how deviating from unit elasticity affects intertemporal choices and the long term stochastic discount factor.

### 3 Methodology

In this section we discuss several choices regarding the model setup and utility specification. In order to do so we first outline the setup for the economy. Since we do not consider mitigation policies in this paper, we opt for assuming a pure exchange economy, where agents are endowed with an exogenous stochastic income stream. Agents can buy risky stocks, which give a claim on the endowment. Consumption goods are perishable, transferring wealth to the future is only possible by buying stocks. The income stream can intuitively be seen as a tree that produces an uncertain amount of fruit every time period. All agents can buy stocks, which are shares in the tree. The fruit is non-storable, so it must be consumed at the period of the endowment. This implies that aggregate endowment equals aggregate consumption at every moment in time. It is assumed that all agents have identical preferences and endowments, so the separate agents can be replaced by one representative agent. That agent then wants to maximize utility from consumption.

The aggregate endowment process follows a geometric Brownian motion with an additional jump component that represents climate disasters. Suppose we have a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  on which a standard Brownian motion  $Z_t$ , a Poisson process  $N_t$  with arrival rate  $\lambda$  and a random variable  $Y_t$  are defined. Jump amplitudes are assumed to be independent over time. The three types of shocks, namely Brownian motions, Poisson arrivals and jump magnitudes, are assumed to be independent. Assume there is a filtration  $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$ . We will use the following notation:

$E_t[\cdot] = E[\cdot|\mathcal{F}_t]$ . Consider the following process for aggregate endowments:

$$dC_t = (\mu - \lambda m)C_t dt + \sigma C_t dZ_t + Y_t C_{t-} dN_t \quad (1)$$

The endowment follows the usual geometric Brownian motion dynamics, with an additional jump process.  $C_{t-}$  denotes aggregate consumption just before the jump ( $C_{t-} = \lim_{h \downarrow 0} C_{t-h}$ ). Clearly, in equilibrium aggregate consumption must equal aggregate endowment and therefore the process is also referred to as the aggregate consumption process. The growth rate  $\mu \geq 0$  and the volatility  $\sigma > 0$  are constant. When a jump arrives at time  $t$ , the jump size is controlled by the random variable  $Y_t$ . Assume that  $Y_t = e^{W_t} - 1$  where  $W_t$  follows a normal distribution with mean  $\mu_J < 0$  and variance  $\sigma_J$ .  $Y_t$  can be seen as the percentage change of aggregate consumption after a jump. The expected jump size equals  $E[Y_t] = m = e^{\mu_J + \frac{1}{2}\sigma_J^2} - 1$ . In line with the subject of climate disasters, we assume that  $m < 0$ . To compensate for the expected effect of the jumps, the term  $\lambda m$  is subtracted from the drift (i.e. we use a compensated compound Poisson process). We use a compensated compound Poisson process because this allows us to focus specifically on the welfare effects of risk reduction: once the correction term is included, the *expected* effect of a jump is zero. A similar assumption is made in the LRR model with climate change disasters (Bansal et al., 2016). This implies that climate disasters increase risk rather than decreasing expected aggregate consumption. It is relatively straightforward to later on add the expected value of the compensation terms to look at results without compensation. Although the focus of the paper is on risk, we will later on shortly look at the impact of leaving out the compensation term on the outcome.

### 3.1 Utility Specification

The representative agent maximizes his utility of consumption over an infinite planning horizon. We consider the continuous time version of the Epstein-Zin preferences (Epstein & Zin, 1989), first developed by Duffie and Epstein (1992b). Epstein and Zin (1989) consider the following class of preferences in discrete time:  $V_t = [(1 - \beta)C_t^{1-1/\epsilon} + \beta\mu_t(V_{t+1})^{1-1/\epsilon}]^{\frac{1}{1-1/\epsilon}}$  where  $\epsilon = EIS$ ,  $\beta$  is the time preference parameter and  $\mu_t(\cdot)$  is a certainty equivalent function. When considering a deterministic consumption program,  $V_t$  is a constant elasticity of substitution (CES) utility function. In the other extreme case where only a static gamble is considered, there are no intertemporal choices and the utility is entirely determined by the certainty equivalent function  $\mu_t(\cdot)$ . The certainty equivalent function (or risk aggregator) that is widely used throughout the literature is  $\mu_t(V_{t+1}) = E_t(V_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}$  where  $\gamma$  is the relative risk-aversion coefficient. This specification of  $\mu_t(\cdot)$  yields a special case of the preferences studied by Kreps and Porteus (1978) and is therefore also called Kreps-Porteus utility. Static gambles are evaluated as if the agent has power utility, but in a dynamic stochastic setting EIS and risk aversion are decoupled under SDU: this specification allows to separate risk aversion  $\gamma$  from the elasticity of intertemporal substitution  $\epsilon$ .

One property of this utility specification is that the agent has preferences for early resolution of uncertainty if  $\epsilon > \frac{1}{\gamma}$ . Consider a consumption program with consumption in two periods:  $c_0$  and  $c_1$ . In consumption program  $c^A$ , at time  $t_0$ , two coins (0 and 1) are tossed, one for each period. When a coin  $i$  yields a head, consumption at time  $i$  equals  $H$ . When a coin  $i$  yields a tail, consumption equals  $L < H$ . Consumption program  $c^B$  is similar, except for the timing of the coin tosses. Instead of tossing both coins at time  $t_0$ , now coin 0 is tossed at time  $t_0$  and coin 1 is tossed at  $t_1$ . So in terms of risk, both consumption programs are identical. When considering expected utility, both consumption programs have equal utility. However, when considering Epstein-Zin utility with  $\epsilon > \frac{1}{\gamma}$ , the agent has preferences for early resolution for uncertainty and prefers  $c^A$  over  $c^B$ . Knowing the outcomes earlier yields a higher utility. It must be noted that even when the agent is not able to improve planning when the outcomes are known earlier, he still prefers to know the outcomes earlier.

We consider a special case of SDU, namely the continuous time equivalent of Kreps-Porteus utility described above. In fact, instead of using Kreps-Porteus utility itself, we use an ordinally equivalent utility process. Similar to the discrete time case, SDU can be represented by a combination of an aggregator  $f$  that determines the degree of intertemporal substitution and a certainty equivalent operator  $m$ . In the case of Kreps-Porteus utility,  $f(C, V) = \frac{\beta}{1-1/\epsilon} \frac{C^{1-1/\epsilon} - V^{1-1/\epsilon}}{V^{-1/\epsilon}}$  and  $m(\sim V) = [E(V^{1-\gamma})]^{\frac{1}{1-\gamma}}$ . In this case the drift of the value function consists of the aggregator  $f(C, V)$  and a variance multiplier  $A$  that belongs to  $m$ . [Duffie and Epstein \(1992b\)](#) show that there exists an ordinally equivalent utility process with aggregator  $f$  as in [\(2\)](#). In this case  $m(\sim V) = E(V)$  and the variance multiplier  $A$  that belongs to  $m$  is zero. This yields a simplification and is easier to work with. For more details, we refer to [Duffie and Epstein \(1992b\)](#). The agent's utility or value function equals:

$$V_t = E_t \left[ \int_t^\infty f(C_s, V_s) ds \right]$$

where

$$f(C, V) = \frac{\beta}{1-1/\epsilon} \frac{C^{1-1/\epsilon} - ((1-\gamma)V)^{\frac{1}{\zeta}}}{((1-\gamma)V)^{\frac{1}{\zeta}-1}} \quad \text{for } \epsilon \neq 1 \quad (2)$$

with  $\zeta = \frac{1-\gamma}{1-1/\epsilon}$ .

Throughout this paper, we refer to this utility specification as stochastic differential utility (SDU) although [Duffie and Epstein \(1992b\)](#) actually consider a more general class of utilities under that label. Similar to the discrete time counterpart,  $\gamma$  denotes risk-aversion,  $\epsilon$  is the elasticity of intertemporal substitution and  $\beta$  equals the time preference parameter. We will focus on the case where  $\epsilon \neq 1$  and therefore will derive our results only for this case. For the case  $\epsilon = 1$  we can take the limit  $\epsilon \rightarrow 1$  or follow the same derivation but with  $f(C, V) = \beta(1-\gamma)V \left( \log C - \frac{1}{1-\gamma} \log((1-\gamma)V) \right)$ . If  $\gamma = \frac{1}{\epsilon}$ , the utility specification reduces to standard power utility.

## 3.2 Ambiguity

There is much uncertainty regarding the arrival rate and magnitude of climate disasters. And, as stressed by Pindyck (2017), we know very little about the damage functions. Where consumption growth and volatility can be estimated more accurately from historical data, the estimation of the jump parameters will be much harder since jumps do not happen that often. In fact in the context of climate change, abrupt changes have only taken place in the very far past (Goosse, 2015). It is fair to state that we simply do not know the exact distribution of climate damages. We consider it therefore desirable to 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 to certain 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 famous Ellsberg Paradox that people are ambiguity averse, i.e. they prefer known probabilities over unknown probabilities.

### 3.2.1 Overview of methods to model (parameter) ambiguity

There are several different approaches that are commonly used in the literature to model ambiguity about parameters. A widely used approach in the static setting is the maxmin approach of Gilboa and Schmeidler (1989). Assume the agent does not know the actual distribution of a certain random variable. The idea is to first specify a set of reasonable probability measures  $\mathbb{Q}$ . The agent is ambiguity averse and given this set of measures he considers the worst case measure. Utility is then of the form  $V_t = \min_{\mathbb{Q} \in \mathcal{P}} E_{\mathbb{Q}}[U(C_t)]$  for some utility specification  $U(\cdot)$ .

It is not straightforward to extend the Gilboa-Schmeidler maxmin preferences to a dynamic setting. We will discuss two approaches that have been proposed by Chen and Epstein (2002) and by Hansen, Sargent, Turmuhambetova, and Williams (2001) in the setting of our model.

Consider the agent’s problem. In the setting without ambiguity, the value function is given by:

$$V_t = E_t \left[ \int_t^\infty f(C_s, V_s) ds \right] \quad (3)$$

However, in the model with ambiguity the agent takes into account the fact that he is not certain about the true value of the arrival rate  $\lambda$  and the jump size  $m$ .

Hansen et al. (2001) propose two approaches to model ambiguity: the *multiplier* approach and *constraint* approach. We first consider the *multiplier* approach. The ‘best estimate’ model or reference model is the agent’s most reliable model with measure  $\mathbb{P}$ . But he also takes into account other, alternative models. The alternative models have measure  $\mathbb{Q}^{a,b}$ , the jump arrival rate becomes  $\lambda_t^{\mathbb{Q}} = e^{a_t} \lambda$  and the jump size parameter becomes  $m_t^{\mathbb{Q}} = e^{\mu_J + b_t \sigma_J^2 + \frac{1}{2} \sigma_J^2} - 1$ . Deviating from the reference model



is penalized since the agent does not choose the ‘best estimate’ model. The size of the penalty is proportional to  $d(a_t, b_t)$ , which represents the distance between the alternative model and the reference model. An alternative model that has a large distance to the reference model is considered less likely to be true and therefore using it receives a larger penalty. The distance function should satisfy  $d(a_t, b_t) \geq 0 \forall (a_t, b_t)$  and  $d(0, 0) = 0$ . Therefore using the reference model carries a zero penalty. The penalty is scaled by  $\theta$ , which is the ambiguity aversion parameter. This parameter controls the importance of the penalty term. Then the agent solves the following problem:

$$V_t = \min_{\{a_s, b_s\}_{s \geq t}} E_t^{\mathbb{Q}} \left[ \int_t^{\infty} \left( f(C_s, V_s) + e^{-\beta(s-t)} \theta d(a_s, b_s) \right) ds \right] \quad (4)$$

$E_t^{\mathbb{Q}}$  denotes the expectation under the alternative model with parameters  $\lambda_t^{\mathbb{Q}}$  and  $m_t^{\mathbb{Q}}$ . The expected utility of consumption is lower for high  $a_t$  and low  $b_t$ . We see that the agent faces a trade-off between how likely a combination of  $(a_s, b_s)$  is in terms of distance to the reference model and how bad it is in terms of expected utility. This trade-off results in optimal values of  $a_s$  and  $b_s$ .

The *constraint* approach is closely related to the *multiplier* approach. Instead of adding a penalty function, the agent can put a constraint on the distance function  $d(a_t, b_t)$ . Then the problem becomes:

$$\begin{aligned} V_t = & \min_{\{a_s, b_s\}_{s \geq t}} E_t^{\mathbb{Q}} \left[ \int_t^{\infty} f(C_s, V_s) ds \right] \\ \text{s.t. } & \int_t^{\infty} e^{-\beta(s-t)} d(a_s, b_s) ds \leq \eta \end{aligned} \quad (5)$$

So in this approach,  $\eta$  controls the size of the set of alternative models that seem reasonable to him.  $\eta$  can again be seen as an ambiguity aversion parameter. A high  $\eta$  implies a large set of priors and therefore corresponds to high ambiguity aversion. Given the constraint the worst-case model is chosen. This approach is very similar to the penalty approach and the two problems are related via Lagrangian optimization where  $\theta$  can be seen as the Lagrange multiplier.

[Hansen and Sargent \(2001\)](#) consider how the penalty and constraint approaches are related. They show that when the consumption process follows a pure geometric Brownian motion (i.e. no jumps), there exists an  $\eta$  for the constraint approach and a  $\theta$  for the multiplier approach such that both problems yield the same optimal outcome. The constraint approach is directly motivated from the Gilboa-Schmeidler maxmin utility. Since the multiplier approach is weakly related to the constraint approach, these approaches are indirectly also motivated by the static maxmin utility. Furthermore, the multiplier utility is axiomatized by [Strzalecki \(2011\)](#).

A disadvantage of both these approaches is that utility is not homothetic. [Maenhout \(2004\)](#) proposes to use a state-dependent Lagrange-multiplier  $\theta(V_t)$  in the framework of the multiplier approach to obtain homothetic utility. This approach is also adopted by [Liu et al. \(2004\)](#). However, by assuming that the ambiguity aversion parameter  $\theta$  can be state dependent, the relation with the constraint preferences is lost. [Pathak \(2002\)](#) gives an extensive discussion about this issue. He argues that

the main motivation of the multiplier approach by Hansen et al. (2001) is through the constraint approach. But with the state-dependent ambiguity aversion parameter this new utility specification cannot be seen anymore as a dynamic extension of the Gilboa-Schmeidler utility. Furthermore, the axiomatic foundation is not valid anymore. Pathak (2002) proposes an alternative method to model ambiguity: *recursive multiple priors utility* developed in continuous time by Chen and Epstein (2002).

We adopt the advise of Pathak (2002) and in contrast to Liu et al. (2004) we choose to use the approach of Chen and Epstein (2002). The approach is closely related to the constraint approach of Hansen et al. (2001), but does preserve the homotheticity of the preferences. Consider the following problem:

$$\begin{aligned} V_t = \min_{\{a_s, b_s\}_{s \geq t}} E_t^{\mathbb{Q}} \left[ \int_t^{\infty} f(C_s, V_s) ds \right] \\ \text{s.t. } d(a_t, b_t) \leq \theta_t \quad \forall t \end{aligned} \quad (6)$$

The main difference compared to the constraint approach of Hansen et al. (2001) is that not the lifetime distance between the reference measure and the alternative measure is bounded, but at every time period  $t$  the distance between the measures is bounded. This approach leads to more tractable solutions. The recursive multiple priors utility is axiomatized by Epstein and Schneider (2003).

Lastly we will briefly discuss the *smooth ambiguity* model, since it is often used in the literature as well. In this approach the agent first constructs a probability distribution that reflects his beliefs on  $\lambda$  and  $m$ . Define  $p(x, y) = P(\lambda = x, m = y)$ . 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 results in the following problem:

$$V_t = \int_{-1}^0 \int_0^{\infty} \left( p(x, y) \phi \left( E_t \left[ \int_t^{\infty} f(C_t, V_t) ds \mid \lambda = x, m = y \right] \right) \right) dx dy \quad (7)$$

Here the function  $\phi$  controls the ambiguity aversion of the agent. When  $\phi$  is a concave function, the agent is ambiguity averse. This approach was introduced by Klibanoff et al. (2005). 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. And since the recursive multiple priors utility is intuitive and leads to tractable results, we chose not to move in the direction of the smooth ambiguity model.

### 3.2.2 Distance measure: Relative Entropy

We have not defined the distance measure  $d(a_t, b_t)$  yet. To measure the distance between the reference model  $\mathbb{P}$  and an alternative model  $\mathbb{Q}^{a,b}$  we use the concept of *relative entropy*, a metric for the distance between two measures. Relative entropy is calculated in equation (12) in the next section. Hansen and Sargent (2008) give two reasons why relative entropy is suitable to measure the distance between two



models. First, relative entropy yields tractable outcomes. And second it is statistically interpretable. To illustrate the concept of statistical interpretability, consider the following example: denote by  $f(x)$  the probability density function (pdf) of a random variable  $X$  under  $\mathbb{P}$  and denote by  $f^{\mathbb{Q}}(x)$  the pdf of  $X$  under  $\mathbb{Q}$ . Assume that we have a random sample of  $n$  observations of  $X$ , namely  $x_1, \dots, x_n$ . Consider the following hypothesis:

$$H_0 : \mathbb{P} \text{ is true } \text{ vs. } H_1 : \mathbb{Q} \text{ is true} \quad (8)$$

We want to know whether, based on our sample, we can conclude that  $\mathbb{Q}$  is the most likely model of the two. The test can have two outcomes: it either rejects  $H_0$  and therefore concludes that  $\mathbb{Q}$  is more likely than  $\mathbb{P}$ , or it does not reject  $H_0$ . We can define the likelihood ratio:  $\Lambda(x) = \frac{\prod_1^n f^{\mathbb{Q}}(x_i)}{\prod_1^n f(x_i)}$ . The Neyman-Pearson lemma says that the most powerful test for this problem is to reject  $H_0$  if  $\Lambda(x) > k$  or equivalently  $\log(\Lambda(x)) > \log(k)$ . Then  $k$  is chosen such that the probability that we make the wrong decision (rejecting  $H_0$  if actually  $H_0$  is true) is very small. Here we see that the log-likelihood ratio determines the outcome of the test. So the question whether one model fits the data better than the other, can be answered using the log-likelihood ratio.

Measuring the distance between two probability densities using the concept of relative entropy, then comes down to taking the expectation of the log likelihood ratio under the alternative measure. Instead of using a random sample, as in the test case, one determines the ratio assuming the alternative model is true.

## 4 The model

In this section we derive an analytic expression for the value function based on the choices made in the previous section. We then present expressions for the pricing kernel, the equilibrium rate of interest, the risk premium and the consumption-to-wealth ratio. This completes the environment within which we will introduce the climate change model in Sections 7 and 8. We start by implementing the recursive multiple priors approach for ambiguity.

### 4.1 Implementation of the recursive multiple priors utility

The first step is to determine the set of admissible priors given the relative entropy constraint. In order to do so, for each  $a = (a_t)_{t \geq 0}$  and  $b = (b_t)_{t \geq 0}$  we define the measure  $\mathbb{Q}^{a,b}$  which is equivalent to  $\mathbb{P}$  and has Radon-Nikodym derivative  $\frac{d\mathbb{Q}^{a,b}}{d\mathbb{P}}|_{\mathcal{F}_t} = \xi_t^{a,b}$  where  $\xi_t^{a,b}$  follows:

$$d\xi_t^{a,b} = -(\lambda_t^{\mathbb{Q}} - \lambda)\xi_t^{a,b} dt - \frac{\lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \lambda m}{\sigma} \xi_t^{a,b} dZ_t + \left( e^{a_t + b_t W_t - b_t \mu_J - \frac{1}{2} b_t^2 \sigma_J^2} - 1 \right) \xi_{t-}^{a,b} dN_t \quad (9)$$

and  $\xi_0^{a,b} = 1$ . Here  $m = e^{\mu_J + \frac{1}{2} \sigma_J^2} - 1$  is the mean jump size under  $\mathbb{P}$ ,  $m_t^{\mathbb{Q}} = e^{\mu_J + b_t \sigma_J^2 + \frac{1}{2} \sigma_J^2} - 1$  is the mean jump size under  $\mathbb{Q}^{a,b}$  and  $\lambda_t^{\mathbb{Q}} = e^{a_t} \lambda$  is the arrival

rate under  $\mathbb{Q}^{a,b}$ .  $\xi_t^{a,b}$  is chosen such that the jump distribution under  $\mathbb{Q}^{a,b}$  has the characteristics described in the following proposition.

**Proposition 1.** *Under the measure  $\mathbb{Q}^{a,b}$ ,  $C_t$  follows:*

$$dC_t = \left( \mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} \right) C_t dt + \sigma C_t dZ_t^{\mathbb{Q}} + Y_t C_{t-} dN_t \quad (10)$$

where  $Z_t^{\mathbb{Q}}$  is a Brownian motion,  $N_t$  is a Poisson process with arrival rate  $\lambda_t^{\mathbb{Q}} = e^{a_t} \lambda$  and  $Y_t = e^{W_t} - 1$  where  $W_t$  follows a normal distribution with mean  $\mu_J + b_t \sigma_J^2$  and variance  $\sigma_J^2$ .

*Proof.* See Theorem 11.6.7 of [Shreve \(2004\)](#).  $\square$

In words this implies that under the measure  $\mathbb{Q}^{a,b}$  the arrival rate of the Poisson process becomes  $\lambda_t^{\mathbb{Q}} = e^{a_t} \lambda$  and the mean jump size changes to  $m_t^{\mathbb{Q}} = e^{\mu_J + \frac{1}{2} \sigma_J^2 + b_t \sigma_J^2} - 1$ . Clearly, a larger  $a_t$  leads to a higher arrival rate and a smaller  $b_t$  leads to a more negative jump size, both leading to a lower utility.

The Radon-Nikodym derivative  $\xi_t^{a,b}$  that we have specified is actually the ratio between the alternative measure  $\mathbb{Q}^{a,b}$  and the reference measure  $\mathbb{P}$ . We can use it to determine the relative entropy RE between the two measures. The relative entropy between  $\mathbb{Q}^{a,b}$  and  $\mathbb{P}$  over time unit  $\Delta$  is defined as  $E_t^{\mathbb{Q}}[\log(\frac{\xi_{t+\Delta}^{a,b}}{\xi_t^{a,b}})]$ . Here  $E_t^{\mathbb{Q}}$  denotes the expectation with respect to the alternative measure  $\mathbb{Q}^{a,b}$ . Then divide by  $\Delta$  and let  $\Delta \rightarrow 0$  to obtain the instantaneous relative entropy:  $RE(a_t, b_t) = \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} E_t^{\mathbb{Q}} \left[ \log \left( \frac{\xi_{t+\Delta}^{a,b}}{\xi_t^{a,b}} \right) \right]$ .

Applying Itô's lemma for jump processes to  $\xi_t^{a,b}$ <sup>6</sup>, we obtain the following dynamics for  $\log(\xi_t^{a,b})$ :

$$\begin{aligned} d \log(\xi_t^{a,b}) = & \left( -(\lambda_t^{\mathbb{Q}} - \lambda) - \frac{1}{2\sigma^2} (\lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \lambda m)^2 \right) dt - \frac{\lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \lambda m}{\sigma} dZ_t \\ & + \left( a_t + b_t W_t - b_t \mu_J - \frac{1}{2} b_t^2 \sigma_J^2 \right) dN_t \end{aligned} \quad (11)$$

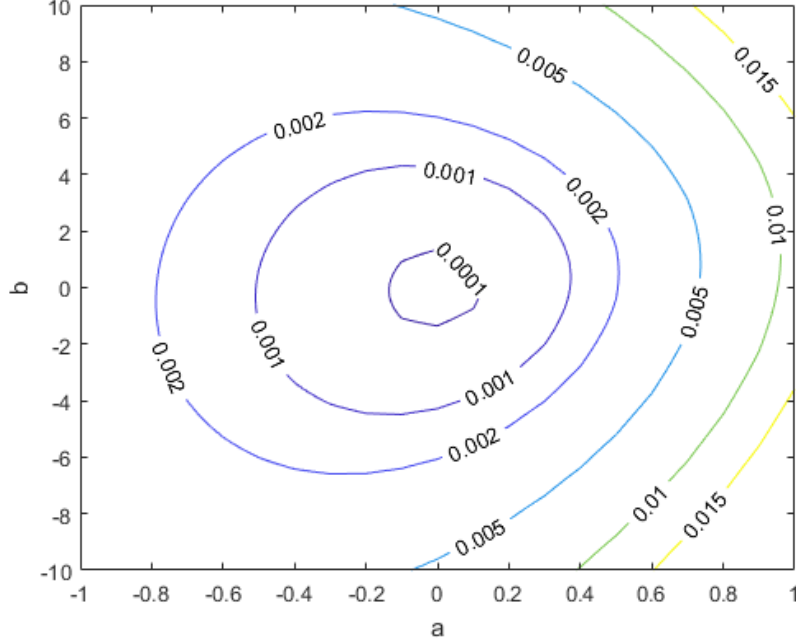
Now we can calculate (instantaneous) relative entropy at time  $t$ . We use here that under  $\mathbb{Q}^{a,b}$ ,  $N_t$  is a Poisson process with arrival rate  $e^{a_t} \lambda$ ,  $W_t$  is a normal random variable with mean  $\mu_J + b_t \sigma_J^2$  and variance  $\sigma_J^2$  and  $dZ_t = dZ_t^{\mathbb{Q}} - \frac{\lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \lambda m}{\sigma} dt$  where  $dZ_t^{\mathbb{Q}}$  is a Brownian motion under  $\mathbb{Q}^{a,b}$ .

$$\begin{aligned} RE(a_t, b_t) = & \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} E_t^{\mathbb{Q}} \left[ \log \left( \frac{\xi_{t+\Delta}^{a,b}}{\xi_t^{a,b}} \right) \right] = \\ & -(\lambda_t^{\mathbb{Q}} - \lambda) + \frac{1}{2\sigma^2} (\lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \lambda m)^2 + \lambda_t^{\mathbb{Q}} \left( a_t + \frac{1}{2} b_t^2 \sigma_J^2 \right) \end{aligned} \quad (12)$$

If we take a look at the relative entropy, it is clear that if  $(a_t, b_t) = (0, 0)$ , the relative entropy equals zero. When one or both of the two variables deviate from 0,

<sup>6</sup>See e.g. [Shreve \(2004\)](#), Theorem 11.5.1.

Figure 1: Relative entropy



The contour lines show the value of relative entropy for different values of  $a_t$  and  $b_t$ . Input parameters are  $\sigma = 0.03$ ,  $\mu_J = -10\%$ ,  $\sigma_J = 10\%$  and  $\lambda = 0.01$ .

the relative entropy increases. Every contour in figure 1 gives a set of combinations  $(a_t, b_t)$  that yields the same relative entropy.

We can now define the admissible set of priors. Define the set of priors by  $\mathcal{P}^\theta = \{Q^{a,b} : RE(a_t, b_t) \leq \theta_t \forall t\} = \{Q^{a,b} : (a_t, b_t) \in \Theta_t \forall t\}$ . The utility is then defined by:

$$V_t = \min_{Q \in \mathcal{P}^\theta} V_t^Q \quad (13)$$

where  $V_t^Q$  is the SDU utility process given the measure  $Q$ :

$$V_t^Q = E_t^Q \left[ \int_t^\infty f(C_s, V_s^Q) ds \right] \quad (14)$$

Following [Chen and Epstein \(2002\)](#) we impose a property on the set  $\mathcal{P}^\theta$  that they call rectangularity. In a loose sense, rectangularity implies invertibility between the set of multiple priors and the instantaneous actions following from the minimax optimization based on that set at any possible node  $\mathcal{F}_t$  in the filtration  $\mathbb{F}$ . At every node  $\mathcal{F}_t$ ,  $\mathcal{P}$  generates a set of conditional probabilities for the "one-step ahead" state variables. Conversely each time/event pair can be associated with a set of measures for the states in the next period. Define the collection of these sets by  $\mathcal{P}'$ . Rectangularity requires  $\mathcal{P}$  to coincide with  $\mathcal{P}'$ . If this is the case, we can replace the global minimization problem with a local minimization problem at every time  $t$ .

The parameter  $\theta_t$  captures the ambiguity aversion of the agent at time  $t$ . We assume that this parameter is constant over time. The parameter  $\theta$  then determines

the admissible set of priors in each time period. Since the parameters are constant over time as well, the set of admissible priors is the same for every  $t$ . A large  $\theta$  then implies that the agent has a high ambiguity aversion and he therefore considers a larger set of priors. Given this set of priors, the agent chooses the worst case measure.  $\theta = 0$  implies that  $\mathcal{P}^\theta = \{\mathbb{P}\}$  and the agent only considers one measure, namely the reference measure. There is no ambiguity aversion in this case. Where the risk aversion parameter  $\gamma$  can be seen a parameter that is relevant for any risky bet, the parameter  $\theta$  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, the set of admissible priors will be smaller compared to a process of which not much is known.

It is not necessary to have a constant  $\theta$ . One could for example incorporate learning by assuming that  $\theta_t$  is a decreasing function over time. The agent obtains more information about a process over time and therefore the set of priors is shrinking over time. However, there does not (yet) seem to exist a decent framework to determine how the set of priors should shrink over time based on observations. Similar to [Chen and Epstein \(2002\)](#) we will therefore focus on the case with a constant  $\theta$ .

## 4.2 Solving for the value function

We first derive for each measure  $\mathbb{Q}^{a,b}$  the corresponding Hamilton-Jacobi-Bellman (HJB) equation and find an expression for the value function  $V_t^\mathbb{Q}$ . At the end of the section we derive the solution for the value function with ambiguity aversion:  $V_t = \min_{\mathbb{Q} \in \mathcal{P}^\theta} V_t^\mathbb{Q}$ .

**Proposition 2.** *The value function  $V_t^\mathbb{Q}$  satisfies the following Hamilton-Jacobi-Bellman equation.*

$$\begin{aligned} 0 = & f(C_t, V_t^\mathbb{Q}) + V_C^\mathbb{Q}(\mu - \lambda_t^\mathbb{Q} m_t^\mathbb{Q}) C_t dt + \frac{1}{2} V_{CC}^\mathbb{Q} \sigma^2 C_t^2 \\ & + \lambda_t^\mathbb{Q} E_t^\mathbb{Q} [V^\mathbb{Q}((1+Y)C_{t-}) - V^\mathbb{Q}(C_{t-})] \end{aligned} \quad (15)$$

*Proof.* See appendix [A](#) □

Here  $V_C^\mathbb{Q}$  denotes the first derivative of  $V^\mathbb{Q}$  with respect to  $C_t$ . Similar notation is used for the second derivative. We can compare the HJB-equation with the baseline model without jumps and with power utility. There are several differences. Firstly, the utility  $u(C)$  is replaced by  $f(C, V)$ . Furthermore, there is an additional term that captures the expected jump in the value function. We also see the compensation term in the drift. Since we look at an infinite time horizon,  $V_t^\mathbb{Q}$  only depends on the state variable  $C_t$  and not directly on time. Therefore, from now on we can write  $V_t^\mathbb{Q} = V^\mathbb{Q}(C_t)$ .

The HJB equation is a partial differential equation. We need to find a more specific form of  $V_t^\mathbb{Q}$  to solve the equation. We first show that  $V_t^\mathbb{Q}$  is homogeneous of degree  $1 - \gamma$ .

**Proposition 3.**  *$V^\mathbb{Q}(C_t)$  is homogeneous of degree  $1 - \gamma$ .*

*Proof.* Conjecture that  $V^{\mathbb{Q}}(kC_t) = k^{1-\gamma}V^{\mathbb{Q}}(C_t)$ . We verify that this is indeed the case. Then:

$$\begin{aligned}
V^{\mathbb{Q}}(kC_t) &= E_t^{\mathbb{Q}} \left[ \int_t^{\infty} \frac{\beta}{1-1/\epsilon} \frac{(kC_s)^{1-1/\epsilon} - ((1-\gamma)V^{\mathbb{Q}}(kC_s))^{\frac{1}{\zeta}}}{((1-\gamma)V^{\mathbb{Q}}(kC_s))^{\frac{1}{\zeta}-1}} ds \right] \\
&= E_t \left[ \int_t^{\infty} \frac{\beta}{1-1/\epsilon} \frac{(kC_s)^{1-1/\epsilon} - ((1-\gamma)k^{1-\gamma}V^{\mathbb{Q}}(C_s))^{\frac{1}{\zeta}}}{((1-\gamma)k^{1-\gamma}V^{\mathbb{Q}}(C_s))^{\frac{1}{\zeta}-1}} ds \right] \\
&= E_t^{\mathbb{Q}} \left[ \int_t^{\infty} \frac{k^{1-1/\epsilon}}{k^{\gamma-1/\epsilon}} \frac{\beta}{1-1/\epsilon} \frac{(C_s)^{1-1/\epsilon} - ((1-\gamma)V^{\mathbb{Q}}(C_s))^{\frac{1}{\zeta}}}{((1-\gamma)V^{\mathbb{Q}}(C_s))^{\frac{1}{\zeta}-1}} ds \right] \\
&= k^{1-\gamma}V^{\mathbb{Q}}(C_t)
\end{aligned} \tag{16}$$

□

**Proposition 4.** *The value function under the measure  $\mathbb{Q}^{a,b}$  is of the following form:*

$$V^{\mathbb{Q}}(C_t) = g \frac{C_t^{1-\gamma}}{1-\gamma} \tag{17}$$

where  $g$  is a constant.

*Proof.* By Proposition 3, we have that  $V^{\mathbb{Q}}(C_t) = k^{\gamma-1}V^{\mathbb{Q}}(kC_t)$ . Now let  $k = 1/C_t$ . This gives  $V^{\mathbb{Q}}(C_t) = V^{\mathbb{Q}}(1)C_t^{1-\gamma}$ .  $V^{\mathbb{Q}}(1)$  does not depend on consumption anymore. Define the constant  $g$ :  $g = (1-\gamma)V^{\mathbb{Q}}(1)$ . Replacing  $V^{\mathbb{Q}}(1)$  by  $g/(1-\gamma)$  yields the form of  $V^{\mathbb{Q}}(C_t)$  given above. □

We can now substitute the form of the value function into the HJB-equation. Furthermore we can calculate the expectation. This gives the following HJB-equation.

**Proposition 5.** *The function  $g$  satisfies the following HJB-equation:*

$$0 = \frac{\beta}{1-1/\epsilon} \left( g^{-\frac{1}{\zeta}} - 1 \right) + (\mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}}) - \frac{\gamma}{2} \sigma^2 + \lambda_t^{\mathbb{Q}} \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \tag{18}$$

*Proof.* See Appendix B. □

From this equation we are able to derive the constant  $g$  for every measure  $\mathbb{Q}^{a,b}$ . Now let us return to the problem with ambiguity. We are not interested in the solution for every single measure  $\mathbb{Q}^{a,b}$ , but we want to find the solution to  $V_t = \min_{\mathbb{Q} \in \mathcal{P}^{\theta}} V_t^{\mathbb{Q}}$ . Since we impose rectangularity on the set  $\mathcal{P}^{\theta}$ , we can replace the global minimization problem of (13) by an instantaneous optimization problem at every time period  $t$ . The HJB-equation of the problem with ambiguity then becomes.

$$0 = \min_{(a_t, b_t) \in \Theta_t} \left\{ \frac{\beta}{1-1/\epsilon} \left( g^{-\frac{1}{\zeta}} - 1 \right) + \mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \frac{\gamma}{2} \sigma^2 + \lambda_t^{\mathbb{Q}} \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \right\} \tag{19}$$

### 4.3 Optimal control variables

From the HJB-equation we can then calculate the optimal control variables  $a^*$  and  $b^*$ . Note that since both the parameters of the model and the budget  $\theta$  are constant over time, both  $a^*$  and  $b^*$  are constant over time as well. In section 7 we extend the model with a climate model and model the arrival rate of climate disasters as a function of temperature. In that case,  $a^*$  and  $b^*$  are not constant anymore.

The minimization problem in the HJB-equation can be reduced to:

$$\min_{(a_t, b_t) \in \Theta_t} \left\{ \lambda_t^{\mathbb{Q}} \left( \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} - m_t^{\mathbb{Q}} \right) \right\} \quad (20)$$

Now we can determine the optimal  $a$  and  $b$ . This is a constrained optimization problem with Lagrangian:

$$L(a_t, b_t, l_t) = \lambda_t^{\mathbb{Q}} \left( \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} - m_t^{\mathbb{Q}} \right) - l_t (RE(a_t, b_t) - \theta) \quad (21)$$

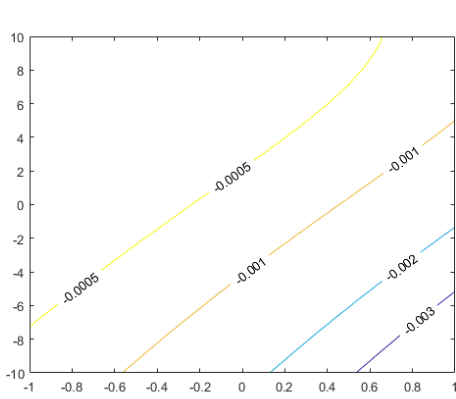
Here  $l_t$  is the Lagrange multiplier.

**Proposition 6.**  $a^*$  and  $b^*$  and the Lagrange-multiplier  $l_t$  are the solutions to the following first order conditions:

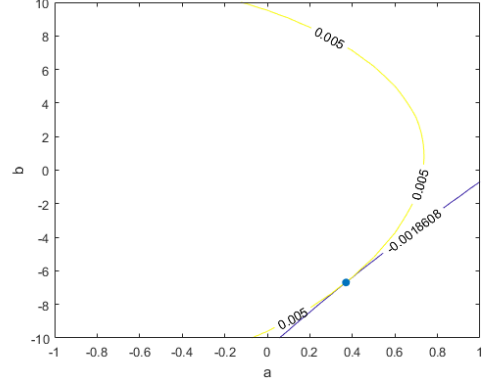
$$\begin{aligned} \frac{\partial}{\partial a_t} L(a_t, b_t, l_t) &= \lambda_t^{\mathbb{Q}} \left( \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} - m_t^{\mathbb{Q}} \right) \\ &\quad - l_t \lambda_t^{\mathbb{Q}} \left( \frac{1}{\sigma^2} m_t^{\mathbb{Q}} (\lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \lambda m) + a_t + \frac{1}{2} b_t^2 \sigma_J^2 \right) = 0 \\ \frac{\partial}{\partial b_t} L(a_t, b_t, l_t) &= \lambda_t^{\mathbb{Q}} \sigma_J^2 \left( e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - e^{\mu_J + \frac{1}{2} \sigma_J^2 + b_t \sigma_J^2} \right) \\ &\quad - l_t \lambda_t^{\mathbb{Q}} \sigma_J^2 \left( b_t + \frac{1}{\sigma^2} (\lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \lambda m) e^{\mu_J + \frac{1}{2} \sigma_J^2 + b_t \sigma_J^2} \right) = 0 \\ \frac{\partial}{\partial l_t} L(a_t, b_t, l_t) &= \theta + (\lambda_t^{\mathbb{Q}} - \lambda) - \frac{1}{2\sigma^2} (\lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} - \lambda m)^2 - \lambda^{\mathbb{Q}} \left( a_t + \frac{1}{2} b_t^2 \sigma_J^2 \right) = 0 \end{aligned} \quad (22)$$

Figure 2 illustrates the optimization problem. Given an entropy budget  $\theta$ , one can determine the feasible set of  $(a, b)$ . Figure 1 shows the feasible set 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 objective function. The goal is to minimize this function, given the relative entropy 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 size:  $\lambda^* = \lambda e^{a^*}$  and  $m^* = e^{\mu_J + \frac{1}{2} \sigma_J^2 + b^* \sigma_J^2}$ .

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 yellow line gives the admissible values for  $a$  and  $b$  that are within the relative entropy budget. Purple line is the objective function.

#### 4.4 The risk-free interest rate, risk premia and the consumption-wealth ratio

Since we know the optimal  $a^*$  and  $b^*$  we can now solve the HJB equation (19) to obtain  $g$ .

$$g = \left( 1 + \frac{(1/\epsilon - 1)}{\beta} \left( \mu - \frac{\gamma}{2} \sigma^2 - \lambda^* m^* + \lambda^* \frac{e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} \right) \right)^{-\zeta} \quad (23)$$

Furthermore, we now explicitly specify the asset market and a pricing kernel to obtain interest rates and risk premia. Assume that the representative agent has the possibility to invest in two assets, namely a risk-free asset and a risky stock. The risk-free asset with price  $B_t$  pays a fixed continuously compounded interest rate  $r$ . The stock pays continuous dividends at a rate  $C_t$  and has ex-dividend price  $S_t$ . We denote the cum-dividend stock price by  $S_t^d$ . Since all dynamics of the model are not time-dependent, we conjecture that the consumption to wealth ratio is constant. We will later verify this conjecture. This implies that  $kS_t = C_t$  for some constant  $k$ , since in equilibrium total wealth must be equal to the aggregate claim on consumption and thus the stock price. Under the reference measure  $\mathbb{P}$  the assets have the following distribution.

$$\begin{aligned} dB_t &= rdt \\ dS_t^d &= dS_t + C_t dt = \left( \mu - \lambda m + k \right) S_t dt + \sigma S_t dZ_t + Y_t S_t^- dN_t \end{aligned} \quad (24)$$

Duffie and Epstein (1992a) derive that the pricing kernel 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 [\(9\)](#).

We can derive an explicit stochastic differential equation for the pricing kernel. Using this pricing kernel, it is possible to calculate the endogenous risk-free rate and the endogenous risk premium of the stock. The risk premium of an asset equals the expected excess return of that asset compared to investing in the risk-free asset. Lastly, we obtain an expression for the consumption-wealth ratio. The following proposition summarizes these results.

**Proposition 7.** *The pricing kernel equals:*

$$d\pi_t = \left\{ -r - \lambda(e^{a^* - \gamma(\mu_J + b^* \sigma_J^2 - \frac{1}{2}\gamma \sigma_J^2)} - 1) \right\} \pi_t dt + \left\{ -\gamma \sigma - \frac{\lambda^* m^* - \lambda m}{\sigma} \right\} \pi_t dZ_t \\ + \left( e^{a^* + (b^* - \gamma)W_t - b^* \mu_J - \frac{1}{2}(b^*)^2 \sigma_J^2} - 1 \right) \pi_{t-} dN_t \quad (25)$$

The interest rate  $r$  is then given by:

$$r = \beta + \frac{1}{\epsilon}(\mu - \lambda^* m^*) - \left(1 + \frac{1}{\epsilon}\right) \frac{\gamma}{2} \sigma^2 - \left(\gamma - \frac{1}{\epsilon}\right) \lambda^* \frac{e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} \\ - \lambda^* (e^{-\gamma(\mu_J + b^* \sigma_J^2 - \frac{1}{2}\gamma \sigma_J^2)} - 1) \quad (26)$$

Furthermore, the risk premium for the dividend paying stock equals:

$$\gamma \sigma^2 + \lambda^* m^* - \lambda^* e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} + \lambda^* e^{-\gamma(\mu_J + b^* \sigma_J^2 - \frac{1}{2}\gamma \sigma_J^2)} \quad (27)$$

Lastly, the consumption-wealth ratio is given by:

$$k = \beta g^{-\frac{1}{\epsilon}} = \beta + (1/\epsilon - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 - \lambda^* m^* + \lambda^* \frac{e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} \right) \quad (28)$$

*Proof.* See appendix [C](#) □

## 5 The impact of different preferences on asset prices and the consumption-wealth ratio

Since we include both ambiguity aversion and stochastic differential utility in our model, we have three preference parameters. Therefore it is useful to look at the results of proposition [7](#) in more detail before integrating the model with a climate dynamics model which we will do in section [7](#). In this section, we use the power utility case as a benchmark to compare the general outcomes with the outcomes if either ambiguity aversion or stochastic differential utility is not present and when both are not present. The first situation is obtained by letting  $\theta \rightarrow 0$ , the second situation by setting  $\frac{1}{\epsilon} = \gamma$ . In order to be able to compare different preference settings, we



make some additional assumptions on the parameters. In particular, for this section we assume that both  $\gamma > 1$  and  $\epsilon > \frac{1}{\gamma}$  when we consider stochastic differential utility, which is in line with empirical evidence (Epstein and Zin (1989), Van Binsbergen et al. (2012)).<sup>7</sup>

## 5.1 The safe rate of interest

Consider first the equilibrium interest rate in more detail. In the climate change debate, the equilibrium interest rate is important since it plays a crucial role in discounting future damages. Therefore it is useful to decompose the interest rate and see how differences in the structure of preferences influence it.

Let us start with the simplest interest rate possible. Without any risk ( $\lambda = 0, \sigma = 0$ ), the interest rate would equal  $\beta + \frac{1}{EIS}\mu$  where  $EIS$  is the elasticity of intertemporal substitution. This is the well-known Ramsey equation (Ramsey, 1928). In the case of power utility, the  $EIS$  equals  $\frac{1}{\gamma}$ , in the stochastic differential utility case it equals  $\epsilon$ .

If we add Brownian risk to the exogenous consumption stream ( $\lambda = 0, \sigma > 0$ ), the interest rate goes down by  $(1 + \frac{1}{EIS})\frac{\gamma}{2}\sigma^2$ . So a higher relative risk aversion  $\gamma$  decreases the interest rate. This is an intuitive result: higher risk aversion increases demand for the safe asset and therefore its price, or, equivalently, lowers the safe rate of interest. The effect of the  $EIS$  depends on whether  $\mu > \frac{\gamma}{2}\sigma^2$ . It is plausible that this holds as one can see by simply looking at the consumption process. A higher  $EIS$  decreases the interest rate, similar to what happens in the Ramsey case. This is again intuitive, since a higher  $EIS$  implies higher propensity to save and therefore also increases demand for the risk-free asset.

If we compare the stochastic differential utility with power utility preferences, we see that the interest rate is lower for the stochastic differential utility case. Now define the interest rate without jumps in the power utility case and the stochastic differential utility (SDU) case by respectively  $r^{Pow}$  and  $r^{SDU}$ . Then  $r^{Pow} = \beta + \gamma\mu - (1 + \gamma)\frac{\gamma}{2}\sigma^2$  and  $r^{SDU} = \beta + \frac{1}{\epsilon}\mu - (1 + \frac{1}{\epsilon})\frac{\gamma}{2}\sigma^2$ .

To isolate the impact of the various deviations from the standard time separable power utility case, we vary them in turn. Consider the four situations described

---

<sup>7</sup>Although empirical evidence seems to offer compelling support for  $\epsilon > 1 > 1/\gamma$ , the EZ approach has not gone unchallenged. In particular Epstein, Farhi, and Strzalecki (2014) argue that this condition, if fulfilled, leads to implausible values for the willingness to pay for early resolution of uncertainty. These values do depend on strong persistence in the stochastic structures however. It seems fair to argue that the jury is still out on EZ preferences.

above:

**1. Power utility, no ambiguity aversion**

$$r_1 = r^{Pow} - \gamma \lambda m - \lambda(e^{-\gamma(\mu_J - \frac{1}{2}\gamma\sigma_J^2)} - 1)$$

**2. SDU utility, no ambiguity aversion**

$$r_2 = r^{SDU} - \frac{1}{\epsilon} \lambda m - \left(\gamma - \frac{1}{\epsilon}\right) \lambda \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} - \lambda(e^{-\gamma(\mu_J - \frac{1}{2}\gamma\sigma_J^2)} - 1)$$

**3. Power utility, ambiguity aversion**

$$r_3 = r^{SDU} - \gamma \lambda^* m^* - \lambda^*(e^{-\gamma(\mu_J + b^*\sigma_J^2 - \frac{1}{2}\gamma\sigma_J^2)} - 1)$$

**4. SDU utility, ambiguity aversion**

$$r_4 = r^{SDU} - \frac{1}{\epsilon} \lambda^* m^* - \left(\gamma - \frac{1}{\epsilon}\right) \lambda^* \frac{e^{(1-\gamma)(\mu_J + b^*\sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} - \lambda^*(e^{-\gamma(\mu_J + b^*\sigma_J^2 - \frac{1}{2}\gamma\sigma_J^2)} - 1)$$

Start with power utility and no ambiguity aversion (case one): including jump risk adds two additional terms to the interest rate. Since the drift of the consumption stream is compensated for the expected loss of disasters, including jump risk also changes the drift of the consumption stream. The first additional term is present because of this compensation. The second term is present because of the risk-aversion of the agent. When the agent is more risk-averse, the demand for the risk-free asset is higher and in equilibrium this leads to a lower interest rate. Jump risk always decreases the interest rate and the effect is larger for a larger risk aversion  $\gamma$ . To verify this, we can write the jump terms as:

$$\begin{aligned} & -\lambda\left(\gamma(e^{c_1} - 1) + (e^{-\gamma c_3} - 1)\right) = \lambda\gamma \sum_{k=1}^{\infty} \left(\frac{(-\gamma)^{k-1} c_3^k - c_1^k}{k!}\right) = \\ & \lambda\gamma \left(\underbrace{c_3 - c_1}_{<0} + \underbrace{\frac{-\gamma c_3^2 - c_1^2}{2}}_{<0} + \underbrace{\frac{\gamma^2 c_3^3 - c_1^3}{6}}_{<0} + \underbrace{\frac{-\gamma^3 c_3^4 - c_1^4}{24}}_{<0} + \dots\right) < 0 \end{aligned} \quad (29)$$

with  $c_1 = \mu_J + \frac{1}{2}\sigma_J^2 < 0$ ,  $c_3 = \mu_J - \frac{1}{2}\gamma\sigma_J^2 < 0$  and  $c_3 < c_1$ . We use  $c_1$  and  $c_3$  here on purpose,  $c_2$  will be defined later. Note that  $m = e^{c_1} - 1$ . With the assumption that  $\gamma > 1$ , clearly all terms are negative. The total sum is decreasing in  $\gamma$ .

Consider next the SDU case, but still without ambiguity aversion (case 2). This

time we can write the jump terms as:

$$\begin{aligned}
& -\frac{1}{\epsilon}\lambda(e^{c_1} - 1) - \left(\gamma - \frac{1}{\epsilon}\right)\lambda\frac{e^{(1-\gamma)c_2} - 1}{1-\gamma} - \lambda(e^{-\gamma c_3} - 1) = \\
& \lambda\frac{1}{\epsilon}\sum_{k=1}^{\infty}\left(\frac{(1-\gamma)^{k-1}c_2^k - c_1^k}{k!}\right) + \lambda\gamma\sum_{k=1}^{\infty}\left(\frac{(-\gamma)^{k-1}c_3^k - (1-\gamma)^{k-1}c_2^k}{k!}\right) = \\
& \lambda\frac{1}{\epsilon}\left(\underbrace{c_2 - c_1}_{<0} + \underbrace{\frac{(1-\gamma)c_2^2 - c_1^2}{2}}_{<-\frac{1}{2}c_1^2} + \underbrace{\frac{(1-\gamma)^2c_2^3 - c_1^3}{6}}_{<-\frac{1}{6}c_1^3} + \dots\right) \\
& + \lambda\gamma\left(\underbrace{c_3 - c_2}_{<0} + \underbrace{\frac{-\gamma c_3^2 + (\gamma-1)c_2^2}{2}}_{<0} + \underbrace{\frac{\gamma^2c_3^3 - (1-\gamma)^2c_2^3}{6}}_{<0} + \dots\right) < 0
\end{aligned} \tag{30}$$

where  $c_1$  and  $c_3$  are already defined,  $c_2 = \mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2 < 0$  and  $c_3 < c_2 < c_1$ . For  $\gamma > 1$ , both sums are negative and therefore jump risk decreases the interest rate ( $r_2 < r^{SDU}$ ). Increasing  $\gamma$  decreases both sums, which implies that a higher  $\gamma$  amplifies the flight to safety effect. Increasing  $\epsilon$  increases the first sum, which implies that  $r_2 - r^{SDU}$  is increasing in  $\epsilon$ . However, note that  $\epsilon$  also affects  $r^{SDU}$ , so the overall effect of increasing  $\epsilon$  will be a lower interest rate. Comparing case one and two, we have for plausible parameter choices that  $r_2 < r_1$ <sup>8</sup>. Due to the additional term we also have that  $r_2 - r^{SDU} > r_1 - r^{Pow}$ . This implies that jump risk decreases the interest rate in both cases, but has a larger effect (lowers the interest rate more) in the power utility case.

Effectively, ambiguity aversion increases the arrival rate  $\lambda$  and decreases the jump size parameter  $\mu_J$ . We did not assume a particular value for  $\lambda$  and  $\mu_J$  yet. Therefore, all the results from case 1 and 2 carry over to the ambiguity aversion cases, but the effects are somewhat stronger in case 3 and 4. The higher the ambiguity aversion parameter  $\theta$ , the higher  $a^*$  and the lower  $b^*$ .

Now we go back to power utility but introduce ambiguity aversion (case 3): when the representative agent is ambiguity averse, introducing jump risk again lowers the interest rate. To analyze the interest rate in case 3 we can replace  $\lambda$  by  $\lambda^*$  and  $c_1$  and  $c_3$  in (29) by  $c_1^* = \mu_J + b^*\sigma_J^2 + \frac{1}{2}\sigma_J^2$  and  $c_3^* = \mu_J + b^*\sigma_J^2 - \frac{1}{2}\gamma\sigma_J^2$ . Clearly, for  $a^* > 0$ ,  $\lambda^* > \lambda$ . We see that for  $b^* < 0$ ,  $c_3^* - c_1^* = c_3 - c_1$  but  $\frac{-\gamma(c_3^*)^2 - (c_1^*)^2}{2} < \frac{-\gamma c_3^2 - c_1^2}{2}$  and this holds for the higher order terms as well. Overall we can therefore conclude that  $r_3 < r_1$  when ambiguity aversion  $\theta > 0$ .

Finally consider the combination of SDU with ambiguity aversion (case 4): similar to the situation without ambiguity aversion, we have that  $r_4 < r_3$  for plausible parameter values<sup>9</sup>. So introducing SDU also leads to lower interest rates when starting with ambiguity aversion but without SDU, just like the impact of introducing SDU starting from just power utility, but in addition to this inequality, we also have:

<sup>8</sup>Specifically,  $r_2 < r_1$  when  $\mu - \lambda m > \frac{\gamma}{2}\sigma^2 - \lambda\frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma}$ .

<sup>9</sup>Specifically,  $r_4 < r_3$  when  $\mu - \lambda^* m^* > \frac{\gamma}{2}\sigma^2 + \frac{e^{(1-\gamma)(\mu_J + b^*\sigma_J^2 - \lambda^*\frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma}$ .

$r_4 - r^{SDU} > r_3 - r^{Pow}$ . This implies that SDU and Ambiguity aversion do not reinforce each other, the impact of introducing SDU without ambiguity aversion present is larger than the impact of introducing SDU when ambiguity aversion is already present. We discuss this issue in more detail later in this section.

## 5.2 The risk premium

The risk premium of the stock only depends on the risk and ambiguity aversion of the agent. The EIS plays no role, as a consequence the risk premium is the same for both power utility and SDU. That reduces our four cases to two:

### 1./2. No ambiguity aversion

$$r^S - r = \gamma\sigma^2 + \lambda m - \lambda e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} + \lambda e^{-\gamma(\mu_J - \frac{1}{2}\gamma\sigma_J^2)}$$

### 3./4. Ambiguity aversion

$$r^S - r = \gamma\sigma^2 + \lambda^* m^* - \lambda^* e^{(1-\gamma)(\mu_J + b^*\sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} + \lambda^* e^{-\gamma(\mu_J + b^*\sigma_J^2 - \frac{1}{2}\gamma\sigma_J^2)}$$

The risk premium can be divided in a diffusive part ( $\gamma\sigma^2$ ) and a jump part. The diffusive part is the standard formula under power utility and remains the same under SDU. Without jump risk, ambiguity aversion plays no role, but with jump risk it obviously does play a role. The two cases show the impact of the introduction of jump risk without and with ambiguity aversion. Using a similar method as for the interest rates, we can analyse the risk premium.

$$\begin{aligned} \lambda(e^{c_1} - 1) - \lambda e^{(1-\gamma)c_2} + \lambda e^{-\gamma c_3} &= \lambda \sum_{k=1}^{\infty} \left( \frac{c_1^k - (1-\gamma)^k c_2^k + (-\gamma)^k c_3^k}{k!} \right) \\ &= \lambda \left( \underbrace{c_1 - c_2}_{>0} + \underbrace{\gamma(c_2 - c_3)}_{>0} + \underbrace{\frac{c_1^2 - (1-\gamma)^2 c_2^2 + \gamma^2 c_3^2}{2}}_{>0} + \underbrace{\frac{c_1^3 - (1-\gamma)^3 c_2^3 - \gamma^3 c_3^3}{6}}_{>0} + \dots \right) > 0 \end{aligned} \quad (31)$$

$c_1$ ,  $c_2$  and  $c_3$  where defined in the previous section where  $c_3 < c_2 < c_1 < 0$ . It is obvious that a higher  $\gamma$  also increases the jump risk premium.

In fact the ambiguity aversion parameter  $\theta$  does not directly influence the risk premium, but does so via  $a^*$  and  $b^*$ . A higher  $\theta$  leads to a higher optimal value for  $a$  and a lower optimal value for  $b$ ; the agent becomes as it were more pessimistic with higher ambiguity aversion. Since the jump risk premium is linear in  $\lambda$  it is straightforward to verify that  $a^* > 0$  increases the jump risk premium. For the ambiguity aversion case, define  $c_2^* = \mu_J + b^*\sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2$ .  $c_1^*$  and  $c_3^*$  are already defined in the previous subsection. Then the following inequalities hold:  $c_3^* < c_3$ ,  $c_2^* < c_2$ ,  $c_1^* < c_1$  and  $c_3^* < c_2^* < c_1^* < 0$ . One can verify that  $c_1^* - c_2^* = c_1 - c_2$  and  $c_2^* - c_3^* = c_2 - c_3$ , but for the second term we have  $\frac{(c_1^*)^2 - (1-\gamma)^2(c_2^*)^2 + \gamma^2(c_3^*)^2}{2} > \frac{c_1^2 - (1-\gamma)^2 c_2^2 + \gamma^2 c_3^2}{2}$ . Similar inequalities hold for higher order terms and therefore the equity premium is larger when ambiguity aversion is present ( $\theta > 0$ ).

### 5.3 The consumption-to-wealth ratio

We look at a stationary problem (i.e. the processes and parameters do not change over time), so the consumption-to-wealth ratio  $k$  is constant over time. We can use the expression for the consumption to wealth ratio in equation (28) to derive the expression for the consumption-to-wealth ratio for the same four cases we distinguished before:

#### 1. Power utility, no ambiguity aversion

$$k_1 = \beta + (\gamma - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 - \lambda m + \lambda \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} \right)$$

#### 2. SDU utility, no ambiguity aversion

$$k_2 = \beta + (1/\epsilon - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 - \lambda m + \lambda \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} \right)$$

#### 3. Power utility, ambiguity aversion

$$k_3 = \beta + (\gamma - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 - \lambda^* m^* + \lambda^* \frac{e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} \right)$$

#### 4. SDU utility, ambiguity aversion

$$k_4 = \beta + (1/\epsilon - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 - \lambda^* m^* + \lambda^* \frac{e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} \right) \quad (32)$$

Without jump risk, all the variants yield the traditional expression; for example, with logarithmic utility ( $\gamma = 1$ ), consumption equals wealth times the time preference parameter  $\beta$ . Introducing jump risk under power utility (case 1) lowers the consumption wealth ratio for risk aversion larger than one. This is the case since:

$$\begin{aligned} & -\lambda(e^{c_1} - 1) + \lambda \frac{e^{(1-\gamma)c_2} - 1}{1 - \gamma} = \lambda \sum_{k=1}^{\infty} \left( \frac{(1-\gamma)^{k-1} c_2^k - c_1^k}{k!} \right) \\ & = \lambda \left( \underbrace{c_2 - c_1}_{<0} + \underbrace{\frac{(1-\gamma)c_2^2 - c_1^2}{2}}_{<-\frac{1}{2}c_1^2} + \underbrace{\frac{(1-\gamma)^2 c_2^3 - c_1^3}{6}}_{<-\frac{1}{6}c_1^3} + \dots \right) < 0. \end{aligned} \quad (33)$$

And there is of course overwhelming empirical evidence in favor of  $\gamma > 1$  (Cochrane, 2009). But that leads to another empirical puzzle. There is also empirical evidence in favor of a negative effect of jump risk on asset prices which would lead one to expect a higher ratio of consumption to wealth (Berkman, Jacobsen, & Lee, 2011).

But note that in our general equilibrium framework, jump risk not only influences the price of risk but also the interest rate. As we have seen in subsection 5.1, jump risk decreases the interest rate and a lower interest rate leads to higher asset prices. The total effect of jump risk on asset prices is a combination of a negative risk effect and a positive interest rate effect. We see from (32) that as long as  $EIS < 1$ , as is implied by power utility in combination with the empirically strongly supported

assumption that  $\gamma > 1$ , the interest rate effect will dominate and prices will increase when jump risk is included, which is a counterfactual prediction.

But the puzzle can be resolved by switching to SDU (case 2). Under SDU, the  $EIS$  and the risk aversion parameter  $\gamma$  can be chosen independently so we are able to have both a realistic risk aversion  $\gamma > 1$  and an overall positive effect of jump risk on the consumption wealth ratio, which requires  $EIS > 1$ . In fact we can interpret the empirical results of [Berkman et al. \(2011\)](#) as support for SDU and the assumption of  $EIS > 1$ .

Furthermore we see that introducing ambiguity aversion (case 3) has the same impact on the introduction of jump risk: introducing ambiguity aversion also leads to a lower consumption to wealth ratio in the presence of jump risk when  $\gamma > 1$ . Similar to the analysis with the safe rate of interest and the equity premium, we can use  $c_1^*$  and  $c_2^*$  here to show this. Remember that  $c_2^* < c_2$ ,  $c_1^* < c_1$  and  $c_2^* < c_1^* < 0$ . When  $\gamma > 1$ , the consumption wealth ratio under ambiguity aversion is even lower than the consumption wealth ratio under power utility ( $k_3 < k_1$ ). This is the case since  $c_2^* - c_1^* = c_2 - c_1$  and  $\frac{(1-\gamma)(c_2^*)^2 - (c_1^*)^2}{2} < \frac{(1-\gamma)c_2^2 - c_1^2}{2}$ . This inequality also holds for all terms for which the power is even. For the third term and the other odd terms the sign depends on the input parameters. For  $\gamma > 2$  it is easy to verify that  $\frac{(1-\gamma)^2(c_2^*)^3 - (c_1^*)^3}{6} < \frac{(1-\gamma)^2c_2^3 - c_1^3}{6}$ . For  $1 < \gamma < 2$  this is not directly clear, but since the second term dominates the third term the result will still be that  $k_3 < k_1$ . The main reason for this is that the interest rate effect dominates when  $\gamma > 1$ , and with ambiguity aversion the interest rate is even lower which implies that the interest rate effect is stronger compared to power utility. The interest rate effect becomes stronger for a higher level of ambiguity aversion.

The results of the combination of SDU and ambiguity aversion (case 4) depend on whether the  $EIS$  is larger or smaller than one. The empirically relevant case is the case where jump risk decreases prices and therefore increases the consumption to wealth ratio. This is the case when the  $EIS > 1$ . Using similar reasoning as in case 3, we obtain the following result for case 4. When the  $EIS > 1$ ,  $k_4 > k_2$ . So the impact of introducing ambiguity aversion actually adds to the effect of using SDU with  $EIS > 1$  in assessing the impact of the introduction of jump risk on the consumption-wealth ratio.

In fact we can go further than just stating that ambiguity aversion adds to the magnification effect triggered by SDU, we can show that the two reinforce each other. To do so compare the incremental effect of adding ambiguity aversion with and without SDU. Label the first  $\Delta_1$  and the second  $\Delta_0$ . If we can prove  $\Delta_1 > \Delta_0$ , we have established that the two (SDU and AA) not only are additive but that they reinforce each other. We get for the incremental effect of introducing ambiguity aversion for power utility,  $\Delta_0$ :

$$\Delta_0 = (k_3 - k_1) = (\gamma - 1) \left( -\lambda^* m^* + \lambda^* \frac{e^{(1-\gamma)c_2^*} - 1}{1 - \gamma} + \lambda m - \frac{e^{(1-\gamma)c_2} - 1}{1 - \gamma} \right) \quad (34)$$

Similar calculations yield for  $\Delta_1$ :

$$\Delta_1 = (k_4 - k_2) = \left(\frac{1}{\epsilon} - 1\right) \left( -\lambda^* m^* + \lambda^* \frac{e^{(1-\gamma)c_2^*} - 1}{1-\gamma} + \lambda m - \frac{e^{(1-\gamma)c_2} - 1}{1-\gamma} \right) \quad (35)$$

Combining equations (34) and (35) gives us after some rearranging of terms the amplification effect, the difference between the incremental impact of introducing AA in an SDU environment with i.e.  $\epsilon, \gamma > 1$  ( $\Delta_1$ ) and the incremental impact of introducing AA for power utility, i.e. with  $\epsilon = 1/\gamma$  ( $\Delta_0$ ):

$$\begin{aligned} \Delta_1 - \Delta_0 &= \left(\frac{1}{\epsilon} - \gamma\right) \left( -\lambda^* m^* + \lambda^* \frac{e^{(1-\gamma)c_2^*} - 1}{1-\gamma} + \lambda m - \frac{e^{(1-\gamma)c_2} - 1}{1-\gamma} \right) \\ &= \underbrace{\left(\frac{1}{\epsilon} - \gamma\right)}_{<0} \underbrace{\left( \frac{(1-\gamma)\left((c_2^*)^2 - c_2^2\right) - (c_1^*)^2 + c_1^2}{2} \right)}_{<0} \\ &\quad + \underbrace{\frac{(1-\gamma)^2\left((c_2^*)^3 - c_2^3\right) - (c_1^*)^3 + c_1^3}{6}}_{<0 \text{ if } \gamma > 2} + \dots > 0 \end{aligned} \quad (36)$$

So introducing Ambiguity Aversion and an SDU environment simultaneously indeed has a larger impact than the sum of the impacts of introducing SDU and Ambiguity Aversion separately: there truly is a reinforcement effect. For the case  $1 < \gamma < 2$ , the second term does not have to be negative. However, since the absolute value of all the  $c$ 's is smaller than one, the first term will dominate and the overall effect will still be the same. A similar argument works for the higher order terms.

## 5.4 The value of resolving risk

As we explained in the introduction, this model is not set up to estimate how costly it will be to reduce disaster risk in terms of the welfare costs of the policy actions required to do so. However, we can calculate how much an agent would be willing to pay in terms of reduced consumption to reduce the risk. Consider the following example. Assume that an ambiguity averse agent has the choice between the situation with climate risk and the same situation but without climate risk. Eliminating the risk is costly (think of abatement policies such as carbon taxes or closure of coal fueled electricity generating plants). The question we can answer using our model is: how much is the agent willing to give up in terms of consumption as the price for eliminating the climate risk. Comparing the utility of the representative agent in the case with jump risk with the case without jump risk allows us to calculate how much consumption reduction would make the agent indifferent between the situation with and the situation without jump risk.

Define utility in the situation without climate risk as  $\bar{V}_0 = \bar{g} \frac{\bar{C}_0^{1-\gamma}}{1-\gamma}$ . Here  $\bar{g} = \left(1 - \frac{(1-1/\epsilon)}{\beta} \left(\mu - \frac{\gamma}{2}\sigma^2\right)\right)^{-\zeta}$ . Then the agent is indifferent if  $V_0 = \bar{V}_0$ . Normalize

consumption at time  $t = 0$  to one ( $C_0 = 1$ ). Then we can calculate for which  $\overline{C}_0 < 1$  the agent is indifferent:

$$\overline{C}_0 = \left(\frac{g}{\bar{g}}\right)^{\frac{1}{1-\gamma}} \quad (37)$$

Maybe it is more realistic to assume that we cannot eliminate climate risk completely, but that we can at least reduce it: assume that we are able to reduce the expected shock size and/or volatility by increasing  $\mu_J$  to  $\overline{\mu}_J$  or by decreasing  $\sigma_J$  to  $\overline{\sigma}_J$  or both. The new expected jump size becomes  $\overline{m} = e^{\overline{\mu}_J + \frac{1}{2}\overline{\sigma}_J^2} - 1$  and in the ambiguity case  $\overline{m}^* = e^{\overline{\mu}_J + \overline{b}^* \sigma_J^2 + \frac{1}{2}\overline{\sigma}_J^2} - 1$ . Furthermore, we assume that it is also possible to reduce the arrival rate from  $\lambda$  to  $\overline{\lambda}$ . The arrival rate with ambiguity aversion becomes  $\overline{\lambda}^* = \overline{\lambda} e^{\overline{a}^*}$ . Note that  $a^*$  and  $b^*$  depend on  $\mu_J$ ,  $\lambda$  and  $\sigma_J$  and therefore also change. We can again calculate  $\overline{C}_0$  using (37) with  $\bar{g} = \left(1 - \frac{(1-1/\epsilon)}{\beta} \left(\mu - \frac{\gamma}{2}\sigma^2 - \overline{\lambda}^* \overline{m}^* + \overline{\lambda}^* \frac{e^{(1-\gamma)(\overline{\mu}_J + \overline{b}^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\overline{\sigma}_J^2) - 1}}{1-\gamma}\right)\right)^{-\zeta}$ . It is more intuitive to write the expression for  $\overline{C}_0$  as a function of the consumption-wealth ratio in both cases. Using  $g = \left(\frac{k}{\beta}\right)^{-\zeta} = \left(\frac{k}{\beta}\right)^{-\frac{1-\gamma}{1-1/\epsilon}}$  we obtain:

$$\overline{C}_0 = \left(\frac{\overline{k}}{k}\right)^{\frac{1}{1-1/\epsilon}} \quad (38)$$

As mentioned in the previous subsection, when  $\epsilon < 1$ , the interest rate effect dominates the consumption-wealth ratio and  $\overline{k} > k$ . Since  $\frac{1}{1-1/\epsilon} < 0$ , we obtain  $\overline{C}_0 < 1$ . For  $\epsilon > 1$ , the consumption-wealth ratio decreases when risk increases and therefore  $\overline{k} < k$ . However, since  $\frac{1}{1-1/\epsilon} > 0$  in this case we again find  $\overline{C}_0 < 1$ . It is not directly clear what the effect of  $\epsilon$  and  $\theta$  are on  $\overline{C}_0$ , since  $\epsilon$  and  $\theta$  affect both  $k$  and  $\overline{k}$ . In the next section we will present a numerical example to get a better idea of the implications of parameter choices.

## 6 The impact of different preferences on asset prices and the consumption-wealth ratio: a numerical example

Next we present numerical examples to quantify the implications of jump risk, stochastic differential utility and ambiguity aversion on the outcomes of the model and the extent to which doing so resolves well known asset pricing puzzles. In the previous section, we analysed the sensitivity of model solutions to the choices made for some key parameters. Closed form solutions are useful for that, but often multiple effects play a role and it is hard to determine which effect dominates without choosing specific parameter values. A numerical example then helps to illustrate the relations between endogenous variables and input parameters and gives a sense of magnitudes. These examples also help with the calibration of the utility parameters in the climate model that we will introduce in the next section.



We start off with the model without jumps. If we set  $\lambda = 0$ , we obtain the standard Consumption-based Capital Asset Pricing Model (CCAPM) developed in Lucas Jr (1978). Mehra and Prescott (1985) found that this model is not consistent with empirical data, coining the by now well known term ‘Equity Premium Puzzle’. Using data from the US between 1889 and 1978 they estimated the average growth rate of consumption to be roughly 2% and the consumption volatility to be roughly 3.5%, although in later years the consumption volatility seems to have decreased. We match those numbers by choosing  $\mu = 0.02$  and  $\sigma = 0.03$ , which is similar to Bansal and Yaron (2004). Furthermore Mehra and Prescott (1985) find a sample average risk-free rate of around 1% and a sample average risk premium of roughly 6%. In a more recent and more global analysis Dimson, Marsh, and Staunton (2011) estimate the global equity premium to be around 3 to 3.5%, still much higher than the basic CCAPM produces.

Using power utility and the plausible assumption that the risk aversion coefficient  $\gamma$  should be at most 10, Mehra and Prescott (1985) show that it is not possible to produce a realistic risk-free rate and equity premium at the same time. These two puzzles are called the Risk-Free Rate Puzzle and the Equity Premium puzzle. For values of  $\gamma$  below 1, it is possible to produce a low risk-free rate close to 1%. However, this results in a very low equity premium. On the other hand, choosing  $\gamma$  close to 10 results in a higher equity premium. The equity premium equals  $\gamma\sigma^2$ . Therefore, even for  $\gamma = 10$  we obtain an equity premium that is far below 6% (1.6%), even below the low value that Dimson et al. (2011) find. And this also results in a very high risk-free rate. These puzzles are known as the equity premium puzzle and the risk-free rate puzzle.

Table 1 gives an overview of all parameter values that are used for the examples. We set  $\gamma = 5$  and  $\beta = 0.02$ . Other parameter values are set to roughly match the values in Mehra and Prescott (1985). If we consider the CCAPM situation with power utility and without jump risk, we see that indeed the model produces a too high risk-free rate (10.7%) and a too low equity premium (0.45%). Now let us take a look what happens if we include jump risk, ambiguity aversion and stochastic differential utility. We assume climate disasters take place on average every 62.5 years ( $\lambda = 0.016$ ) and the expected jump size equals -9% ( $\mu_J = -10\%$ ,  $\sigma_J = 10\%$ ). In the next section we will look at a model where the arrival rate is temperature dependent. The calibration of the model with temperature-dependent arrival rate is such that an arrival rate of 0.016 corresponds with the temperature anomaly of  $2^\circ\text{C}$  compared to the pre-industrial level. This arrival rate is particularly interesting to study since the target of the Paris climate agreement is to keep global warming well below  $2^\circ\text{C}$ .

## 6.1 Interest rate

Figure 3 shows the interest rates for different values of the elasticity of intertemporal substitution  $\epsilon$  and the ambiguity aversion parameter  $\theta$ . Subfigure 3a shows the interest rate as a function of  $\epsilon$  in the case of SDU without ambiguity aversion (Case 2.). Subfigure 3b shows  $r$  as a function of  $\theta$  in the power utility case with ambiguity aversion (Case 3.). The special case with power utility and no ambiguity aversion

Table 1: Exogenous Parameters

Variable	Description	Value
$\gamma$	Risk Aversion	5
$\beta$	Pure Time Preference	0.02
$\mu$	Endowment Growth Rate	0.02
$\sigma$	Endowment Volatility	0.03
$\lambda$	Jump Arrival Rate	0.016
$\mu_J$	Jump Size Parameter	-0.1
$\sigma_J$	Jump Size Parameter	0.1

(Case 1.) is obtained by either putting  $\epsilon = \frac{1}{\gamma} = .2$  in subfigure 3a or setting  $\theta = 0$  in subfigure 3b. In this special case, we see that  $r = 10\%$ . So including jump risk already reduces the interest rate with approximately 0.7%. It can be seen that increasing the EIS has the largest effect on the interest rate. The graph is convex, so small increases of  $\epsilon$  from 0.2 already has quite some effect. If  $\epsilon > 1$ , which is in line with the empirical findings of Berkman et al. (2011) that prices go down when disaster risk is increasing, the difference in  $r$  between the power utility and SDU is even more than 6%. Ambiguity aversion also lowers the interest rate, but has a smaller impact than choosing  $\epsilon > 1$ .

Combining both cases gives the subfigures 3c (Contour plot) and 3d (Surface plot) where the interest rate  $r$  as function of both  $\epsilon$  and  $\theta$  is presented. Also here it can be seen that the interest rate is more sensitive to changes in  $\epsilon$  than changes in  $\theta$ . The contour plot shows that for several  $(\epsilon, \theta)$  combinations the model produces quite realistic risk free rates. The interest rate is still above 1% but a combination of  $\epsilon = 1.5$  and  $\theta = 0.01$  yields an interest rate of 2.2%.

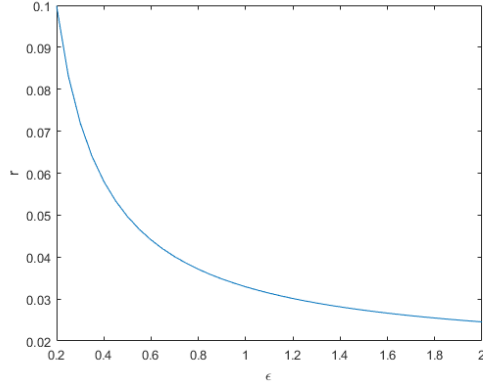
The difference in the value of an uncertain cash-flows can be huge if the risk-free rate is 2% instead of 10%. Climate disasters in the future can be seen as uncertain cash-flows. A widely used measure for the impact of climate change is the social cost of carbon (SCC), which is defined as the present value of the long-term damage done by an extra emission of one ton  $CO_2$ . We will later show in more detail the implications of discount rates on the valuation of climate damages.

## 6.2 Risk premium

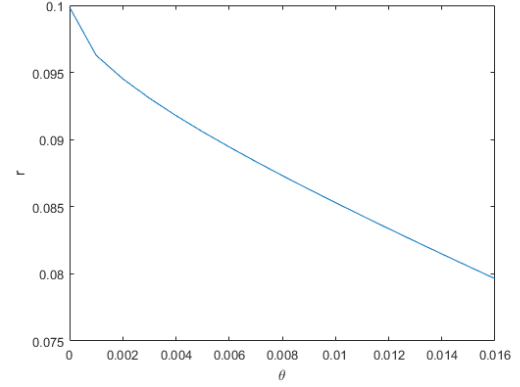
Figure 4 shows the risk premium for different values of  $\theta$ . The risk premium can be split in a diffusion risk and jump risk part. The risk premium does not depend on the EIS, so we obtain cases 1. and 2. by setting  $\theta = 0$ . We see that around two-thirds of the risk premium consists of diffusion risk premium and the total risk premium equals 0.7%. Increasing ambiguity aversion of course increases the risk premium. We see that the model is not able to solve the equity premium puzzle since even with high ambiguity aversion the risk premium is at most around 1.5% which is well below 3%.

So it is clear that climate disasters that arrive on average every 62.5 years with a

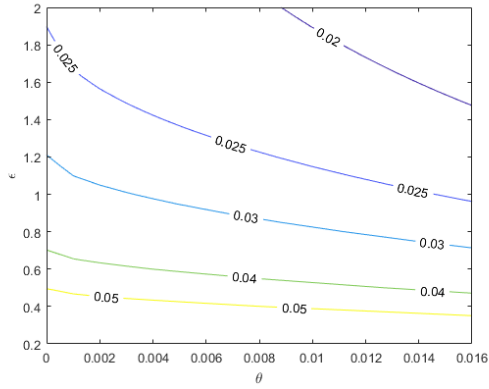
Figure 3: Interest rates for different values of  $\epsilon$  and  $\theta$ .



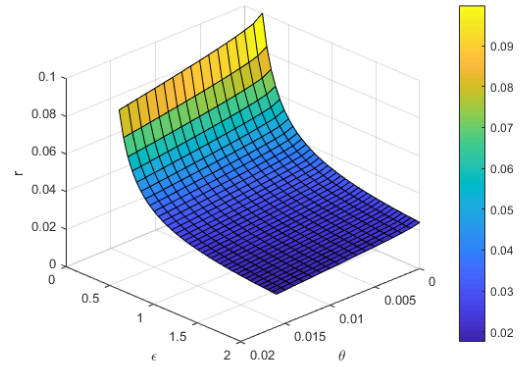
(a) Interest rate as function of elasticity of intertemporal substitution  $\epsilon$ .



(b) Interest rate as function of ambiguity aversion parameter  $\theta$ .



(c) Interest rate as function of both  $\epsilon$  and  $\theta$ , contour plot.



(d) Interest rate as function of both  $\epsilon$  and  $\theta$ , surf plot.

Figure 4: Risk premium as function of ambiguity aversion parameter  $\theta$ .

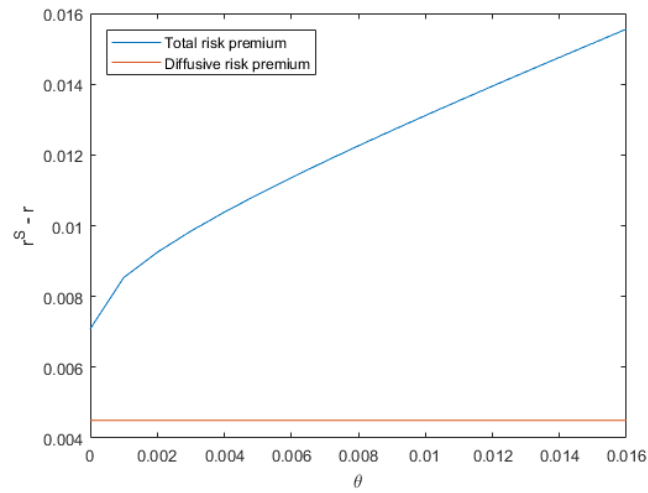
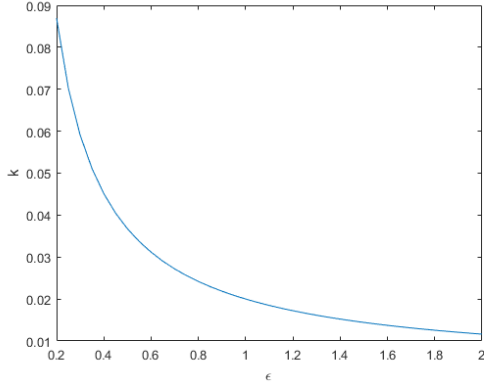
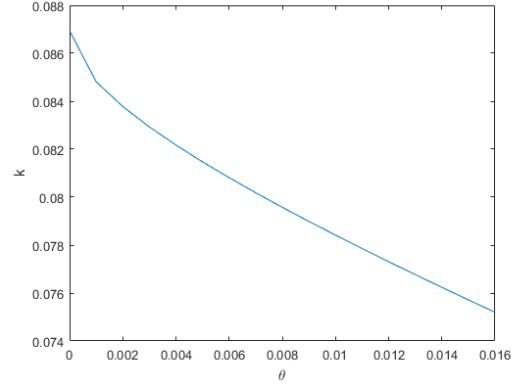


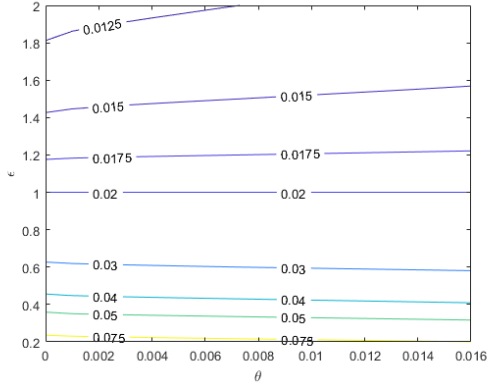
Figure 5: Consumption wealth ratio for different values of  $\epsilon$  and  $\theta$ .



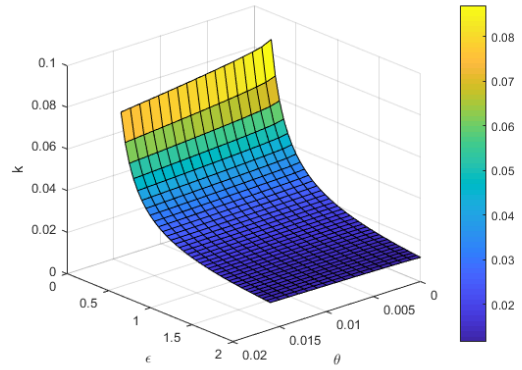
(a) Consumption wealth ratio for different values of  $\epsilon$ .



(b) Consumption wealth ratio as function of ambiguity aversion parameter  $\theta$ .



(c) Consumption wealth ratio as function of both  $\epsilon$  and  $\theta$ , contour plot.



(d) Consumption wealth ratio as function of both  $\epsilon$  and  $\theta$ , surf plot.

mean size of  $-9\%$  are not enough to generate a reasonable equity premium. But it is still likely that we are not going to manage to stay below  $2^\circ\text{C}$ , and more global warming would lead to a higher risk premium. But it is clearly necessary to consider different extensions to explain the equity premium puzzle, climate risk with an arrival rate that is constant over time alone is not enough.

Our model is not set up to perfectly match market data. However, we propose two possible ways to obtain a more realistic equity premium. One way is to introduce another Poisson process that captures damages from non-climate related disasters. One can think of wars or large crises such as the Great Depression. Furthermore, it is possible to introduce leveraged assets in the model (Abel, 1999). Instead of a stock that pays  $C_t$  as dividends, one can introduce assets that pay  $C_t^x$  as dividends with  $x > 1$ . Then the risk premium of those leveraged assets is higher. Leverage is for example used in Wachter (2013) and Bansal and Yaron (2004) to explain the equity premium and risk-free rate puzzles. We will not consider these extensions.

### 6.3 The Consumption wealth ratio

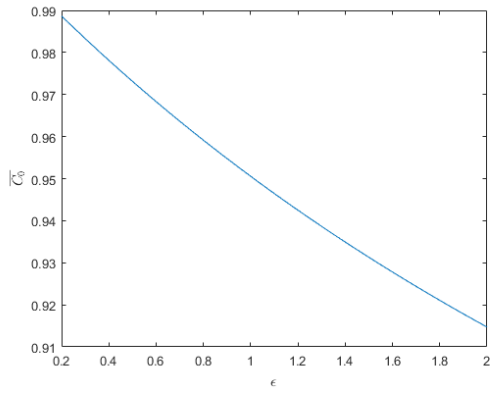
As discussed before, whether jump risk increases or decreases prices depends on the *EIS*. Subfigure 6a shows that the consumption wealth ratio is very sensitive to the choice of  $\epsilon$ . The value of the stock is around 8 times as much for  $\epsilon = 2$  compared to  $\epsilon = 0.2$ . Subfigure 6b illustrates the effect of ambiguity aversion on prices in the power utility case. Compared to the effect of  $\epsilon$ , the effect of ambiguity aversion on prices is small. Subfigures 6c and 6d illustrate a similar result. The surf plot shows that the consumption wealth ratio is insensitive to changes in  $\theta$ . This implies that intertemporal choices are dominant in determining future prices in this case, climate risk is too small for ambiguity aversion to have large effects.

### 6.4 The value of resolving risk

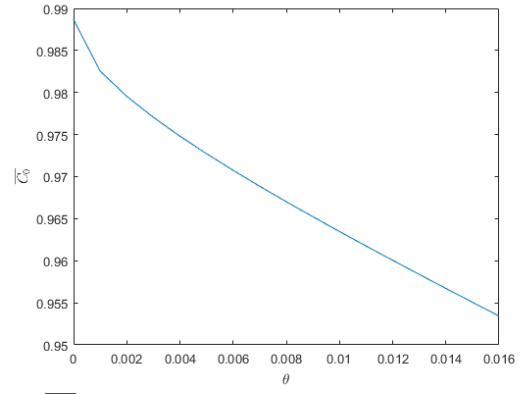
Lastly, let us again consider the example of subsection 5.4. Assume that the representative agent can choose between a world with climate risk ( $\mu_J = -10\%$ ,  $\sigma_J = 10\%$ ) and without climate risk ( $\mu_J = \sigma_J = 0$ ). Then we can calculate what percentage of initial consumption  $C_0$  the agent would be willing to give up if that would get him/her from the case with to the case without climate risk. The results are presented in figure 6. In contrast to the previous subsections, both  $\epsilon$  and  $\theta$  have in magnitude a similar impact. With power utility and without ambiguity aversion, the agent is indifferent for an initial consumption  $\bar{C}_0$  of 98.9%. Increasing the *EIS* from 0.2 to 1.5 already implies that the agent is willing to give up almost 6% extra:  $\bar{C}_0 = 93.1\%$ . If we consider subfigure 6c, we see that the impact of ambiguity aversion is also quite large, choosing  $\theta = 0.01$  leads to  $\bar{C}_0 = 96.3\%$ . Combining SDU with ambiguity aversion shows that the impact of the preferences on  $\bar{C}_0$  can indeed be large. As said before,  $\bar{C}_0 = 98.9\%$  with power utility and without ambiguity aversion, but when we set  $(\epsilon, \theta) = (1.5, 0.01)$  we obtain  $\bar{C}_0 \approx 80.6\%$ , which is quite a difference.

Similarly, it is possible to obtain graphs for the case where climate risk can not be eliminated but only reduced. In that situation, the graphs will have similar shapes but will be flatter. Obviously, in reality it is not possible to (partly) eliminate climate risk within a short time period, but this example does illustrate the sensitivity of the outcomes to different parameter choices.

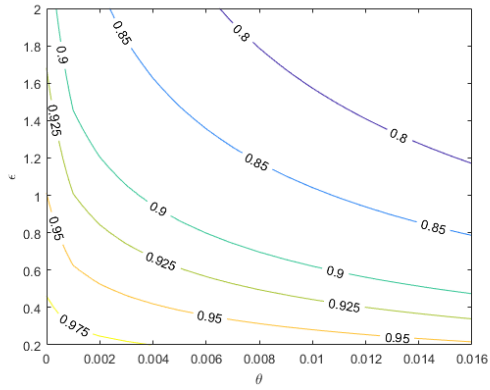
Figure 6: Initial consumption  $\overline{C}_0$  for which the agent is indifferent between  $(\mu_J, \sigma_J) = (-0.1, 0.1)$  with  $C_0 = 1$  and  $(\mu_J, \sigma_J) = (0, 0)$  with  $\overline{C}_0$ .



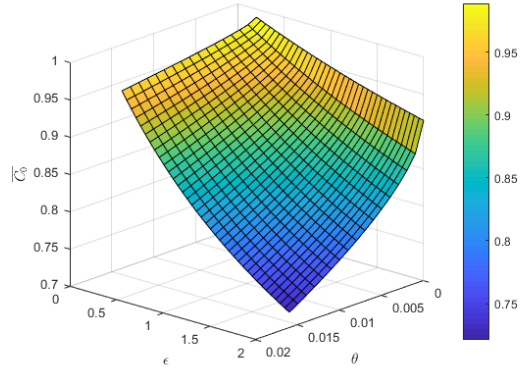
(a)  $\overline{C}_0$  for different values of  $\epsilon$



(b)  $\overline{C}_0$  as function of ambiguity aversion parameter  $\theta$



(c)  $\overline{C}_0$  as function of both  $\epsilon$  and  $\theta$ , contour plot.



(d)  $\overline{C}_0$  as function of both  $\epsilon$  and  $\theta$ , surf plot.

## 7 Climate Change, the social cost of carbon and the structure of preferences

So far we have focused on the asset price implications of disaster risks with a constant arrival rate. This allowed us to explore the dynamics of the model analytically and the consequences of parameter choices, specifically those affecting the structure of preferences. This has helped to form intuition about the likely impact of introducing EZ preferences and ambiguity aversion once we introduce an integrated assessment model or similar climate dynamics modeling, which is what we do in this section where we add a climate sub-model. We make some simplifying assumptions to still be able to obtain analytic solutions of the model. The main requirement is that the state variables of the climate submodel are deterministic, an assumption we will relax in follow up work. The key characteristic of climate change we do focus on is the time dimension: emissions today have effects in the far future. In such a setting several additional dynamics play a role. An increase of atmospheric carbon concentration takes decades or even centuries to be reversed. Furthermore, given an amount of atmospheric carbon concentration, the increase of surface temperature will in turn be a delayed response to the amount of atmospheric carbon concentration. So emissions today will keep warming the earth for decades to come and possible damages will occur into the far future. As a consequence choices about the structure of preferences with respect to time and the dynamics of uncertainty choices play an important role in how these future damages are valued today. Specifically, we will calculate the social cost of carbon (SCC), which is defined as the long-term damage of emitting one ton of carbon emissions today in dollar terms.

We base the climate equations of our model on the widely used DICE model (W. Nordhaus, 2014). Note that although we use a similar climate model as DICE, there are still several differences. In our setup, we use a stochastic endowment economy with climate disasters. Preferences are modeled by multiple priors recursive utility. In this paper, we focus on climate risk, not on expected losses, hence the use of compensated Poisson processes, as we discussed extensively in Sections 4 and 5<sup>10</sup>. So climate change increases risk but in our setup does not change the expected growth of the economy due to the compensation terms in the Poisson processes<sup>11</sup>. Furthermore, we do not explicitly model population growth. Finally our model setup is designed to focus on the benefit side of the cost-benefit analysis. The purpose of the model is to calculate how much one is willing to pay to reduce climate risks. Therefore we only look at the so called Business-As-Usual (no policy measures) scenario but we do not (yet) explicitly incorporate optimal or for that matter suboptimal abatement policies.

The DICE model is formulated in discrete time so we have to transform all their

---

<sup>10</sup>See Shreve (2004), Ch. 11 for precise definitions.

<sup>11</sup>To avoid misunderstanding, we do this because we want to focus on the value of risk reduction, not because we think that expected losses are unimportant in themselves. Thus our value of the SCC should be extended by incorporating the effect of these expected losses to get a more complete assessment. We will show in the last section what the effect is of including the expected losses on the SCC.

equations to adjust them to our continuous time setting. Preferences are modeled using power utility in DICE and population growth is explicitly modeled. DICE uses a production-based economy without uncertainty. Technology grows at an exogenous rate, output is Cobb-Douglas in capital and labour. Every period the representative agent can choose to spend output on investment in capital, consumption and abatement. Damages are quadratically increasing in the temperature and are modeled as a certain reduction of output each period. DICE considers a Business-As-Usual scenario, but also optimal and suboptimal abatement scenarios.

## 7.1 A climate model

Preferences of the agent and the consumption dynamics are similar to and based on what we analysed in the first part of the paper, sections 3-5:

$$dC_t = (\mu - \lambda_t m)C_t dt + \sigma C_t dZ_t + Y_t C_t - dN_t \quad (39)$$

where  $N_t$  is again a Poisson process. The size of the disaster has the same distribution as in the first part of the paper:  $Y_t = e^{W_t} - 1$  where  $W_t$  follows a normal distribution with mean  $\mu_J$  and variance  $\sigma_J^2$ . The main difference is that the arrival rate now is temperature-dependent. Similar to the DICE model we assume that damages are quadratically increasing in temperature:  $\lambda_t = \lambda_T T_t^2$ .

Industrial emissions (from fossil fuel burning) are usually modeled as the product of the carbon intensity of aggregate output and aggregate output (or aggregate consumption) itself. Of course in addition to industrial emissions, land-use change such as deforestation also causes carbon emissions<sup>12</sup>. We simplify the problem by modeling emissions as exogenous, which in the current setting is not all that important because output growth itself is not yet endogenized. Thus we directly model total emissions, which are the sum of industrial emissions and emissions caused by for example land-use change. This simplification is necessary to keep the state variables deterministic, which in turn is necessary for analytical solvability. If we would not make this assumption, emissions are stochastic and this would make it impossible to solve the model analytically. We therefore assume that emissions are growing at a rate  $g_t^E$ . The growth rate itself moves towards the long-run equilibrium  $g_\infty^E$  at a rate  $\delta_{g^E}$ . By assuming that  $g_\infty^E < 0$ , this specification allows us to have growing emissions today, but in the long run the growth rate will then become negative and emissions will go to zero. This is a logical assumption since there is a point where the stock of fossil fuels will be depleted. This gives us the following process for emissions:

$$\begin{aligned} dE_t &= g_t^E E_t dt \\ dg_t^E &= \delta_{g^E} (g_\infty^E - g_t^E) dt \end{aligned} \quad (40)$$

In our setup, it is not a great loss to lose the direct connection between the economy and the carbon emissions since we use a Lucas-tree model where the economy already

---

<sup>12</sup>For an extensive report on the relation between land-use change and emissions we refer to the special IPCC report (Noble, Bolin, Ravindranath, Verardo, & Dokken, 2000).



has an exogenous growth rate. We do not analyse optimal policy and therefore the causes of economic growth and emissions are not of first order importance. What is important for the valuation of the risk is that the climate dynamics of the model are matching the dynamics of other climate models such as DICE. In the next section we show that this specification is able to match the DICE model surprisingly well in spite of the much greater simplicity of our stylized model.

We do admit that specifying emissions as an exogenous process also means that we implicitly assume that there is no correlation between climate risk and the economy. In reality, emissions are low when the economy is in a recession and vice-versa. There is a substantial correlation between economic growth and worldwide carbon emissions. However, due to thermal inertia temperature reacts very slowly on emissions and the contemporaneous correlation between consumption and temperature will be much lower. When climate risk is high in good states, one would be willing to pay less to reduce the risk. So the correlation between climate risk and the consumption process does play a role in the valuation of damages. However, since the correlation between aggregate consumption and climate risk is not very large due to thermal inertia we expect that this does not play a large role.

Following DICE we consider three carbon reservoirs: atmospheric carbon  $M_t^{at}$ , carbon in the upper oceans and biosphere  $M_t^{up}$  and carbon in the deep oceans  $M_t^{lo}$ . Carbon moves between these reservoirs. The parameters  $\delta_{a \rightarrow b}$  capture the rate at which carbon moves from reservoir  $a$  to reservoir  $b$ . The carbon cycle is given by:

$$\begin{aligned} dM_t^{at} &= \left( -\delta_{at \rightarrow up} M_t^{at} + \delta_{up \rightarrow at} M_t^{up} + E_t \right) dt \\ dM_t^{up} &= \left( \delta_{at \rightarrow up} M_t^{at} - (\delta_{up \rightarrow at} + \delta_{up \rightarrow lo}) M_t^{up} + \delta_{lo \rightarrow up} M_t^{lo} \right) dt \\ dM_t^{lo} &= \left( \delta_{up \rightarrow lo} M_t^{up} - \delta_{lo \rightarrow up} M_t^{lo} \right) dt \end{aligned} \quad (41)$$

The next step is to model the impact of carbon concentration on temperature. This requires modeling what is called radiative forcing: radiative forcing is 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. Following [W. Nordhaus \(2014\)](#) we propose a logarithmic relation between atmospheric carbon concentration and radiative forcing.

$$F_t = \kappa_1 \frac{\log(M_t^{at}/M_{pre}^{at})}{\log(2)} \quad (42)$$

When  $M_t^{at} = 2M_{pre}^{at}$ , we obtain  $F_t = \kappa_1$ .  $\kappa_1$  can therefore be interpreted as the equilibrium forcing when the atmospheric carbon concentration is twice the pre-industrial level. Further we include radiative forcing from other greenhouse gases, aerosols, ozone and other factors. We call all these effects together exogenous forcing. The state variable that captures exogenous forcing is assumed to grow over time to the asymptotic level  $EF_\infty$ .

$$dEF_t = \kappa_2 (EF_\infty - EF_t) dt \quad (43)$$

In the DICE model, exogenous radiative forcing is linearly increasing until 2100 and after that it stays constant. In our model, exogenous forcing is also increasing and converges to some pre-specified level. Define total radiative forcing by  $F_t^{tot} = F_t + EF_t$ .

The final step moves from  $F_t^{tot}$  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  $\tau_1$  while  $\tau_2$  equals the equilibrium temperature impact of a doubling of carbon concentration in the atmosphere compared to the pre-industrial level.  $\tau_3$  captures the speed of temperature transfer between the upper layers and the deep layers of the ocean.  $\tau_4$  equals the heat capacity of the deeper layers of the ocean.

$$\begin{aligned} dT_t &= \frac{1}{\tau_1} \left( F_t^{tot} - \frac{\kappa_1}{\tau_2} T_t - \frac{\tau_4}{\tau_3} (T_t - T_t^{oc}) \right) dt \\ dT_t^{oc} &= \frac{1}{\tau_4} \frac{\tau_4}{\tau_3} (T_t - T_t^{oc}) dt \end{aligned} \tag{44}$$

From this equation, one can verify that the long run equilibrium temperature for a given level of radiative forcing equals:  $T_t^{eq} = \frac{\tau_2}{\kappa_1} F_t^{tot}$ . Therefore we can rewrite the first equation to:

$$dT_t = \frac{1}{\tau_1} \left( \frac{\kappa_1}{\tau_2} (T_t^{eq} - T_t) - \frac{\tau_4}{\tau_3} (T_t - T_t^{oc}) \right) \tag{45}$$

This 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  $T_t^{eq}$ .

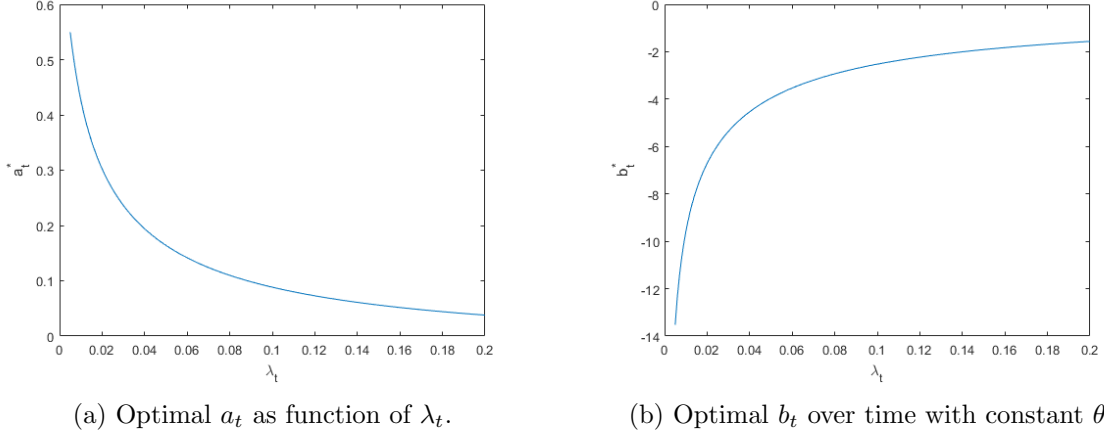
## 7.2 Ambiguity

In section [4.3](#) we derived the first order conditions to calculate the optimal  $a^*$  and  $b^*$ . These two parameters were constant since the ambiguity parameter  $\theta$  and the jump parameters  $\lambda$  and  $m$  were constant as well. However, in the climate setting we assume that  $\lambda_t$  is time-varying. This implies that the optimal parameters  $a_t^*$  and  $b_t^*$  are also time-varying, since the relative entropy is a function of  $\lambda_t$ . Since  $a_t^*$  and  $b_t^*$  are a function of  $\lambda_t$ , they are implicitly a function of temperature  $T_t$  as well.

We again assume that  $\theta$  is constant over time. Figure [7](#) shows the optimal  $a_t^*$  and  $b_t^*$  as a function of  $\lambda_t$ . For each  $\lambda_t$  one finds the corresponding  $a_t^*$  and  $b_t^*$  by solving the first order conditions, similar as in the constant  $\lambda$  case. A constant relative entropy budget implies that  $a_t$  is decreasing in  $\lambda_t$  and  $b_t$  is increasing in  $\lambda_t$ . From now on we define  $\lambda_t^* = e^{a_t^*} \lambda_t$  and  $m_t^* = e^{\mu_J + b_t^* \sigma_J^2 + \frac{1}{2} \sigma_J^2} - 1$ .

The idea behind the decreasing parameters is illustrated in the following example. Assume  $\theta = 0.01$ . At time  $t$ , the arrival rate equals 0.02 and at time  $t'$  the arrival rate equals 0.04. At every time point the following equality must hold at the optimum:  $RE(a_t, b_t, \lambda_t) = \theta$ . For  $\lambda_t = 0.02$  the optimal parameters are  $(a_t^*, b_t^*) = (0.3, -6.7)$  and  $RE(0.3, -6.7, 0.02) = 0.01$ . Now consider time  $t'$  with arrival rate 0.04. If we

Figure 7: Optimal parameters  $a_t$  and  $b_t$  with time-varying arrival rate  $\lambda_t$  and constant ambiguity aversion parameter  $\theta$ . Input parameters:  $\theta = 0.01$ ,  $\sigma = 0.03$ ,  $\mu_J = -0.1$ ,  $\sigma_J = 0.1$ ,  $\gamma = 5$ .



would use the same optimal parameters as at time  $t$ , the relative entropy exceeds the budget:  $RE(0.3, -6.7, 0.04) > 0.01$ . In our setup the relative entropy is increasing in  $\lambda_t$ . For a larger arrival rate, an x% increase in the arrival rate generates a larger relative entropy or ‘distance’ between the two measures. Therefore the optimal  $a_t$  and  $b_t$  must adjust to make sure that the relative entropy remains within the constant budget. At time  $t'$ , the optimal parameters become:  $(a_t^*, b_t^*) = (0.19, -4.53)$ .

### 7.3 Consumption strips

It is typically not possible to solve the partial differential equation of the problem with climate state variables (except when the highly restrictive assumption of a unit EIS is made). 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 use the solution method proposed in Tsai and Wachter (2018); The details are described in appendix D, but we sketch the approach here. Similar to the case with a constant arrival rate, we can derive the pricing kernel for this problem. We then obtain an expression that depends on an unknown function  $g(X_t)$ , where  $X_t = [g_t^E, E_t, M_t^{at}, M_t^{up}, M_t^{lo}, EF_t, T_t, T_t^{oc}]$  is a vector of all the state variables of the climate model. But by substituting the HJB-equation into the pricing kernel we obtain an expression that only depends on known parameters. The step-by-step elaboration is in Appendix D.

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  $s > t$ . Call its value at time  $t$ :  $H_t(C_t, X_t, u)$ , where  $u$  denotes the time to maturity;  $u = s - t$ . We can define a stock that yields an infinite collection of these consumption strips. The value of such a stock then obviously becomes:

$$S_t = \int_0^\infty H(C_t, X_t, u) du \quad (46)$$

The value of aggregate wealth  $W_t$  equals  $S_t$  in equilibrium. Dividing  $C_t$  by  $W_t$  gives the consumption-wealth ratio.

The price of a consumption strip paying out at time  $s > t$  equals:

$$\begin{aligned} H_t &= H(C_t, X_t, u) \\ &= E_t \left[ \frac{\pi_s}{\pi_t} C_s \right] \end{aligned} \quad (47)$$

We can use the fact that every asset multiplied by the pricing kernel must be a martingale to calculate the value of such an asset. We then get for  $H_t$  the following expression:

**Proposition 8.** *The price of a consumption strip at time  $t$  that pays aggregate consumption  $C_s$  and has time to maturity  $u = s - t$  equals:*

$$H_t = C_t \exp \left\{ - \underbrace{ku}_A - \underbrace{(1/\epsilon - 1)}_B \underbrace{\int_t^{t+u} \lambda_s^* \left( \frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_s^* \right) ds}_C \right\} \quad (48)$$

where  $\lambda_t^* = \lambda_T e^{a_t^* T_t^2}$ ,  $T_t = T(X_t)$ ,  $m_t^* = e^{\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}\sigma_J^2}$  and  $k = \beta + (1/\epsilon - 1) \left( \mu - \frac{\gamma}{2}\sigma^2 \right)$ .

*Proof.* See Appendix [D](#) □

Consider first the case without any temperature risk, i.e. with the arrival rate independent of temperature:  $\lambda_T = 0$ ; equation (48) then simplifies to  $H_t = C_t \exp(-ku)$  where  $k$  is the consumption-wealth ratio in the absence of climate risk. When there is no jump risk,  $\lambda_T = 0$  and thus only term  $A$  plays a role in the value of the consumption strip, the terms  $B$  and  $C$  in the exponential function drop out in this particular case. In words, with  $\lambda_T = 0$  the value today of a consumption strip in the future declines exponentially at a rate equal to the consumption-wealth ratio, which can thus be interpreted as the discount rate for the risky consumption strip. Under the plausible assumption that  $\mu - \frac{\gamma}{2}\sigma^2 > 0$ , the consumption-wealth ratio is decreasing in  $\epsilon$ , and therefore the price of the consumption strip is increasing in  $\epsilon$ .

Assume now that temperature risk does play a role; the value of the consumption strip then becomes state dependent. The additional two terms in (48) have a clear interpretation. Term  $B$  determines whether climate risk, whatever the magnitude of its impact, leads to a higher or lower consumption-to-wealth ratio. See below for further discussion.

Term  $C$  captures the effect of climate damages on the consumption strip. Total climate risk is determined by the disaster arrival rate  $\lambda_t^*$  and the expected jump size. It is

important to note that the Poisson process that triggers the climate disasters is compensated, so the expected value of climate disasters is zero. In our model, disasters do not damage the economy on average (because of the compensation term in the drift) but rather increase the variability of the aggregate consumption process. If we would not (or only partially) compensate the Poisson process, such that on average disasters have a negative impact, the social cost of carbon will obviously become commensurately larger.  $m_t^*$  equals the expected jump size under  $\mathbb{Q}^{a^*, b^*}$ .  $\frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma}$  can be interpreted as the certainty equivalent of the climate shock. The difference between the actual expectation and the certainty equivalent determines how damages are valued. Term  $C$  is equal to zero when the agent is risk-neutral. Since the disasters are compensated, climate related risks do not matter for the agent if he/she is risk neutral. Temperature plays a role through the arrival rate  $\lambda_t^*$ . It is important to note that not only  $T_{t+u}$  matters for the price of the consumption strip that pays off at time  $t + u$ , but the whole path of  $T_s$  between  $t$  and  $t + u$ . All climate disasters that possibly occur within this period affect aggregate consumption at time  $t + u$ . Therefore the integral over the damages determines the magnitude of climate risk. The other climate state variables and parameters do not show up in the formula but do influence  $T_t$  and therefore indirectly also influence asset prices.

As discussed in the previous paragraph, the term  $C$  determines the impact of total climate risk on the value of a consumption strip. In section 5.3 we already saw that how risk affects prices ultimately depends on whether  $\epsilon$  is larger or smaller than one. We see a similar result here. When  $\epsilon > 1$ , term  $B$  is negative which implies that climate risk decreases the price of the consumption strip. Now consider what happens if  $\epsilon$  is increased. On the one hand, this implies that term  $B$  becomes more negative, so that the same amount of climate risk has a more negative impact on prices. On the other hand,  $k$  is decreasing in  $\epsilon$ <sup>13</sup> and therefore term  $A$  will be smaller. The latter effect normally dominates, which gives the intuitive result that the price of the consumption strip is increasing in the elasticity of intertemporal substitution.

When  $\epsilon < 1$ , term  $B$  is positive and additional risk increases asset prices. Increasing  $\epsilon$  has a different effect now. Term  $B$  will become smaller, which implies that a certain amount of climate risk increases the price by a smaller amount.  $k$  is again decreasing in  $\epsilon$ . Overall, the effect of  $\epsilon$  on  $k$  will normally again dominate and therefore  $H_t$  is increasing in this situation as well. If  $\epsilon = 1$ , the price of the consumption strip reduces to  $H_t = C_t \exp(-\beta u)$ . Risk plays no role in the price, since the interest rate effect and the risk effect cancel out. In that case the price of the consumption strip in the future is simply the value of the consumption strip today discounted with the pure rate of time preference  $\beta$ .

We already defined an asset paying a continuous stream of dividends  $C_t$  at time  $t$  as a stock with price  $S_t$ . Aggregate wealth must be equal to the total value of all such stocks in equilibrium. Denote aggregate wealth by:  $W(C_t, X_t)$ . The state-dependent consumption-wealth ratio therefore equals the inverse of the Wealth-to-Consumption ratio  $S_t$ :

---

<sup>13</sup>This is the case for plausible parameter values, i.e. if  $\mu - \frac{\gamma}{2}\sigma^2 > 0$ .

$$k(X_t) = \frac{C_t}{W(C_t, X_t)} = \frac{C_t}{\int_0^\infty H(C_t, X_t, u) du} \quad (49)$$

When there is no climate risk ( $\lambda_T = 0$ ), the expression in equation (49) simplifies to  $k(X_t) = \left( \int_0^\infty \exp(-ku) du \right)^{-1} = k$ .

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 condition that  $f_C = V_W$ . Furthermore, we conjecture

$$V(C_t, X_t) = \frac{g(X_t)C_t^{1-\gamma}}{1-\gamma} \quad (50)$$

Using the chain rule we get:

$$\begin{aligned} V_W &= V_C \frac{\partial C}{\partial W} = V_C k(E_t, M_t, \tilde{T}_t) \\ &= g(X_t) C_t^{-\gamma} k(X_t) \end{aligned} \quad (51)$$

Also we have for the intertemporal aggregator:

$$f_C = \beta g(X_t)^{\frac{1/\epsilon - \gamma}{1-\gamma}} C_t^{-\gamma} \quad (52)$$

Together this gives us:

$$g(X_t) = \left( \frac{k(X_t)}{\beta} \right)^{-\frac{1-\gamma}{1-1/\epsilon}} \quad (53)$$

Inserting this expression for  $g(X_t)$  into equation (50) then yields the expression for the value function.

## 7.4 Social cost of carbon

Given the value function, we can calculate the social cost of carbon, which we define 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 to obtain the social cost of carbon in "dollar" terms (i.e. in terms of the price of time  $t$  consumption units). Simple differentiation of equation (50) gives:

$$\begin{aligned} SCC_t &= -\frac{\partial V_t / \partial E_t}{\partial V_t / \partial C_t} = -\frac{\frac{\partial}{\partial E_t} g(X_t)}{(1-\gamma)g(X_t)} C_t = -\frac{\frac{\partial}{\partial E_t} \left( \frac{k(X_t)}{\beta} \right)^{-\frac{1-\gamma}{1-1/\epsilon}}}{(1-\gamma) \left( \frac{k(X_t)}{\beta} \right)^{-\frac{1-\gamma}{1-1/\epsilon}}} C_t = -\frac{C_t}{1/\epsilon - 1} \frac{\frac{\partial}{\partial E_t} k(X_t)}{k(X_t)} \\ &= -C_t \int_0^\infty \int_t^{t+u} \underbrace{\frac{\partial}{\partial E_t} (\lambda_s^*)}_A \underbrace{\left( \frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_s^* \right)}_B ds \underbrace{\left( \frac{H(C_t, X_t, u)}{W_t} \right)}_C du \end{aligned} \quad (54)$$

We will first discuss the general formula and then the implications of different preferences.

Equation (54) shows that the social cost of carbon is proportional to  $C_t$ , the aggregate consumption level: when the current aggregate consumption level doubles, the SCC doubles as well. For given consumption level, the SCC depends on the product of three terms, labeled  $A$ ,  $B$  and  $C$  respectively. Term  $A$  captures the marginal effect of emitting one unit of carbon at time  $t$  on the arrival rate  $\lambda_s = \lambda_T e^{a_s^*} T_s^2$  at time  $s$ . This term implicitly depends on the other state variables of the climate model through the temperature  $T_t$ . It takes into account the fact that emitting one ton of carbon at time  $t$  has effects on the temperature in the far future.

This marginal effect of one unit of carbon emissions on the arrival rate is then multiplied by term  $B$ ,  $\left( \frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_s^* \right)$ , which captures how the damages are valued. This term can be seen as the difference between the expected damage of a disaster and the certainty equivalent. The total climate risk over the time period  $t$  to  $t + u$  is then obtained by integrating the product of term  $A$  and  $B$  over this time period. This can be seen as the cumulative additional climate risk due to emitting one unit of carbon at time  $t$ . For each time to maturity  $u$ , this cumulative additional climate risk is then multiplied by the ratio of the price of the consumption strip  $H_t(C_t, X_t, u)$  strip paying out at time  $u$  over total wealth  $W_t$  (term  $C$ ). Since  $W_t = \int_0^\infty H_t(C_t, X_t, u) du$ , this fraction determines the weight of each maturity  $u$ . The weights integrate to one. The relative importance of each maturity  $u$  is determined by the derivative of  $H_t$  with respect to  $u$ . When  $H_t$  is declining very fast in  $u$ , much weight is on the short term, which are the periods where climate risk is small. When  $H_t$  declines slowly, more weight is on the long maturities which will therefore also lead to a higher SCC.

Now consider the impact of preferences on the  $SCC$ . Once again we consider four polar cases: (1) Power Utility without ambiguity aversion, the benchmark case; (2) SDU, but still no ambiguity aversion; (3) Back to Power Utility, but now with Ambiguity Aversion; and (4) Both SDU and Ambiguity Aversion.

### 1. Power utility, no ambiguity aversion

$$SCC_t = -C_t \int_0^\infty \int_t^{t+u} \frac{\partial}{\partial E_t}(\lambda_s) \left( \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m \right) ds \left( \frac{H_1(C_t, X_t, u)}{\int_0^\infty H_1(C_t, X_t, u) du} \right) du$$

where

$$H_1(C_t, X_t, u) = C_t \exp \left\{ -k_1 u - (\gamma - 1) \int_t^{t+u} \lambda_s \left( \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m \right) ds \right\},$$

$$k_1 = \beta + (\gamma - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 \right) \text{ and } \lambda_t = \lambda_T T_t^2.$$

### 2. SDU utility, no ambiguity aversion

$$SCC_t = -C_t \int_0^\infty \int_t^{t+u} \frac{\partial}{\partial E_t}(\lambda_s) \left( \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m \right) ds \left( \frac{H_2(C_t, X_t, u)}{\int_0^\infty H_2(C_t, X_t, u) du} \right) du$$

where

$$H_2(C_t, X_t, u) = C_t \exp \left\{ -k_2 u - (1/\epsilon - 1) \int_t^{t+u} \lambda_s \left( \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m \right) ds \right\},$$

$$k_2 = \beta + (1/\epsilon - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 \right) \text{ and } \lambda_t = \lambda_T T_t^2.$$

### 3. Power utility, ambiguity aversion

$$SCC_t = -C_t \int_0^\infty \int_t^{t+u} \frac{\partial}{\partial E_t}(\lambda_s^*) \left( \frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_s^* \right) ds \left( \frac{H_3(C_t, X_t, u)}{\int_0^\infty H_3(C_t, X_t, u) du} \right) du$$

where

$$H_3(C_t, X_t, u) = C_t \exp \left\{ -k_3 u - (\gamma - 1) \int_t^{t+u} \lambda_s^* \left( \frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_s^* \right) ds \right\},$$

$$k_3 = \beta + (\gamma - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 \right) \text{ and } \lambda_t^* = e^{a_t^*} \lambda_T T_t^2.$$

### 4. SDU utility, ambiguity aversion

$$SCC_t = -C_t \int_0^\infty \int_t^{t+u} \frac{\partial}{\partial E_t}(\lambda_s^*) \left( \frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_s^* \right) ds \left( \frac{H_4(C_t, X_t, u)}{\int_0^\infty H_4(C_t, X_t, u) du} \right) du$$

where

$$H_4(C_t, X_t, u) = C_t \exp \left\{ -k_4 u - (1/\epsilon - 1) \int_t^{t+u} \lambda_s^* \left( \frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_s^* \right) ds \right\},$$

$$\text{and } k_4 = \beta + (1/\epsilon - 1) \left( \mu - \frac{\gamma}{2} \sigma^2 \right).$$



Before we discuss the first case, it is useful to consider the following proposition on the relation between  $\epsilon$  and the maturity weighting functions defined in proposition 9 below.

**Proposition 9.** *Consider two cases. In case one we assume that the  $EIS = \epsilon_1$  and in the second case we assume  $EIS = \epsilon_2 > \epsilon_1$ . Then there exists a maturity  $u^*$  for which the maturity weighting functions  $\frac{H(C_t, X_t, u)}{\int_0^\infty H(C_t, X_t, u)}$  in both cases are equal. For all  $u < u^*$ , the weighting function in case 1 is larger. For  $u > u^*$ , the weighting function in case 2 is larger.*

*Proof.* See appendix E. □

Consider the case of power utility (case 1). Assume that  $\gamma > 1$ . An increase in risk aversion then implies that  $\left( \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m \right)$  becomes more negative which leads to a higher SCC. On the other hand, a larger  $\gamma$  also leads to a smaller  $EIS$ . Proposition 9 shows that a decrease of the  $EIS$  leads to a shift of weights from large maturities to small maturities. The additional damage of emitting one ton of carbon today is increasing with the maturity  $u$ , the largest damages will occur in the (far) future. However, with a low  $EIS$  the weight of a maturity  $u$  is very close to zero in the (far) future. So there is not much weight at the time period where the large damages occur and therefore the integral will not be very large. To sum up, power utility  $\gamma$  captures both the risk aversion and willingness to substitute over time. A higher  $\gamma$  yields a higher risk aversion, which leads to a higher social cost of carbon. However, this also implies that the  $EIS$  goes down, which yields a lower weight on maturities in the far future and therefore the climate damages in the future will not have a high value. These two effects work in opposite direction which explains why the SCC will not be very large when considering power utility for all reasonable choices of  $\gamma$ .

SDU utility enables the disentangling of risk aversion and the elasticity of intertemporal substitution (case 2.). Increasing risk aversion again makes  $\left( \frac{e^{(1-\gamma)(\mu_J + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m \right)$  more negative. Risk aversion thus leads to a higher discrepancy between the expected damage and the certainty equivalent. Risk aversion also has an effect on  $k_2$ , but since  $\sigma^2$  is small this effect is almost negligible. Therefore the  $SCC$  is increasing in risk aversion. The  $EIS$  now only plays a role in the weighting function  $\frac{H_2(C_t, X_t, u)}{\int_0^\infty H_2(C_t, X_t, u) du}$ . When  $\epsilon$  increases, weights shift from short-term maturities to long-term maturities. Since climate damages are larger in the future, putting more weight on the long-term maturities leads to a higher SCC, so a higher  $\epsilon$  given  $\gamma$  leads to a higher SCC.

To analyse the cases with ambiguity aversion, we need the following proposition.

**Proposition 10.** *Consider two cases. In case one we assume that ambiguity aversion  $\theta = \theta_1$  and in the second case we assume  $\theta = \theta_2 > \theta_1$ . Then there exists a maturity  $u^*$  for which the maturity weighting functions  $\frac{H(C_t, X_t, u)}{\int_0^\infty H(C_t, X_t, u)}$  in both cases are equal. If the  $EIS < 1$ , we have that for all  $u < u^*$ , the weighting function in case 1 is larger. For  $u > u^*$ , the weighting function in case 2 is larger. When  $EIS > 1$ , the opposite is true. For  $u < u^*$ , the weighting function in case 2 is larger and for  $u > u^*$  the weighting function in case 1 is larger.*

*Proof.* See appendix [E](#). □

Assume that again the agent has power utility preferences, but he is ambiguity averse (case 3.). We assume that  $\gamma > 1$ , which directly implies that  $EIS < 1$ . Ambiguity aversion again influences both the climate risk and the weighting function. Since  $\lambda_t^*$  is increasing in  $\theta$  and  $\left(\frac{e^{(1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_t^*\right)$  is decreasing in  $\theta$ , ambiguity aversion increases climate risk. Furthermore, from proposition [10](#) we derive that increasing  $\theta$  implies that more weight will be on the long-term maturities. Both effects lead to a higher SCC and therefore the SCC is increasing in  $\theta$ .

In the last case both ambiguity aversion and SDU are present (case 4.). Increasing the  $EIS$  leads to the same (qualitative) effects as we derived in case 2. The effect of ambiguity aversion now depends on the  $EIS$ . Irrespective of the  $EIS$ , ambiguity aversion still increases  $\lambda_t^*$  and it also increases the discrepancy between the expected disaster size and the risk- and ambiguity-adjusted disaster size. This climate risk effect leads to a higher SCC. However,  $\theta$  also affects the weighting function. When  $EIS < 1$ , we are in the same situation as case 3. A higher  $\theta$  leads to more weight on long-term maturities which in turn leads to a higher SCC. When  $EIS = 1$ , ambiguity has no effect on the weighting function since the price of a consumption strip is just  $H_t = C_t \exp(-\beta u)$ . If  $\epsilon > 1$ , several signs change. Increasing  $\theta$  now leads to more weight on short-term maturities. These are the maturities where the damages of emitting a unit of carbon are smaller. The effect of ambiguity on the weighting function therefore leads to a lower SCC. Therefore increasing ambiguity aversion has two offsetting effects in this case.

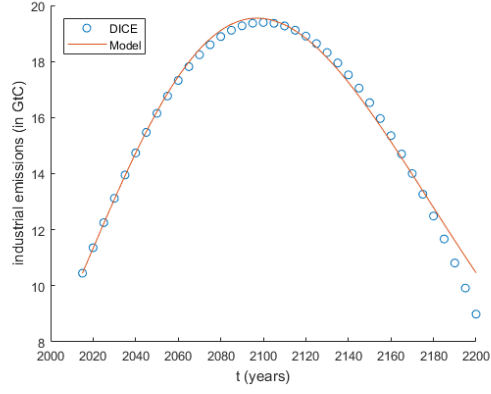
## 8 Climate Change, the social cost of carbon and the structure of preferences: numerical results

### 8.1 Calibration

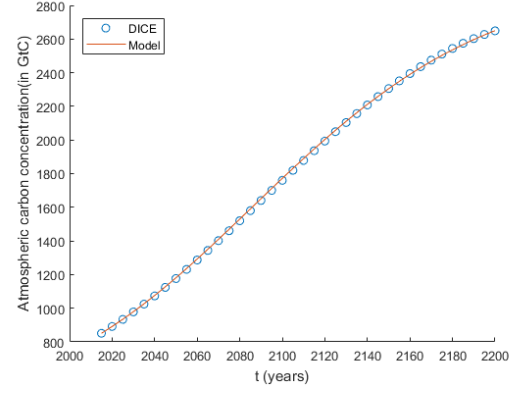
In appendix [F](#), the calibration of the climate model is given. Several parameters are based on the DICE-2016 calibration. Since the DICE model uses a time step of five years, several parameters are calibrated on that time step. In our model, it is most intuitive to express growth and decline rates per year. Therefore we transform all parameters to yearly parameters. Since our model uses a continuous time framework, there will still be some differences between continuous and discrete compounding. The parameters that are not obtained from DICE are calibrated to match the DICE model.

Figure [8](#) shows the results of the comparison of our climate model with the DICE model. We compare our results with the base case of the DICE model, which can be seen as the Business-As-Usual case. Except from the emissions specification, our model replicates the DICE climate dynamics. Figure [8a](#) shows that the exogenous emissions in our model are very similar to the emissions in the DICE model. The emissions peak at the end of the century. Given these emissions, also the other climate variables are almost identical to the DICE model.

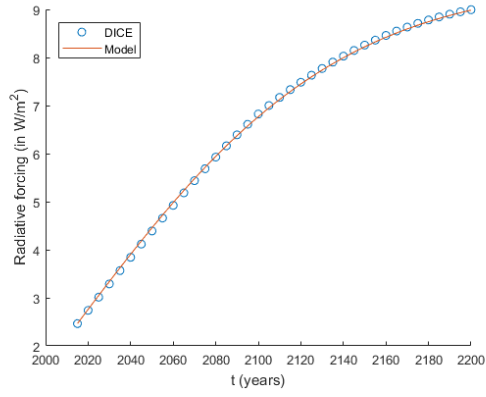
Figure 8: Trajectories of our climate model and the damage calibration compared with the DICE model.



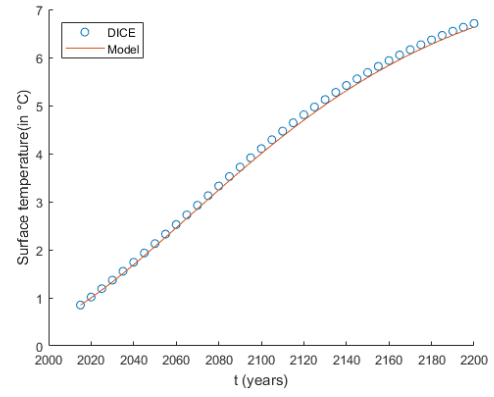
(a) Industrial emissions



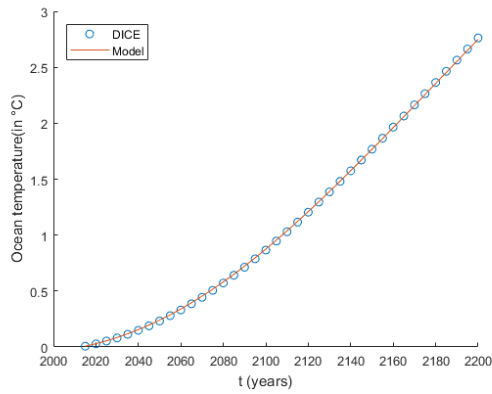
(b) Atmospheric carbon concentration



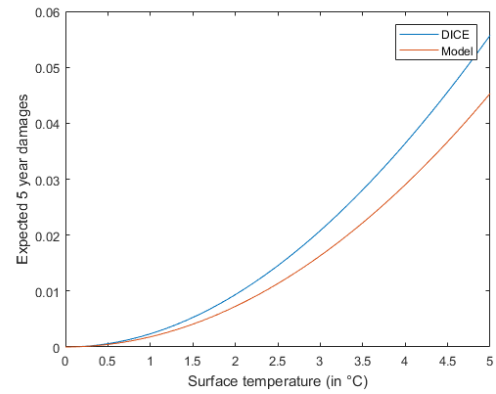
(c) Radiative forcing



(d) Surface temperature



(e) Ocean temperature



(f) Expected damages over a period of 5 years

Table 2: Parameters for the economic model

Par.	Description	Value
$C_t$	Initial consumption level (in trillion \$)	105
$\mu$	Growth rate of consumption (per year)	0.02
$\sigma$	Volatility of consumption (per year)	0.03
$\gamma$	Risk aversion	5
$\epsilon$	Elasticity of substitution	1.5
$\theta$	Ambiguity aversion	0.01
$\beta$	Pure rate of time preference	0.02
$\lambda_T$	Arrival rate parameter (per year)	0.004
$\mu_J$	Disaster size parameter	-0.1
$\sigma_J$	Disaster size parameter	0.1
$m$	Expected disaster size	-0.09

The calibration of the economic parameters is given in table 2. The expected disaster size and the arrival rate parameter of disasters are also chosen to match the DICE model. We assume that the expected disaster size  $m$  is -9% and that the arrival rate parameter  $\lambda_T$  equals 0.04. This corresponds to an arrival rate of 0.0029 today, but the arrival rate increases to 0.016 for  $T_t = 2^\circ C$  and 0.1 for  $T_t = 5^\circ C$ . To be able to compare these numbers to the DICE model we look at the expected disaster damages over a period of 5 years. This specification yields expected disaster damages of 0.7% per 5 years for a temperature increase of 2 degrees and 4.5% for a temperature increase of 5 degrees. For a given temperature level, the expected damages over a period of 5 years are equal to  $5\lambda(T_t)m = 5\lambda_T T_t^2 (e^{\mu_J + \frac{1}{2}\sigma_J^2} - 1)$ . A comparison between our expected damages and the damages of the DICE model are given in figure 8.

Further we set risk aversion equal to 5 and the EIS equal to 1.5. An EIS of 1.5 is common in the literature for Epstein-Zin preferences. Generally there is strong empirical evidence of an EIS larger than one (Van Binsbergen et al. (2012); Vissing-Jørgensen and Attanasio (2003)). Our risk aversion can even be seen as conservative.<sup>14</sup> However, it must be noted that these parameters are often calibrated to match market data (risk premia, interest rates). Ambiguity aversion also lowers the interest rate and increases the risk premium of the stock. We set the ambiguity aversion parameter equal to 0.01. We use this value for our base calibration, but we show that ambiguity aversion is of second order importance for the social cost of carbon. The optimal  $a_t^*$  and  $b_t^*$  were already illustrated in figure 2. This implies that today  $\lambda_t^* = 1.82\lambda_t = 0.0053$  and  $m_t^* = -0.2376$ . For  $T_t = 2^\circ C$  and  $T_t = 5^\circ C$  we obtain respectively  $(\lambda_t^*, m_t^*) = (0.0225, -0.1568)$  and  $(\lambda_t^*, m_t^*) = (0.1092, -0.1133)$ . We use these estimates for the base calibration, but will later present the social cost of carbon for a whole range of values for  $\epsilon$  and  $\theta$  to illustrate the effect of preferences on the SCC. This calibration yields an equilibrium interest rate of 2.5% today and a risk premium of 1% today.<sup>15</sup> So climate risk and ambiguity are not enough to solve the

<sup>14</sup>For an overview of several parameter choices in the literature, see Cai et al. (2015) Ch. 3.3.

<sup>15</sup>The formulas for the time dependent interest rate  $r_t$  and the risk premium  $r_t^S - r_t$  are in

equity premium puzzle in this setup. Note that the interest rate is decreasing over time and that the risk premium is increasing over time since climate risk is small today but will increase in the future. DICE uses a risk aversion parameter of 1.45, but the main reason for this low value is that the risk aversion also determines the EIS in the DICE model. The pure rate of time preference equals 0.02, which can be considered conservative compared to the DICE model which uses  $\beta = 0.015$ .

## 8.2 Social cost of carbon

Our base calibration yields a social cost of carbon of \$1256 per ton of carbon (\$ 342 per ton of  $CO_2$ ). In what follows all numbers for the social cost of carbon will be expressed in dollars per ton carbon <sup>16</sup>. Figure 9 shows the social cost of carbon for different choices of  $\epsilon$  and  $\theta$ . The figure shows that the SCC is very sensitive to the choice of  $\epsilon$ , but much less sensitive to the ambiguity aversion parameter  $\theta$ . When we would assume that the representative investor has power utility preferences, the SCC is much lower: \$13. This is about 100 times less than it is in our base calibration. The social cost of carbon is most sensitive to the choice of the EIS. We saw in the previous section that the social cost of carbon takes into account all future damages. How future damages are valued depends ultimately on the discount rate, it is clear from our analysis that the EIS plays a large role in the effective discount rate.

Figure 9 confirms that the effect of ambiguity aversion on the SCC depends on the EIS. When  $\epsilon < 1$ , the SCC is increasing in ambiguity aversion. When  $\epsilon = 1/\gamma = 0.2$ , choosing  $\theta = 0.01$  doubles the SCC compared to power utility. However, when  $\epsilon > 1$  as discussed two offsetting effects play a role. On the one hand ambiguity increases the discrepancy between the expected value and the certainty equivalent of the damages, but on the other hand it also influences the weighting function that plays a role in the social cost of carbon. A high ambiguity aversion in combination with  $\epsilon > 1$  leads to more weight on short term maturities which in turn leads to a lower SCC. For  $\epsilon = 1.5$  we see that the SCC is actually decreasing in  $\theta$ . Overall we can conclude that ambiguity aversion has a second order effect only on the social cost of carbon.

Overall, these numbers are substantially higher than the estimates of the social cost of carbon from the DICE model. W. D. Nordhaus (2017) estimates the social cost of carbon in the DICE-2016R model to be equal to \$114. Our model directly illustrates the reason for this lower value. With power utility the EIS is calibrated way below the value that is empirically supported and this in turn generates a very low social cost of carbon. The reason is that discount rates are very large when considering power utility for plausible parameter values. The fact that damages are modeled as disaster risk instead of reducing the output by a certain amount every time period also generates a larger SCC.

Several IAMs find that effect of the risk aversion parameter  $\gamma$  on the SCC is of second order importance as well when considering Epstein-Zin preferences. But we find that the SCC is very sensitive to the choice of  $\gamma$ . The reason for the difference is

---

Appendix D.

<sup>16</sup>To obtain the SCC in dollars per ton  $CO_2$ , the SCC per ton of carbon has to be divided by 3.67.

that we model climate risk by increasing the variability of the consumption process instead of decreasing the level or growth rate of output. The social cost of carbon is zero when  $\gamma = 0$  and it is therefore very dependent on the choice of  $\gamma$ . Figure 10 shows the effect of combinations of  $(\gamma, \epsilon)$  on the SCC when there is no ambiguity aversion.

### 8.3 The Total SCC: Adding in Expected Losses

In this paper, the focus is on climate risk, not on the expected loss component of the SCC. This approach is chosen to illustrate that it actually matters a lot whether damages are stochastic or deterministic. In order to focus on risk, the jump process is compensated such that on average climate disasters do not decrease economic growth. In reality of course there will be both a risk component and an expected loss component in the SCC: climate change does have an impact on the growth of the economy. It is relatively straightforward to solve our model without the compensation term if we assume that the agent is not ambiguity averse ( $\theta = 0$ ). If the agent would be ambiguity averse, the relative entropy changes when the compensation term is left out. This changes the dynamics of the model and is beyond the scope of this paper.

To add in the expected loss component, we solve exactly the same problem, with the following exogenous endowment stream:

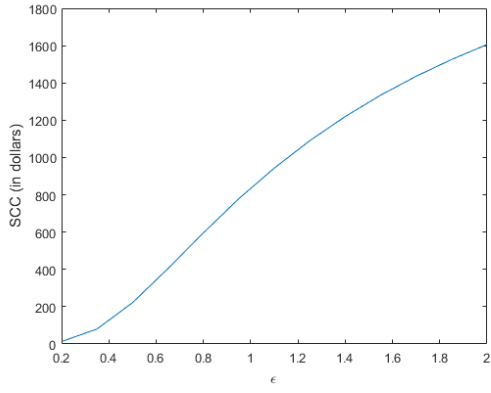
$$dC_t = \mu C_t dt + \sigma C_t dZ_t + Y_t C_{t-} dN_t \quad (55)$$

Calculating the social cost of carbon with our base calibration from table 2 (with  $\theta = 0$  instead of  $\theta = 0.01$ ) yields a SCC of \$2178 per ton of carbon. So if we relax the compensation assumption and include expected losses, the SCC almost doubles.

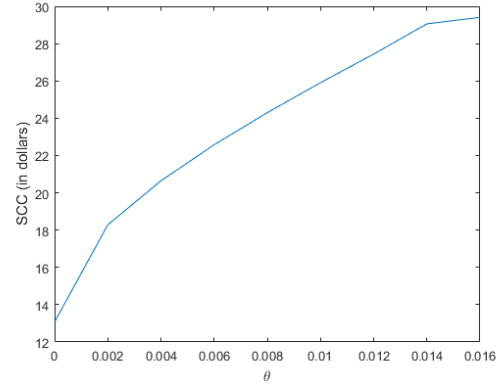
## 9 Conclusions

Climate change will beyond reasonable doubt have a large impact on economic growth in the future. 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, but that these changes will happen far into the future, while mitigating policies are under consideration right now. That discrepancy puts 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 future damages with costs today those future damages need to be discounted back towards today. The debate 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 squarely focus on the discounting question, but we take a different approach. Rather than discussing numerical values of certain parameters given the standard discounting framework (time separable expected utility

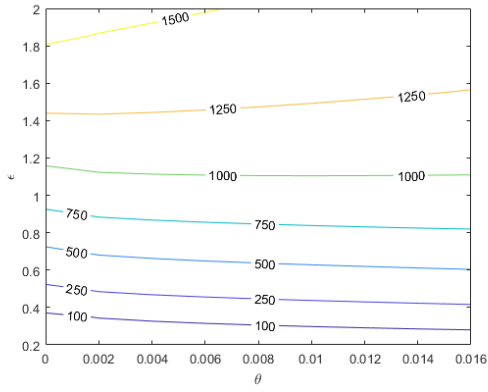
Figure 9: Social cost of carbon in dollars per ton carbon for different values of  $\epsilon$  and  $\theta$ .



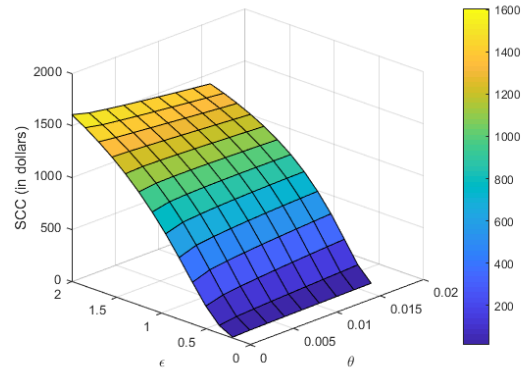
(a) Social cost of carbon as function of  $\epsilon$ .



(b) Social cost of carbon as function of ambiguity aversion parameter  $\theta$ .

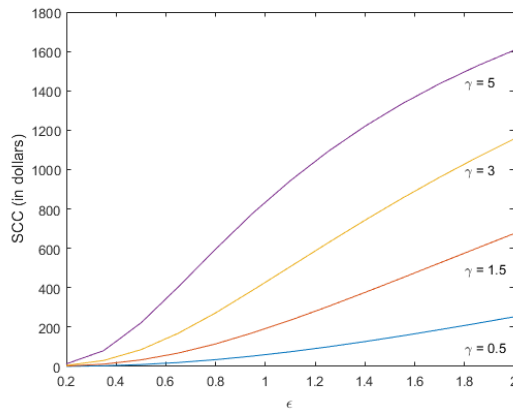


(c) Social cost of carbon as function of both  $\epsilon$  and  $\theta$ , contour plot.



(d) Social cost of carbon as function of both  $\epsilon$  and  $\theta$ , surf plot.

Figure 10: Social cost of carbon as function of  $\epsilon$  for different values of the risk aversion parameter  $\gamma$  (no ambiguity).





maximization, or what we call power utility), and its rigid link between the degree of risk aversion ( $\gamma$ ) and (the inverse of) the intertemporal substitution elasticity, we explore alternative specifications of preferences, and will report SCC values orders of magnitudes larger than obtained using the conventional discounting framework.

We choose to focus on the effect of Epstein-Zin recursive preferences on outcomes of the model and on the presence of unmeasurable risk ("ambiguity"). Both the extensions of stochastic differential utility (the continuous time implementation of Epstein-Zin preferences), breaking the link between  $\gamma$  and the EIS and 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) is very likely to play an important role. 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 then is a natural framework (unmeasurable risk is often referred to as "Knightian uncertainty").

In the first part of the paper we propose a disaster asset pricing model and present closed-form solutions for interest rates, risk premia and wealth consumption ratios. The solutions show that the effect of the two preference extensions on interest rates and wealth consumption ratios is substantial. Especially the effect of different EIS-choices is large. These outcomes give us intuition in how these variables interact with the valuation of damages, which is useful for the second part of the paper. In that second part, we extend the asset pricing approach of the first part by adding a climate risk model designed to mimic the dynamics of much larger climate models.

We then propose an analytic IAM by extending the disaster risk model with a climate change model and a temperature dependent arrival rate. This IAM is able to capture the climate dynamics of the numerical IAMs very well. 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 due to the closed form solutions for the social cost of carbon. Where stochastic numerical IAMs can take hours to be solved, solving this model only requires numerical integration and is therefore solved within seconds.

Our base calibration generates a substantial social cost of carbon of \$1256 per ton of carbon (which increases to \$2178 per ton of carbon once expected losses are added in). This is an order of magnitude higher than for example the estimate of \$114 that is obtained using the DICE-2016R model (W. D. Nordhaus, 2017). Our model highlights how the calibration of preference parameters changes the SCC. The elasticity of intertemporal substitution and risk aversion are the dominant parameters in determining this amount. The effect of ambiguity aversion depends on the elasticity of intertemporal substitution and turns out to be of second order importance.

With power utility, the parameter  $\gamma$  has two offsetting effects on the social cost of carbon. On the one hand, increasing risk aversion leads to a larger SCC. However, since  $\gamma$  equals the inverse of the elasticity of intertemporal substitution as well, a high  $\gamma$  also leads to a representative agent that is focused on the short-term. Long-term damages have little effect on utility and will lead to a low SCC. When the restriction of  $\gamma = \frac{1}{EIS}$  is released, both increasing  $\gamma$  and the *EIS* lead to a higher SCC. Introducing



Epstein-Zin utility with an  $EIS > 1$ , in line with empirical evidence, therefore leads to a much larger value for the SCC.

Analysing the effect of ambiguity aversion on the SCC is a bit more involved since multiple effects play a role. Ambiguity aversion has both an effect on the certainty equivalent (more ambiguity aversion leads to a lower certainty equivalent) and on the discounting component. The effect of ambiguity aversion on discounting depends on the  $EIS$ . When  $EIS < 1$ , increasing ambiguity aversion leads to a smaller effective discount rate on climate damages. For the interesting case  $EIS > 1$ , the opposite is true and therefore increasing ambiguity aversion has two offsetting effects on the SCC. This is the reason that the influence of ambiguity aversion on the SCC is not very large when  $EIS > 1$ . An interesting extension of the model is to assume that the state-variables are stochastic as well. Traeger (2018) shows that introducing randomness in the carbon cycle and the temperature response to radiative forcing can have substantial effects on the SCC. In reality the exact flows of carbon between the reservoirs and the temperature response to atmospheric carbon concentrations are uncertain so this is a useful direction for future research.

## References

- Abel, A. B. (1999). Risk premia and term premia in general equilibrium. *Journal of Monetary Economics*, 43(1), 3–33.
- 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.
- 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.
- Berkman, H., Jacobsen, B., & Lee, J. B. (2011). Time-varying rare disaster risk and stock returns. *Journal of Financial Economics*, 101(2), 313–332.
- Branger, N., & Larsen, L. S. (2013). Robust portfolio choice with uncertainty about jump and diffusion risk. *Journal of Banking & Finance*, 37(12), 5036–5047.
- Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3), 265–296. Retrieved from <https://EconPapers.repec.org/RePEc:eee:jfinec:v:7:y:1979:i:3:p:265-296>
- Bremer, T., & van der Ploeg, F. (2018). The risk-adjusted carbon price. *OxCarre Research Paper*, 203.
- Bretscher, L., & Vinogradova, A. (2018). Escaping damocles’ sword: Endogenous climate shocks in a growing economy. *Economics Working Paper Series*, 18.
- Brock, W. A., & Hansen, L. P. (2017). Wrestling with uncertainty in climate economic models.

- Cai, Y., Judd, K. L., & Lontzek, T. S. (2015). The social cost of carbon with economic and climate risks. *arXiv preprint arXiv:1504.06909*.
- 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.
- Dimson, E., Marsh, P., & Staunton, M. (2011). Equity premia around the world. *Available at SSRN 1940165*.
- 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., Farhi, E., & Strzalecki, T. (2014). How much would you pay to resolve long-run risk? *American Economic Review*, 104(9), 2680–97.
- Epstein, L. G., & Schneider, M. (2003). Recursive multiple-priors. *Journal of Economic Theory*, 113(1), 1–31.
- 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](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>
- Hansen, L. P., & Sargent, T. J. (2001). Robust control and model uncertainty. *American Economic Review*, 91(2), 60–66.
- Hansen, L. P., & Sargent, T. J. (2008). *Robustness*. Princeton university press.
- Hansen, L. P., Sargent, T. J., Turmuhambetova, G., & Williams, N. (2001). Robustness and uncertainty aversion. *Manuscript, University of Chicago*.
- 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).
- 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.
- Jin, X., Luo, D., & Zeng, X. (2017). Dynamic asset allocation with uncertain jump risks: A pathwise optimization approach. *Mathematics of Operations Research*.
- 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*.
- Kreps, D. M., & Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica: journal of the Econometric Society*, 185–200.
- 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.
- Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of monetary Economics*, 15(2), 145–161.
- Munk, C. (2015). *Dynamic asset allocation*.
- Naik, V., & Lee, M. (1990). General equilibrium pricing of options on the market portfolio with discontinuous returns. *The Review of Financial Studies*, 3(4), 493–521.
- Noble, I., Bolin, B., Ravindranath, N., Verardo, D., & Dokken, D. (2000). *Land use, land use change, and forestry*. Cambridge University Press.
- 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. (2007). A review of the stern review on the economics of climate change. *Journal of economic literature*, 45(3), 686–702.
- 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.
- Pathak, P. (2002). Notes on robust portfolio choice. *unpublished paper, Harvard University*, 3, 14.
- Pindyck, R. S. (2017). The use and misuse of models for climate policy. *Review of Environmental Economics and Policy*, 11(1), 100–114.
- Pindyck, R. S., & Wang, N. (2013). The economic and policy consequences of catastrophes. *American Economic Journal: Economic Policy*, 5(4), 306–39.
- 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.

- Shreve, S. E. (2004). *Stochastic calculus for finance ii: Continuous-time models* (Vol. 11). Springer Science & Business Media.
- Stern, N. (2015). *Why are we waiting?: The logic, urgency, and promise of tackling climate change*. Mit Press.
- Stern, N., Peters, S., Bakhshi, V., Bowen, A., Cameron, C., Catovsky, S., ... others (2006). *Stern review: The economics of climate change* (Vol. 30). HM treasury London.
- Strzalecki, T. (2011). Axiomatic foundations of multiplier preferences. *Econometrica*, 79(1), 47–73.
- 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. (2014). Why uncertainty matters: discounting under intertemporal risk aversion and ambiguity. *Economic Theory*, 56(3), 627–664.
- Traeger, C. P. (2018). Ace-analytic climate economy (with temperature and uncertainty).
- Tsai, J., & Wachter, J. A. (2015). Disaster risk and its implications for asset pricing. *Annual Review of Financial Economics*, 7, 219–252.
- Tsai, J., & Wachter, J. A. (2018). *Pricing long-lived securities in dynamic endowment economies* (Tech. Rep.). National Bureau of Economic Research.
- 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.
- 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 Proof of Proposition 2

We will first derive the Hamilton-Jacobi-Bellman equation for every measure  $\mathbb{Q}^{a,b}$ . Duffie and Epstein (1992b) show that the HJB-equation for stochastic differential utility equals:

$$0 = f(C_t, V_t^{\mathbb{Q}}) + \mathcal{D}\mathcal{V}^{\mathbb{Q}} \quad (56)$$

Here  $\mathcal{D}\mathcal{V}^{\mathbb{Q}}$  is the drift of the value function. In order to calculate the drift of the value function, we will apply Itô's lemma. The derivative of  $V_t^{\mathbb{Q}}$  with respect to  $C$  is denoted by  $V_C^{\mathbb{Q}}$ . Similar notation is used for the second derivative. By Itô's lemma for jump processes we have:

$$dV_t^{\mathbb{Q}} = V_C^{\mathbb{Q}} \left( (\mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}}) C_t dt + \sigma C_t dZ_t^{\mathbb{Q}} \right) + \frac{1}{2} V_{CC}^{\mathbb{Q}} \sigma^2 C_t^2 dt + \left( V^{\mathbb{Q}}((1+Y_t)C_{t-}) - V^{\mathbb{Q}}(C_{t-}) \right) dN_t \quad (57)$$

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

$$\mathcal{DV}^{\mathbb{Q}} = V_C^{\mathbb{Q}} \left( \mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} \right) C_t + \frac{1}{2} V_{CC}^{\mathbb{Q}} \sigma^2 C_t^2 + \lambda_t^{\mathbb{Q}} E_t^{\mathbb{Q}} [V^{\mathbb{Q}}((1 + Y_t)C_{t-}) - V^{\mathbb{Q}}(C_{t-})] \quad (58)$$

This gives the following Hamilton-Jacobi-Bellman equation:

$$\begin{aligned} 0 = & f(C_t, V_t^{\mathbb{Q}}) + V_C^{\mathbb{Q}} \left( (\mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}}) C_t dt + \frac{1}{2} V_{CC}^{\mathbb{Q}} \sigma^2 C_t^2 \right. \\ & \left. + \lambda_t^{\mathbb{Q}} E_t^{\mathbb{Q}} [V^{\mathbb{Q}}((1 + Y_t)C_{t-}) - V^{\mathbb{Q}}(C_{t-})] \right) \end{aligned} \quad (59)$$

## B Proof of Proposition 5

First we calculate the derivatives that we need to solve this equation. The first and second derivative of  $V^{\mathbb{Q}}(C_t)$  with respect to  $C_t$  equal:

$$V_C^{\mathbb{Q}} = gC_t^{-\gamma}, \quad V_{CC}^{\mathbb{Q}} = -\gamma gC_t^{-\gamma-1} \quad (60)$$

Furthermore, we can substitute  $V^{\mathbb{Q}}(C_t) = \frac{gC_t^{1-\gamma}}{1-\gamma}$  into  $f(C, V)$ :

$$\begin{aligned} f(C_t, V^{\mathbb{Q}}(C_t)) &= \frac{\beta}{1-1/\epsilon} \frac{C_t^{1-1/\epsilon} - (gC_t^{1-\gamma})^{\frac{1}{\xi}}}{(gC_t^{1-\gamma})^{\frac{1}{\xi}-1}} = \frac{\beta}{1-1/\epsilon} \frac{C_t^{1-1/\epsilon} - g^{\frac{1}{\xi}} C_t^{1-1/\epsilon}}{g^{\frac{1}{\xi}-1} C_t^{\gamma-1/\epsilon}} \\ &= \frac{\beta}{1-1/\epsilon} \left( g^{1-\frac{1}{\xi}} C_t^{1-\gamma} - gC_t^{1-\gamma} \right) = \frac{\beta}{1-1/\epsilon} \left( g^{-\frac{1}{\xi}} - 1 \right) gC_t^{1-\gamma} \end{aligned} \quad (61)$$

Now we can substitute (60) together with (61) into the HJB-equation. Furthermore we use that  $E_t^{\mathbb{Q}} [V^{\mathbb{Q}}((1 + Y_t)C_{t-}) - V^{\mathbb{Q}}(C_{t-})] = \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} gC_t^{1-\gamma}$ . This gives us the following equation:

$$\begin{aligned} 0 = & \frac{\beta}{1-1/\epsilon} \left( g^{-\frac{1}{\xi}} - 1 \right) gC_t^{1-\gamma} + (\mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}}) gC_t^{1-\gamma} - \frac{\gamma}{2} \sigma^2 gC_t^{1-\gamma} \\ & + \lambda_t^{\mathbb{Q}} \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} gC_t^{1-\gamma} \end{aligned} \quad (62)$$

Then we divide (62) by  $gC_t^{1-\gamma}$  to obtain:

$$0 = \frac{\beta}{1-1/\epsilon} \left( g^{-\frac{1}{\xi}} - 1 \right) + (\mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}}) - \frac{\gamma}{2} \sigma^2 + \lambda_t^{\mathbb{Q}} \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \quad (63)$$

## C Proof of Proposition 7

### C.1 Pricing Kernel

We will start with deriving the explicit stochastic differential equation of the pricing kernel. We have that  $\pi_t = \xi_t^{a^*, b^*} \exp \left\{ \int_0^t f_V(C_s, V_s) ds \right\} f_C(C_t, V_t)$ . First we calculate

the derivatives of  $f(C_t, V_t)$  with respect to  $C_t$  and  $V_t$ .

$$f_C(C_t, V_t) = \frac{\beta C^{-1/\epsilon}}{((1-\gamma)V)^{\frac{1}{\zeta}-1}} \quad (64)$$

$$f_V(C_t, V_t) = \beta \zeta \left\{ \left(1 - \frac{1}{\zeta}\right) \left((1-\gamma)V_t\right)^{-\frac{1}{\zeta}} C^{1-1/\epsilon} - 1 \right\} \quad (65)$$

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

$$f_C(C_t, V_t) = \beta g^{1-\frac{1}{\zeta}} C_t^{-\gamma} \quad (66)$$

$$f_V(C_t, V_t) = \beta \zeta \left\{ g^{-\frac{1}{\zeta}} \left(1 - \frac{1}{\zeta}\right) - 1 \right\} \quad (67)$$

This gives:

$$\pi_t = \xi_t^{a^*, b^*} \exp \left( \int_0^t \beta \zeta \left( g^{-\frac{1}{\zeta}} \left(1 - \frac{1}{\zeta}\right) - 1 \right) ds \right) \beta g^{1-\frac{1}{\zeta}} C_t^{-\gamma} \quad (68)$$

We take the logarithm and rewrite (68) towards a stochastic differential equation.

$$d \log(\pi_t) = \beta \zeta \left( g^{-\frac{1}{\zeta}} \left(1 - \frac{1}{\zeta}\right) - 1 \right) dt - \gamma d \log(C_t) + d \log(\xi_t^{a^*, b^*}) \quad (69)$$

Calculating and substituting  $\log(C_t)$  and  $\log(\xi_t^{a^*, b^*})$  gives:

$$\begin{aligned} d \log(\pi_t) = & \left\{ \beta \zeta \left( g^{-\frac{1}{\zeta}} \left(1 - \frac{1}{\zeta}\right) - 1 \right) - \gamma \left( \mu - \lambda m - \frac{\sigma^2}{2} \right) + \lambda - \lambda^* \right. \\ & \left. - \frac{1}{2\sigma^2} (\lambda^* m^* - \lambda m)^2 \right\} dt + \left\{ -\gamma \sigma - \frac{\lambda^* m^* - \lambda m}{\sigma} \right\} dZ_t \\ & + \left( a^* + (b^* - \gamma) W_t - b^* \mu_J - \frac{1}{2} (b^*)^2 \sigma_J^2 \right) dN_t \end{aligned} \quad (70)$$

Applying Ito's lemma once more to  $\log(\pi_t)$  we obtain:

$$\begin{aligned} d\pi_t = & \left\{ \beta \zeta \left( g^{-\frac{1}{\zeta}} \left(1 - \frac{1}{\zeta}\right) - 1 \right) - \gamma \left( \mu - \lambda^* m^* - (\gamma + 1) \frac{\sigma^2}{2} \right) + \lambda - \lambda^* \right\} \pi_t dt \\ & + \left\{ -\gamma \sigma - \frac{\lambda^* m^* - \lambda m}{\sigma} \right\} \pi_t dZ_t + \left( e^{a^* + (b^* - \gamma) W_t - b^* \mu_J - \frac{1}{2} (b^*)^2 \sigma_J^2} - 1 \right) \pi_t dN_t \end{aligned} \quad (71)$$

Substituting  $g$  into (71) gives:

$$\begin{aligned} d\pi_t = & \left\{ -\beta - \frac{1}{\epsilon} \left( \mu - \lambda^* m^* \right) + \left( 1 + \frac{1}{\epsilon} \right) \frac{\gamma}{2} \sigma^2 + \left( \gamma - \frac{1}{\epsilon} \right) \left( \lambda^* \frac{e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \right) \right. \\ & \left. + \lambda - \lambda^* \right\} \pi_t dt + \left\{ -\gamma \sigma - \frac{\lambda^* m^* - \lambda m}{\sigma} \right\} \pi_t dZ_t \\ & + \left( e^{a^* + (b^* - \gamma) W_t - b^* \mu_J - \frac{1}{2} (b^*)^2 \sigma_J^2} - 1 \right) \pi_t dN_t \end{aligned} \quad (72)$$

## C.2 Interest rate

By the no-arbitrage argument,  $r$  should be such that  $\pi_t B_t$  is a martingale. Now write  $d\pi_t = \mu_\pi \pi_t dt + \sigma_\pi \pi_t dZ_t + J_\pi(W_t) \pi_{t-} dN_t$ .

$$d\pi_t B_t = (r + \mu_\pi) \pi_t B_t dt + \sigma_\pi \pi_t B_t dZ_t + J_\pi(W_t) \pi_{t-} B_t dN_t \quad (73)$$

This is a martingale if  $r + \mu_\pi + \lambda E_t[J_\pi(W_t)] = r + \mu_\pi + \lambda(e^{a^* - \gamma\mu_J - \gamma b^* \sigma_J^2 + \frac{1}{2}\gamma^2 \sigma_J^2} - 1) = 0$ . Therefore the interest rate equals:

$$\begin{aligned} r &= -\mu_\pi - \lambda(e^{a^* - \gamma\mu_J - \gamma b^* \sigma_J^2 + \frac{1}{2}\gamma^2 \sigma_J^2} - 1) \\ &= \beta + \frac{1}{\epsilon}(\mu - \lambda^* m^*) - \left(1 + \frac{1}{\epsilon}\right) \frac{\gamma}{2} \sigma^2 - \left(\gamma - \frac{1}{\epsilon}\right) \lambda^* \frac{e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} \\ &\quad - \lambda^*(e^{-\gamma(\mu_J + b^* \sigma_J^2 - \frac{1}{2}\gamma \sigma_J^2)} - 1) \end{aligned} \quad (74)$$

Substituting  $r$  into the pricing kernel gives:

$$\begin{aligned} d\pi_t &= \left\{ -r - \lambda(e^{a^* - \gamma(\mu_J + b^* \sigma_J^2 - \frac{1}{2}\gamma \sigma_J^2)} - 1) \right\} \pi_t dt + \left\{ -\gamma\sigma - \frac{\lambda^* m^* - \lambda m}{\sigma} \right\} \pi_t dZ_t \\ &\quad + \left( e^{a^* + (b^* - \gamma)W_t - b^* \mu_J - \frac{1}{2}(b^*)^2 \sigma_J^2} - 1 \right) \pi_{t-} dN_t \end{aligned} \quad (75)$$

## C.3 Risk premium and consumption wealth ratio

Using a similar argument as for the bond, by no-arbitrage the discounted dividend paying stock price should be a martingale. Using the Ito product rule for jump processes, we find:

$$\begin{aligned} d\pi_t S_t^d &= \left( \mu - \lambda m + k - r - \lambda(e^{a^* - \gamma(\mu_J + b^* \sigma_J^2 - \frac{1}{2}\gamma \sigma_J^2)} - 1) - \gamma\sigma^2 - (\lambda^* m^* - \lambda m) \right) \pi_t S_t^d dt \\ &\quad + \left( \sigma - \gamma\sigma - \frac{\lambda^* m^* - \lambda m}{\sigma} \right) \pi_t S_t^d dZ_t + \left( e^{a^* + (b^* - \gamma + 1)W_t - b^* \mu_J - \frac{1}{2}(b^*)^2 \sigma_J^2} - 1 \right) \pi_{t-} S_{t-}^d dN_t \end{aligned} \quad (76)$$

This is a martingale when:

$$\mu - \lambda^* m^* + k - r - \lambda(e^{a^* - \gamma(\mu_J + b^* \sigma_J^2 - \frac{1}{2}\gamma \sigma_J^2)} - 1) - \gamma\sigma^2 + \lambda(e^{a^* + (1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1) = 0 \quad (77)$$

From this equation, we obtain that the consumption wealth ratio  $k$  equals:

$$k = \beta g^{-\frac{1}{\epsilon}} \quad (78)$$

Lastly, we can calculate the risk premium for the stock. Define the expected return of the dividend paying stock as:  $r^S = \mu + k$ . Then the risk premium is the excess return of the stock over a risk-free bond:

$$r^S - r = \mu + k - r = \gamma\sigma^2 + \lambda^* m^* - \lambda^* e^{(1-\gamma)(\mu_J + b^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} + \lambda^* e^{-\gamma(\mu_J + b^* \sigma_J^2 - \frac{1}{2}\gamma \sigma_J^2)} \quad (79)$$

## D Solution method for the problem with state variables

In this appendix we propose a solution method to find the value function of the agent in the climate change economy. The idea of the method is the following. We first derive the HJB equation. Then we calculate the pricing kernel and substitute the HJB equation into this pricing kernel. This gives us an explicit expression for the pricing kernel in terms of parameters, we can use this to price consumption strips. The integral over all future consumption strips equals the stock price and is therefore equal to total wealth. Given the stock price and the value of a consumption strip, we can calculate the wealth consumption ratio. From the wealth consumption ratio, we can calculate  $g(X_t)$  and the value function. This in turn enables us to obtain an expression for the social cost of carbon.

### D.1 HJB equation

Similar to proposition [2](#), we can obtain the HJB equation for the problem with a climate model. This time, several state variables will enter the HJB equation.  $X_t$  is a column vector of  $n$  state variables. Furthermore, there is no uncertainty in the state variables. The state variables then follow:  $dX_t = \mu_X(X_t)dt$ . The arrival rate is state dependent:  $\lambda_t = \lambda(X_t)$ ,  $\lambda_t^{\mathbb{Q}} = e^{a_t}\lambda_t$ . The growth rate and the volatility of consumption are still constant. Denote by  $V_X^{\mathbb{Q}}$  the row vector of partial derivatives of the value function  $V_t^{\mathbb{Q}}$  with respect to the vector of state variables  $X_t$ ,  $V_X^{\mathbb{Q}} = \left[ \frac{\partial V^{\mathbb{Q}}(C_t, X_t)}{\partial X_{1t}} \quad \dots \quad \frac{\partial V^{\mathbb{Q}}(C_t, X_t)}{\partial X_{nt}} \right]$ . Under the measure  $\mathbb{Q}^{a,b}$ , we obtain the following HJB equation:

$$0 = f(C_t, V_t^{\mathbb{Q}}) + V_C(\mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}})C_t dt + \frac{1}{2} V_{CC}^{\mathbb{Q}} \sigma^2 C_t^2 + V_X \mu_X(X_t) + \lambda_t^{\mathbb{Q}} E_t^{\mathbb{Q}} [V^{\mathbb{Q}}((1 + Y_t)C_{t-}, X_t) - V^{\mathbb{Q}}(C_{t-}, X_t)] \quad (80)$$

Conjecture again that  $V^{\mathbb{Q}}(C_t, X_t) = \frac{g(X_t)C_t^{1-\gamma}}{1-\gamma}$ , the only difference with the constant  $\lambda$  case is that the function  $g$  is now state dependent. Substituting our conjecture  $V^{\mathbb{Q}}(C_t, X_t) = \frac{g(X_t)C_t^{1-\gamma}}{1-\gamma}$  into  $f(C_t, V_t)$  gives:

$$\begin{aligned} f(C_t, V^{\mathbb{Q}}(C_t, X_t)) &= \frac{\beta}{1 - 1/\epsilon} \frac{C_t^{1-1/\epsilon} - \left(g(X_t)C_t^{1-\gamma}\right)^{\frac{1}{\zeta}}}{\left(g(X_t)C_t^{1-\gamma}\right)^{\frac{1}{\zeta}-1}} \\ &= \frac{\beta}{1 - 1/\epsilon} \left(g(X_t)^{\frac{1}{\zeta}-1} C_t^{1-\gamma} - g(X_t)C_t^{1-\gamma}\right) \\ &= \beta \zeta \left(g(X_t)^{-\frac{1}{\zeta}} - 1\right) V^{\mathbb{Q}}(C_t) \end{aligned} \quad (81)$$



The partial derivatives of  $V$  are given by:

$$\begin{aligned} V_C^{\mathbb{Q}} &= g(X_t)C_t^{-\gamma}, \quad V_{CC}^{\mathbb{Q}} = -\gamma g(X_t)C_t^{-\gamma-1} \\ V_X^{\mathbb{Q}} &= \frac{g_X(X_t)C_t^{1-\gamma}}{1-\gamma} \end{aligned} \quad (82)$$

Here  $g_X$  denotes the row vector with partial derivatives to each of the state variables, similar to  $V_X^{\mathbb{Q}}$ . Substituting  $f(C_t, V^{\mathbb{Q}}(C_t, X_t))$  together with the partial derivatives of  $V_t^{\mathbb{Q}}$  into (80) yields the following equation.

$$\begin{aligned} 0 &= \frac{\beta}{1-1/\epsilon} \left( g(X_t)^{-\frac{1}{\zeta}} - 1 \right) g(X_t)C_t^{1-\gamma} + \left( \mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} \right) g(X_t)C_t^{1-\gamma} \\ &\quad - \frac{\gamma}{2} \sigma^2 g(X_t)C_t^{1-\gamma} + \frac{g_X(X_t)C_t^{1-\gamma}}{1-\gamma} \mu_X(X_t) \\ &\quad + \lambda_t^{\mathbb{Q}} \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} g(X_t)C_t^{1-\gamma} \end{aligned} \quad (83)$$

Dividing by  $g(X_t)C_t^{1-\gamma}$  gives:

$$\begin{aligned} 0 &= \frac{\beta}{1-1/\epsilon} \left( g(X_t)^{-\frac{1}{\zeta}} - 1 \right) + \left( \mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} \right) - \frac{\gamma}{2} \sigma^2 + \frac{g_X(X_t)}{g(X_t)(1-\gamma)} \mu_X(X_t) \\ &\quad + \lambda_t^{\mathbb{Q}} \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \end{aligned} \quad (84)$$

This is the HJB-equation given the measure  $\mathbb{Q}^{a,b}$ . If we look at the problem with ambiguity aversion, we consider multiple measures. The HJB-equation for that problem becomes:

$$\begin{aligned} 0 &= \min_{(a_t, b_t) \in \Theta_t} \left\{ \frac{\beta}{1-1/\epsilon} \left( g(X_t)^{-\frac{1}{\zeta}} - 1 \right) + \left( \mu - \lambda_t^{\mathbb{Q}} m_t^{\mathbb{Q}} \right) - \frac{\gamma}{2} \sigma^2 + \frac{g_X(X_t)}{g(X_t)(1-\gamma)} \mu_X(X_t) \right. \\ &\quad \left. + \lambda_t^{\mathbb{Q}} \frac{e^{(1-\gamma)(\mu_J + b_t \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \right\} \end{aligned} \quad (85)$$

We can again use proposition 6 to determine the optimal  $a_t$  and  $b_t$ . However, note that  $\lambda_t$  is not constant anymore. Therefore,  $a_t^*$  and  $b_t^*$  are also time dependent. Define  $\lambda_t^* = e^{a_t^*} \lambda_t$  and  $m_t^* = e^{\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2} - 1$ . Given the optimal control variables  $a_t^*$  and  $b_t^*$  and after rearranging we obtain the following equation:

$$\begin{aligned} \frac{g_X(X_t)}{g(X_t)(1-\gamma)} \mu_X(X_t) &= \frac{\beta}{1-1/\epsilon} \left( 1 - g(X_t)^{-\frac{1}{\zeta}} \right) - \left( \mu - \lambda_t^* m_t^* \right) + \frac{\gamma}{2} \sigma^2 \\ &\quad - \lambda_t^* \frac{e^{(1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \end{aligned} \quad (86)$$

## D.2 Pricing kernel

Similar to Appendix C we now derive the pricing kernel for this problem. (68) is still valid if we replace the constant function  $g$  by the state dependent function  $g(X_t)$ .

$$\pi_t = \xi_t^{a^*, b^*} \exp \left( \int_0^t \beta \zeta \left( g(X_t)^{-\frac{1}{\zeta}} \left( 1 - \frac{1}{\zeta} \right) - 1 \right) ds \right) \beta g(X_t)^{1-\frac{1}{\zeta}} C_t^{-\gamma} \quad (87)$$

Take the logarithm and write as a differential equation:

$$\begin{aligned} d \log(\pi_t) &= \beta \zeta \left( g(X_t)^{-\frac{1}{\zeta}} \left( 1 - \frac{1}{\zeta} \right) - 1 \right) dt - \gamma d \log(C_t) + d \log(\xi_t^{a^*, b^*}) \\ &\quad + \left( 1 - \frac{1}{\zeta} \right) d \log(g(X_t)) \end{aligned} \quad (88)$$

Apply Ito's lemma to  $\log(C_t)$ ,  $\log(\xi_t^{a^*, b^*})$  and  $\log(g(X_t))$  and substitute the result gives the following differential equation.

$$\begin{aligned} d \log(\pi_t) &= \left\{ \beta \zeta \left( g(X_t)^{-\frac{1}{\zeta}} \left( 1 - \frac{1}{\zeta} \right) - 1 \right) - \gamma \left( \mu - \lambda_t m_t - \frac{\sigma^2}{2} \right) + \lambda_t - \lambda_t^* \right. \\ &\quad \left. - \frac{1}{2\sigma^2} (\lambda_t^* m_t^* - \lambda m)^2 + (1/\epsilon - \gamma) \frac{g_X(X_t)}{g(X_t)(1-\gamma)} \mu_X(X_t) \right\} dt \\ &\quad + \left\{ -\gamma\sigma - \frac{\lambda_t^* m_t^* - \lambda m}{\sigma} \right\} dZ_t + \left( a_t^* + (b_t^* - \gamma)W_t - b^* \mu_J - \frac{1}{2}(b_t^*)^2 \sigma_J^2 \right) dN_t \end{aligned} \quad (89)$$

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

$$\begin{aligned} d\pi_t &= \left\{ \beta \zeta \left( g(X_t)^{-\frac{1}{\zeta}} \left( 1 - \frac{1}{\zeta} \right) - 1 \right) - \gamma \left( \mu - \lambda_t^* m_t^* - (\gamma + 1) \frac{\sigma^2}{2} \right) + \lambda_t - \lambda_t^* \right. \\ &\quad \left. + (1/\epsilon - \gamma) \frac{g_X(X_t)}{g(X_t)(1-\gamma)} \mu_X(X_t) \right\} \pi_t dt + \left\{ -\gamma\sigma - \frac{\lambda_t^* m_t^* - \lambda_t m_t}{\sigma} \right\} \pi_t dZ_t \\ &\quad + \left( e^{a_t^* + (b_t^* - \gamma)W_t - b_t^* \mu_J - \frac{1}{2}(b_t^*)^2 \sigma_J^2} - 1 \right) \pi_t dN_t \end{aligned} \quad (90)$$

We can now substitute the HJB equation (86) into the pricing kernel. Several terms cancel out and we are left with:

$$\begin{aligned} d\pi_t &= \left\{ -\beta - \frac{1}{\epsilon} \left( \mu - \lambda_t^* m_t^* \right) + \left( 1 + \frac{1}{\epsilon} \right) \frac{\gamma}{2} \sigma^2 + \left( \gamma - \frac{1}{\epsilon} \right) \left( \lambda_t^* \frac{e^{(1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \right. \right. \\ &\quad \left. \left. + \lambda_t - \lambda_t^* \right\} \pi_t dt + \left\{ -\gamma\sigma - \frac{\lambda_t^* m_t^* - \lambda_t m_t}{\sigma} \right\} \pi_t dZ_t \\ &\quad + \left( e^{a_t^* + (b_t^* - \gamma)W_t - b_t^* \mu_J - \frac{1}{2}(b_t^*)^2 \sigma_J^2} - 1 \right) \pi_t dN_t \end{aligned} \quad (91)$$

Similar to subsection C.2 we have that the interest rate  $r_t = -\mu_{\pi_t} - \lambda_t (e^{a_t^* - \gamma \mu_J - \gamma b_t^* \sigma_J^2 + \frac{1}{2} \gamma^2 \sigma_J^2} - 1)$ . This gives the following time-dependent interest rate.

$$\begin{aligned} r_t &= \beta + \frac{1}{\epsilon} \left( \mu - \lambda_t^* m_t^* \right) - \left( 1 + \frac{1}{\epsilon} \right) \frac{\gamma}{2} \sigma^2 - \left( \gamma - \frac{1}{\epsilon} \right) \lambda_t^* \frac{e^{(1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} \\ &\quad - \lambda_t^* (e^{-\gamma(\mu_J + b_t^* \sigma_J^2 - \frac{1}{2} \gamma \sigma_J^2)} - 1) \end{aligned} \quad (92)$$

### D.3 Consumption strips

Let  $H_t = H(C_t, X_t, s-t) = E_t \left[ \frac{\pi_s}{\pi_t} C_s \right]$  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  $H(C_t, X_t, u) = C_t \exp \left\{ -ku - (1/\epsilon - 1) \int_t^{t+u} \left( \frac{e^{(1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_t^* \right) \lambda^*(X_s) ds \right\}$ .  $u$  denotes the time to maturity of the consumption strip and  $k = \beta + (1/\epsilon - 1)(\mu - \frac{\gamma}{2}\sigma^2)$  is the consumption-wealth ratio in the absence of climate risk. Remember that  $a_t^*$  and  $b_t^*$  are implicit functions of  $X_t$  since they depend on time only through  $\lambda(X_t)$ . Therefore we can write:  $H(C_t, X_t, u) = C_t \exp \left\{ -ku - (1/\epsilon - 1) \int_t^{t+u} f(X_s) ds \right\}$  with  $f(X_t) = \left( \frac{e^{(1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_t^* \right) \lambda^*(X_t)$ . Clearly,  $H(C_t, X_t, 0) = C_t$ . If we can verify that  $H_t$  multiplied by the pricing kernel is indeed a martingale, by the no-arbitrage condition  $H_t$  must be the price of the asset. Applying Ito's lemma to  $H_t$  gives:

$$\begin{aligned} dH_t &= H_C dC_t + H_X dX_t - \frac{\partial H_t}{\partial u} dt = \frac{1}{C_t} H_t dC_t \\ &\quad - (1/\epsilon - 1) \frac{\partial}{\partial X_t} \left( \int_t^{t+u} f(X_s) ds \right) \mu_X(X_t) H_t dt \\ &\quad + k + (1/\epsilon - 1) \frac{\partial}{\partial u} \left( \int_t^{t+u} f(X_s) ds \right) H_t dt \end{aligned} \quad (93)$$

We can calculate both derivatives:

$$\begin{aligned} \frac{\partial}{\partial X_t} \left( \int_t^{t+u} f(X_s) ds \right) \mu_X(X_t) &= \frac{\partial}{\partial t} \left( \int_t^{t+u} f(X_s) ds \right) \frac{\partial t}{\partial X_t} \mu_X(X_t) \\ &= \frac{\partial}{\partial t} \left( \int_t^{t+u} f(X_s) ds \right) = f(X_{t+u}) - f(X_t) \end{aligned} \quad (94)$$

$$\frac{\partial}{\partial u} \left( \int_t^{t+u} f(X_s) ds \right) = f(X_{t+u}) \quad (95)$$

Therefore  $dH_t$  becomes:

$$\begin{aligned} dH_t &= \left\{ \mu - \lambda_t m_t + k + (1/\epsilon - 1) \lambda^*(X_t) \left( \frac{e^{(1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1-\gamma} - m_t^* \right) \right\} H_t dt \\ &\quad + \sigma H_t dZ_t + Y_t H_t dN_t \end{aligned} \quad (96)$$

Now define  $d\pi_t = \mu_\pi \pi_t dt + \sigma_\pi \pi_t dZ_t + J_\pi(W_t) \pi_t dN_t$  and  $dH_t = \mu_H H_t dt + \sigma H_t dZ_t + Y_t H_t dN_t$ . Now we can verify that  $\pi_t H_t$  is martingale.

$$\begin{aligned} d\pi_t H_t &= (\mu_\pi + \mu_H + \sigma \sigma_\pi) \pi_t H_t dt + (\sigma + \sigma_\pi) \pi_t H_t dZ_t \\ &\quad + \left( Y_t + J_\pi(W_t) + Y_t J_\pi(W_t) \right) \pi_t H_t dN_t \end{aligned} \quad (97)$$

We can calculate the expectation of the jump term.

$$\begin{aligned} E_t[Y_t + J_\pi(W_t) + Y_t J_\pi(W_t)] &= E_t[e^{a_t^* + (b_t^* - \gamma + 1)W_t - b_t^* \mu_J - \frac{1}{2}(b_t^*)^2 \sigma_J^2} - 1] \\ &= e^{a_t^* + (1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1 \end{aligned} \quad (98)$$

Therefore  $\pi_t H_t$  is a martingale if:

$$0 = \mu_\pi + \mu_H + \sigma\sigma_\pi + \lambda \left( e^{a_t^* + (1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1 \right) \quad (99)$$

Calculation yields:

$$\mu_\pi + \mu_H = \gamma\sigma^2 + \lambda_t^* m_t^* - \lambda_t m_t - \lambda_t \left( e^{a_t^* + (1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1 \right) \quad (100)$$

Since  $\sigma\sigma_\pi = -\gamma\sigma^2 - \lambda_t^* m_t^* + \lambda_t m_t$ ,  $\pi_t H_t$  is a martingale.

## D.4 Equity premium

To calculate the equity premium, note that  $S_t = \frac{C_t}{k(X_t)}$ .

$$\begin{aligned} dS_t^d &= dS_t + C_t dt = \frac{1}{k(X_t)} dC_t - \frac{C_t}{k(X_t)^2} dk(X_t) + k(X_t) S_t dt \\ &= \left( \mu - \frac{k_X(X_t)}{k(X_t)} \mu_X(X_t) + k(X_t) - \lambda m \right) S_t dt + \sigma S_t dZ_t + Y_t S_t - dN_t \end{aligned} \quad (101)$$

The expected return on the stock equals  $r_t^S = \mu - \frac{k_X(X_t)}{k(X_t)} \mu_X(X_t) + k(X_t)$ . By subtracting the interest rate we then find the equity premium. Using the HJB-equation and equation (92) we find that:

$$\begin{aligned} r_t^S - r_t &= \mu - \frac{k_X(X_t)}{k(X_t)} \mu_X(X_t) + k(X_t) - r_t = \gamma\sigma^2 + \lambda_t^* m_t^* - \lambda_t e^{(1-\gamma)(\mu_J + b_t^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} \\ &\quad + \lambda_t^* e^{-\gamma(\mu_J + b_t^* \sigma_J^2 - \frac{1}{2}\gamma\sigma_J^2)} \end{aligned} \quad (102)$$

## E Proof of propositions 9 and 10

### E.1 Proposition 9

We first make an assumption on some parameter values. This assumption bounds the climate risk, to make sure that a higher elasticity of substitution leads to higher future prices. The assumption is:

$$\left( \mu - \frac{\gamma}{2}\sigma^2 \right) u + \int_t^{t+u} \lambda_s^* \left( \frac{e^{(1-\gamma)(\mu_J + b_s^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} - m_s^* \right) ds > 0 \quad \forall u \quad (103)$$

This assumption is not very restrictive and is satisfied for reasonable parameter values. Given this assumption,  $H(C_t, X_t, u)$  is increasing in  $\epsilon$ . Then we can take the derivative of  $H(C_t, X_t, u)$  with respect to the maturity  $u$ .

$$\begin{aligned} \frac{\partial}{\partial u} H(C_t, X_t, u) &= \left\{ -\beta - (1/\epsilon - 1) \left( \mu - \frac{\gamma}{\sigma^2} \right) \right. \\ &\quad \left. - (1/\epsilon - 1) \lambda_{t+u}^* \left( \frac{e^{(1-\gamma)(\mu_J + b_{t+u}^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} - m_{t+u}^* \right) \right\} H(C_t, X_t, u) \end{aligned} \quad (104)$$

Given the assumption we made in this proposition and that  $\frac{\partial H(C_t, X_t, u)}{\partial \epsilon} > 0$  it is clear that  $\frac{\partial H(C_t, X_t, u)}{\partial u}$  is increasing in  $\epsilon$ .

For this proof we will use the shorthand notation  $H(u; \epsilon)$ . Now consider two cases:  $H(u; \epsilon_1)$  and  $H(u; \epsilon_2)$  where the only difference between the two is that in the latter case the EIS is larger ( $\epsilon_2 > \epsilon_1$ ). Since the derivative of  $H(u; \epsilon)$  with respect to  $\epsilon$  is strictly positive and  $H(0; \epsilon) = C_t$  for all  $\epsilon$  we have that  $H(u; \epsilon_2) > H(u; \epsilon_1)$  if  $u > 0$ . The weight per maturity in the SCC equals  $\frac{H(u; \epsilon_1)}{\int_0^\infty H(u; \epsilon_1) du}$  in the first case and  $\frac{H(u; \epsilon_2)}{\int_0^\infty H(u; \epsilon_2) du}$ . Clearly,  $\int_0^\infty H(u; \epsilon_2) du > \int_0^\infty H(u; \epsilon_1) du$ . Therefore:

$$\frac{H(0; \epsilon_2)}{\int_0^\infty H(u; \epsilon_2) du} = \frac{C_t}{\int_0^\infty H(u; \epsilon_2) du} < \frac{H(0; \epsilon_1)}{\int_0^\infty H(u; \epsilon_1) du} \quad (105)$$

There must be a maturity  $u$  for which the opposite is true. This can be proven by contradiction. Assume that this is not the case. Then  $\frac{H(u; \epsilon_2)}{\int_0^\infty H(u; \epsilon_2) du} < \frac{H(u; \epsilon_1)}{\int_0^\infty H(u; \epsilon_1) du}$  for every  $u$ . But integrating over all possible maturities then yields:

$$1 = \int_0^\infty \frac{H(C_t, X_t, u; \epsilon_2)}{\int_0^\infty H(C_t, X_t, u; \epsilon_2) du} du < \int_0^\infty \frac{H(C_t, X_t, u; \epsilon_1)}{\int_0^\infty H(C_t, X_t, u; \epsilon_1) du} du = 1 \quad (106)$$

By contradiction, there is a maturity  $u$  for which  $\frac{H(u; \epsilon_2)}{\int_0^\infty H(u; \epsilon_2) du} > \frac{H(u; \epsilon_1)}{\int_0^\infty H(u; \epsilon_1) du}$ . Since  $H(u; \epsilon)$  is continuous, this implies that there exists a  $u^*$  for which the weights are the same in both cases.

Since  $\frac{\partial H(u; \epsilon)}{\partial u}$  is increasing in  $\epsilon$ , we also know that  $\frac{\partial H(u; \epsilon_2)}{\partial u} > \frac{\partial H(u; \epsilon_1)}{\partial u}$ . Therefore,  $u^*$  is the only maturity at which the two weighting functions intersect. For  $u < u^*$ ,  $\frac{H(u; \epsilon_2)}{\int_0^\infty H(u; \epsilon_2) du} < \frac{H(u; \epsilon_1)}{\int_0^\infty H(u; \epsilon_1) du}$  and for  $u > u^*$  we have the opposite inequality.

## E.2 Proposition 10

Note that  $b_t^*$  is decreasing in  $\theta$  and  $a_t^*$  is increasing in  $\theta$ . Assume  $\epsilon < 1$ , then  $\frac{\partial H(C_t, X_t, u)}{\partial \theta} > 0$ . We discussed this feature before, namely that additional risk (or ambiguity in this case) increases prices when  $\epsilon < 1$ . Now take a look at the derivative if  $H(C_t, X_t, u)$  with respect to  $u$ :

$$\begin{aligned} \frac{\partial}{\partial u} H(C_t, X_t, u) = & \left\{ -\beta - (1/\epsilon - 1)(\mu - \frac{\gamma}{\sigma^2}) \right. \\ & \left. - (1/\epsilon - 1)\lambda_{t+u}^* \left( \frac{e^{(1-\gamma)(\mu_J + b_{t+u}^* \sigma_J^2 + \frac{1}{2}(1-\gamma)\sigma_J^2)} - 1}{1 - \gamma} - m_{t+u}^* \right) \right\} H(C_t, X_t, u) \end{aligned} \quad (107)$$

This derivative is increasing in  $\theta$ . The proof of the proposition then follows using similar reasoning as in proposition 9.

When  $\epsilon > 1$ , we obtain the opposite case. Additional ambiguity decreases prices and  $\frac{\partial H(C_t, X_t, u)}{\partial \theta} < 0$ . Besides,  $\frac{\partial H(C_t, X_t, u)}{\partial u}$  is now decreasing in  $\theta$ . Using similar reasoning as proposition 9 again yields the proposition for  $\epsilon > 1$ .

## F Calibration

Table 3: Parameters for the Climate model

Par.	Description	Value
$E_t$	Initial level of total emissions (in $GtC$ , 2015)	10.45*
$g_t^E$	Initial growth rate of emissions (per year)	0.017
$g_\infty^E$	Long-run growth rate of emissions (per year)	-0.02
$\delta_{g^E}$	Speed of convergence of growth rate of emissions (per year)	0.0075
$M_t^{at}$	Initial atmospheric carbon concentration (in $GtC$ , 2015)	851*
$M_t^{up}$	Initial carbon concentration upper layer (in $GtC$ , 2015)	460*
$M_t^{at}$	Initial carbon concentration lower layer (in $GtC$ , 2015)	1740*
$M_{pre}^{at}$	Pre-industrial atmospheric carbon concentration (in $GtC$ )	588*
$\delta_{at \rightarrow up}$	Carbon flow rate from atmosphere to upper layer (per year)	0.0252*
$\delta_{up \rightarrow at}$	Carbon flow rate from upper layer to atmosphere (per year)	0.0412*
$\delta_{up \rightarrow lo}$	Carbon flow rate from upper layer to lower layer (per year)	0.0014*
$\delta_{lo \rightarrow up}$	Carbon flow rate from lower layer to upper layer (per year)	0.0003*
$\kappa_1$	Equilibrium forcing of $CO_2$ doubling (in $W/m^2$ )	3.68*
$EF_t$	Initial level of exogenous forcing (in $W/m^2$ )	0.5*
$EF_\infty$	Long-run level of exogenous forcing (in $W/m^2$ )	1*
$\kappa_2$	Speed of convergence exogenous forcing (per year)	0.02
$T_t$	Initial surface temperature compared to pre-industrial (in $^\circ C$ , 2015)	0.85*
$T_t^{oc}$	Initial ocean temperature compared to pre-industrial (in $^\circ C$ , 2015)	0.0068*
$\tau_1$	Thermal inertia parameter of the upper layer (in $Wyears/m^2/^\circ C$ )	47.71*
$\tau_2$	Equilibrium temperature impact of $CO_2$ doubling (in $^\circ C$ )	3.1*
$\tau_3$	Temp. transfer coefficient between upper and deep ocean (in years)	197.99*
$\tau_4$	Thermal inertia parameter of the deep ocean (in $Wyears/m^2/^\circ C$ )	17.42*

\*Based on the DICE 2016 calibration: <https://sites.google.com/site/williamdnordhaus/dice-rice>.

All parameters that depend on the time step of the model are transformed to yearly.