Environ Resource Econ (2013) 55:21–46
DOI 10.1007/s10640-012-9612-0

Scientific Ambiguity and Climate Policy

Antony Millner · Simon Dietz · Geoffrey Heal

Accepted: 4 October 2012 / Published online: 20 October 2012

Abstract Economic evaluation of climate policy traditionally treats uncertainty by appealing to expected utility theory. Yet our knowledge of the impacts of climate policy may not be of sufficient quality to be described by unique probabilistic beliefs. In such circumstances, it has been argued that the axioms of expected utility theory may not be the correct standard of rationality. By contrast, several axiomatic frameworks have recently been proposed that account for ambiguous knowledge. In this paper, we apply static and dynamic versions of a smooth ambiguity model to climate mitigation policy. We obtain a general result on the comparative statics of optimal abatement and ambiguity aversion, and then extend our analysis to a more realistic, dynamic setting, where we introduce scientific ambiguity into the well-known DICE model of the climate-economy system. For policy-relevant exogenous mitigation policies, we show that the value of emissions abatement increases as ambiguity aversion increases, and that this 'ambiguity premium' can in some plausible cases be very

An earlier version of this paper appeared as NBER Working Paper no. 16050. We thank Martin Weitzman, Christian Traeger, Larry Karp, Cameron Hepburn, seminar participants at Columbia, Harvard, Berkeley, Stanford, AERE, EAERE, and the 12th Occasional California Workshop on Environmental Economics, and two anonymous referees for helpful comments. We are grateful to Malte Meinshausen for supplying us with the empirical estimates of the climate sensitivity distributions.

A. Millner (✉)
Department of Agricultural and Resource Economics, University of California, Berkeley,
Berkeley, CA, USA
e-mail: A.Millner@lse.ac.uk

A. Millner · S. Dietz
Grantham Research Institute on Climate Change and the Environment,
London School of Economics and Political Science, London, UK
e-mail: s.dietz@lse.ac.uk

S. Dietz
Department of Geography and Environment, London School of Economics and Political Science,
London, UK

G. Heal
Columbia Business School, Columbia University, New York, NY, USA

large. In these cases the effect of ambiguity aversion on welfare is comparable to that of other much studied welfare parameters. Thus ambiguity aversion may be an important neglected aspect of climate change economics, and seems likely to provide another argument for strong abatement policy.

Keywords Climate change · Uncertainty · Ambiguity

JEL Classification Q54 · D81

1 Introduction

The literature on optimal climate change mitigation policy has thus far remained faithful to the long tradition of welfare analysis based on expected utility theory. The integrated assessment models that are widely used for policy evaluation all have a common welfare-analytic core. Most studies employ deterministic models (e.g. Manne and Richels 1992; Nordhaus 2008; Tol 1997), allowing for efficient determination of optimal policies, which are then subjected to sensitivity analysis in order to test their robustness to changes in model parameters. Other studies employ monte carlo style simulation models (e.g Hope 2006), and generally do not find optimal policies, but rather provide welfare assessments of exogenously specified greenhouse gas emissions pathways. A few authors have combined these two approaches by solving stochastic-dynamic control problems to determine optimal policies that account for learning about future risks (e.g. Keller et al. 2004; Kelly and Kolstad 1999; Karp and Zhang 2006). Thus, while there are several models of varying complexity and emphasis, they share a common commitment to the expected utility framework.

The reasons for the primacy of expected utility theory as a normative model of rational choice are well known to economists. Its axiomatic foundations have been developed by several authors (von Neumann and Morgenstern 1944; Savage 1954; Anscombe and Aumann 1963; Herstein and Milnor 1953). Savage's presentation is widely considered the most satisfactory, since it derives both utility functions and subjective probabilities from primitive preferences over very general mathematical objects known as 'acts'—maps between states and outcomes. Indeed the Savage axioms are often considered to be synonymous with rational choice. Nevertheless, Savage himself took a cautious approach to his theory, suggesting that it should only be applied in small worlds, in which it is possible to 'look before you leap', i.e. imagine every possible contingency, and identify a complete ordering of acts over these contingencies (see Binmore 2009 for a discussion). Further potential limitations on the domain of applicability of Savage's theory were famously identified by Ellsberg (1961), who showed that when our state of knowledge is more accurately described as uncertainty (i.e. unknown probabilities) rather than risk (i.e. known probabilities—we use these terms in the sense of Knight 1921), we may wish to violate Savage's second axiom, the 'sure-thing principle'.

A strong Bayesian would interpret Ellsberg's results as a contribution to positive, rather than normative, decision theory. Bayesians believe that Savage's axioms define rational choice, in which case the preferences Ellsberg observes in his investigations are deemed irrational, and all uncertainty is always describable by a unique subjective probability distribution. This viewpoint has however been strongly contested. As noted by Ellsberg (1961) and Slovic and Tversky (1974), people often stick to choices that violate the sure-thing principle in Ellsberg's choice experiments, even when this violation is pointed out to them. This is in stark contrast to other decision theoretic 'paradoxes', such as the Allais paradox, where people often revert to the prescriptions of expected utility once their violation of the axioms is explained. It has been argued, we think convincingly, that when our information about the world is incomplete, inconsistent, or nonexistent, Savage's axioms need not be the correct standard of rationality (Gilboa et al. 2008, 2009), and it does not necessarily make sense to describe our state of knowledge with a unique probability distribution over states of the world.[1]

Given these views, our assessment of the validity of the expected utility approach to the welfare analysis of climate change policy must depend on how structured our beliefs about the climate system[2] are. If we have sufficiently high quality information to justify probabilistic beliefs, then the approach adopted thus far in the literature is unequivocally useful. If not, we need to justify why this approach is a useful approximation, or attempt to define welfare measures that are true to our actual state of knowledge about the climate system, and reflect our preferences over bets with unknown probabilities. Very often a good way of justifying an approximation is to embed it in a more general framework, and show that the increased power of this framework does not materially alter the results achieved by the approximation. Thus, provided we suspect that our knowledge of the climate system is not very high quality, there would seem to be good reason to develop approaches to policy evaluation which account for uncertainty and not just risk, since these will either justify our reliance on existing methods, or provide appropriate tools for future work.

What is our state of knowledge about the climate system, and can it be described by unique probabilities? We feel it is important to break the state of scientific knowledge about climate into two categories: broad scientific principles, and detailed empirical predictions. In the first category belong concepts such as the laws of thermodynamics, fluid dynamics, and statements of fact such as 'CO_2 traps outgoing long-wave radiation, causing warming'. We believe these principles to be unimpeachable. In the second category belong the sophisticated models scientists use to convert these principles into predictions—energy balance models (EBMs), earth systems models of intermediate complexity (EMICs), and full-scale general circulation models (GCMs). These models can be enormously complex, and attempt to predict, among other things, the response of the global climate to increases in the concentrations of greenhouse gases. Because of the complexity of their task, and the intrinsic difficulties of prediction in highly nonlinear multi-dimensional physical systems (see Smith 2002, 2007; Stainforth et al. 2007; Frame et al. 2007 for illuminating discussions of the scientific and philosophical challenges of climate prediction), these models are not always in agreement with one another. As an example of this, consider Fig. 1, which plots the results of several recent studies' attempts to use different models and observational data to estimate climate sensitivity—the amount by which global mean surface temperature rises for a doubling of CO_2 concentration, in equilibrium. Climate sensitivity is a crucial summary statistic that coarsely captures how the global climate responds to increases in CO_2 concentrations, and features prominently in integrated assessment models.

From the figure it is clear that there are many inconsistent estimates of this important quantity. This suggests that we may indeed be in an environment characterized by uncertainty, rather than risk. There are two common responses to this assertion: Why not simply

[1] Ellsberg himself emphasizes that '...either the postulates failed to be acceptable in those circumstances as normative rules, or they failed to predict reflective choices...But from either point of view, it would follow that *there would be simply no way to infer meaningful probabilities for those events from their choices*, and theories which purported to describe their uncertainty in terms of probabilities would be quite inapplicable in that area.'

[2] Our work focuses on uncertainty about a key parameter of the climate system—climate sensitivity—and for the most part assumes that economic parameters are known. More on this later.

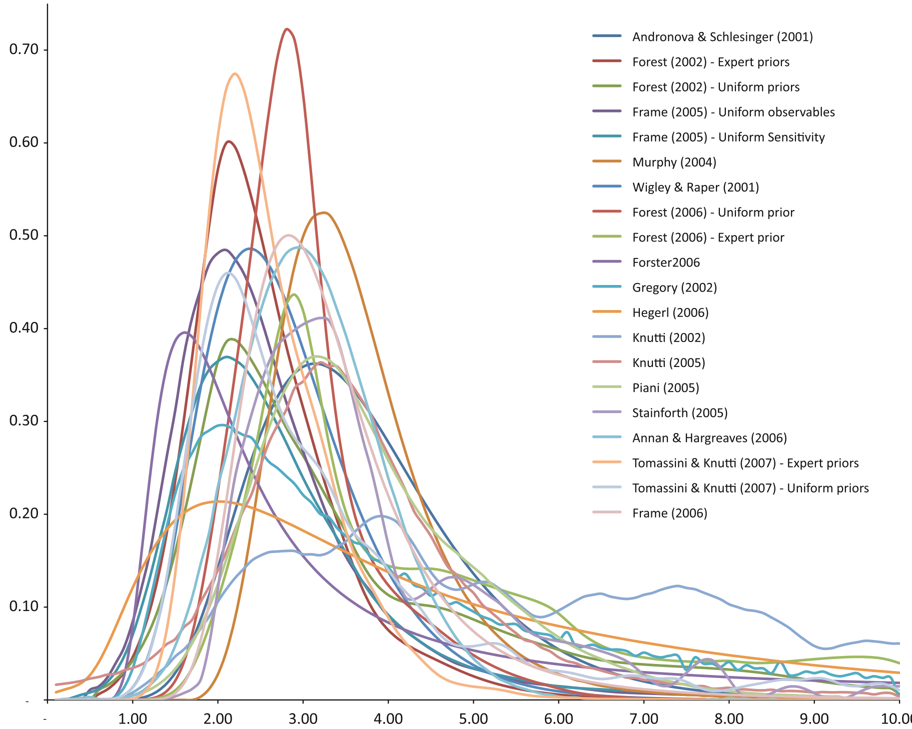

Fig. 1 Estimated probability density functions for climate sensitivity from a variety of published studies, collated by Meinshausen et al. (2009)

aggregate the different estimates into a single probability density ("Bayesian response"); and surely all the estimates are not equally valid, so why not simply choose the 'best' estimate ("scientists' response")? To answer these questions, it is important to understand that the estimates of climate sensitivity in Fig. 1 are based on different climate model structures (i.e. different representations of physical processes, and choices of physical parameters), different statistical methodologies, and different observational data. Climate science is currently unable to identify how these different estimates might best be combined. The problem of model comparison is made especially difficult by the fact that the estimates are not independent, and the fact that the historical instrumental record has already been used to generate the estimates in Fig. 1, thus precluding us from using it again to determine which of the models fits the data best (see Tebaldi and Knutti 2007 for a discussion of the difficulties of combining predictions from multiple climate models.). Thus we cannot quantify how the different estimates relate to one another, or objectively single out a 'best' study.

Any putative weights that we might assign to the different estimates would thus reflect largely subjective choices. Importantly, the standard expected utility paradigm does not distinguish between subjective choices (e.g. the weights on models in our application) and the conditionally objective knowledge generated by assuming a given model structure is correct and comparing it to the data (e.g. the probability densities in Fig. 1 in our application). In the standard Bayesian model subjective weights combine linearly with conditionally objective probabilities to give rise to a unique compound distribution which is indistinguishable from an equivalent objective lottery—we have lost sight of the fact that this compound distribution combines two very different kinds of knowledge.[3] If policy makers are sensitive to the difference between arbitrary subjective weights and objective probabilities that arise from empirical comparisons of models to data (surely a desirable quality in a planner), and would prefer to choose policies that, all else being equal, lead to less ambiguous (or more objective) distributions of outcomes, then there is a strong case that the traditional tools are not capable of reflecting policy makers' preferences.

Thankfully there has been a series of theoretical advances in decision theory which provide axiomatic representations of preferences accounting for the distinction between uncertainty and risk, and for ambiguity aversion. Seminal contributions include Arrow and Hurwicz (1977), Schmeidler (1989), Gilboa and Schmeidler (1989) and Klibanoff et al. (2005). These are elegant models, and have found application in several areas of economics, especially finance (e.g. Dow and da Costa Werlang 1992; Bassett et al. 2004; Gollier 2009; Bossaerts et al. 2010). Hansen and Sargent (2007) have applied similar techniques, later axiomatized by Maccheroni et al. (2006), in macroeconomics in order to derive policies that are robust to model misspecification. These methods have been applied to environmental problems by e.g. Gonzalez (2008) and Athanassoglou and Xepapadeas (2012).[4] Yet such models have only just begun to filter into the argument around climate policy. Henry and Henry (2002) is perhaps the first paper to view climate policy through this lens, and focuses on formalizing precautionary policies as policies that account for ambiguity in scientific knowledge. Lange and Treich (2008) provide some limited comparative statics results on the effect of ambiguity on optimal abatement in a parametric two-period model. Although not confined to climate applications, the work of Traeger (2009) and Gollier and Gierlinger (2008) on the effect of ambiguity aversion on the social discount rate is clearly relevant and has important implications for the assessment of mitigation investments. A recent working paper by Lemoine and Traeger (2012) extends the work of Kelly and Kolstad (1999) to account for learning about uncertain tipping points in the climate system. Knowledge about these tipping points is represented as ambiguous and the effect of ambiguity aversion on the social cost of carbon is examined.[5]

In this paper we hope to provide a further step along the path sketched out by these authors. Our contribution to the literature is two fold. In Sect. 2 we introduce the model of ambiguity averse choice obtained in Klibanoff et al. (2005) and derive some new and quite general conditions which, in a static context, ensure that an arbitrary increase in ambiguity aversion leads to an increase in optimal abatement. Examining the static case helps to build intuition for the new effects that ambiguity aversion introduces into decision making, demonstrates how the preference representation manages to maintain the distinction between subjective

[3] This somewhat complex point will be elaborated on in the next section.

[4] While the Hansen–Sargent robust control methods have analytic appeal due to their tractability, they are limited in other important respects. They assume the most extreme aversion to ambiguity (by adopting a max-min decision rule), are not able to reflect the likelihood information available in the scientific literature, and what analytic insights they do deliver are in practice limited to linear-quadratic models of questionable relevance to the climate problem.

[5] Our work focuses on ambiguity about climate sensitivity, and excludes tipping points, for several reasons. First, climate sensitivity has long been the most important summary statistic in climate change science, yet even this well known quantity is not well characterized in the scientific literature, and is arguably 'ambiguous'. Second, climate sensitivity features in most integrated assessment models in a similar manner, whereas those integrated assessment models that do account for tipping points do so idiosyncratically. Third, while climate sensitivity has a widely recognized meaning, and is very much part of the scientific mainstream, the science of tipping points, while undoubtedly of considerable economic importance, is still in its infancy (see Lenton et al. 2008). Since our aim in this paper is to stay as close as possible to well established science, we do not account for tipping points in our analysis.

and objective knowledge, and allows us to gain a high-level insight into how ambiguity aversion is likely to affect the choice of climate policy. In Sect. 3 we extend this analysis to the dynamic case, and attempt to understand how scientific ambiguity affects the welfare evaluation of dynamic climate change abatement policies. We compute welfare measures for realistic exogenous abatement pathways using the integrated assessment model DICE (Nordhaus 2008) and assess how ambiguity over the correct probability distribution for climate sensitivity affects the welfare benefits of abatement over a business as usual policy. We find that an increase in ambiguity aversion favors abatement—this ambiguity effect is small when damages are flat, and very significant when damages are steep at high temperatures. For steep damages, welfare is as sensitive to ambiguity aversion as it is to the elasticity of marginal utility, which controls risk aversion and consumption smoothing. We interpret these empirical results by appealing to theoretical work on the social discount rate under ambiguity. Section 4 discusses the results of our analysis, and concludes.

2 The Smooth Ambiguity Model and Optimal Abatement

A potential difficulty with several of the decision models that account for ambiguity is that they do not achieve a separation between ambiguous beliefs and attitudes towards ambiguity. This was overcome by the contribution of Klibanoff et al. (2005), who provided a preference representation that separates tastes from beliefs, and allows us to parameterize attitudes to ambiguity via a differentiable function, in a manner analogous to the way utility functions represent risk preferences.[6] Their formalism is thus perfectly suited to understanding how different degrees of ambiguity aversion affect policies and welfare estimates. We introduce their model below.

Define an 'act' *a la* Savage as a map between states and outcomes (see e.g. Gilboa 2009 for a detailed explanation of Savage acts). Klibanoff et al. (2005) define a set of axioms for preferences over ambiguous acts. They first assume that acts that are evaluated on *unique* lotteries have an expected utility representation. Next they define 'second order acts', which map probability distributions (not states) into outcomes. They assume that these acts also have an expected utility representation. Finally they assume consistency between first and second order acts—a second order act evaluated on a unique lottery yields the same valuation as the first order value of the lottery. Combining these axioms they show that act f is preferred to act g if and only if

$$\mathbf{E}_p\phi(\mathbf{E}_\pi u \circ f) > \mathbf{E}_p\phi(\mathbf{E}_\pi u \circ g), \tag{1}$$

where u is a von Neumann–Morgenstern utility function, ϕ is an increasing function, and p is a subjective second-order probability over a set Π of probability measures π that the decision maker (DM) deems to be relevant to her decision problem. To immediately give this model a climatic interpretation, suppose that Π is the set of distributions for climate sensitivity (S) in Fig. 1, indexed by $m \in \mathcal{M}$, and that the the choice variable is the level of abatement a of greenhouse gas emissions. For each probability model, write the expected utility obtained under that model as a function of a as $EU_m(a) = \int U(a(S))\pi_m(S)dS$.[7] Then the policy maker's objective function can be written as

[6] Ghirardato et al. (2004) achieve a similar separation between tastes and beliefs, but their framework is not well suited to the comparative statics analysis and dynamic applications we consider in this paper.

[7] Note that a denotes an act which maps states of the world (values of S) into payoffs ($a(S)$). We integrate over S, but sum over models m in Eq. (1), because within-model uncertainty is represented through a PDF over the continuous climate sensitivity parameter, but there is a discrete set of distributions for S.

$$V(a) = \sum_m p_m \phi(EU_m(a)) \tag{2}$$
$$= \mathbf{E}_{\text{subj}} \phi(\mathbf{E}_{\text{obj}} U(a)) \tag{3}$$

where p_m is a second order weight on probability model m, and the subscript notation in the second line emphasizes that the p_m are *subjective* weights between models, while the EU_m denote conditionally *objective* expectations taken within a given model. Notice that if ϕ is nonlinear, we can never write the objective function as a single expectation over a compound distribution—the nonlinearity of ϕ drives a wedge between subjective and objective knowledge.

If we assume further that $\phi'' < 0$, i.e. ϕ is concave, then we have that

$$\mathbf{E}_{\text{subj}} \phi(\mathbf{E}_{\text{obj}} U(a)) \leq \phi\left(\mathbf{E}_{\text{subj}\times\text{obj}} U(a)\right) \tag{4}$$

where $\mathbf{E}_{\text{subj}\times\text{obj}}$ denotes an expectation over the compound distribution $\sum_m p_m \pi_m$. This elementary inequality (a version of Jensen's inequality) demonstrates a vital property of the model. The concavity of ϕ ensures that if the subjective uncertainty were in fact objective, thus allowing all uncertainty to be described by a unique compound distribution [as in the expectation on the right hand side of (4)], this would always be preferred to the equivalent ambiguous situation where part of the uncertainty is subjective and part of it is objective [as in the left hand side of (4)]. The decision maker thus always prefers a policy whose consequences are objectively known to an equivalent policy whose consequences are in part determined by subjective choices. When ϕ is concave, we will say that the decision maker is ambiguity averse (Klibanoff et al. 2005).

From (2) we find that the first-order conditions can be written as

$$\sum_m \hat{p}_m(a^*) \left.\frac{dEU_m}{da}\right|_{a=a^*} = 0 \tag{5}$$

where a^* is the optimal abatement level, and we have defined 'ambiguity-adjusted' second-order probabilities $\hat{p}_m$ as:

$$\hat{p}_m(a^*) = \frac{\phi'(EU_m(a^*))p_m}{\sum_n \phi'(EU_n(a^*))p_n}. \tag{6}$$

Equation (5) just says that the weighted sum over models of marginal expected utility with respect to abatement should be zero, where the weighting factors are just the $\hat{p}_m$. It is identical to the condition one would obtain under ambiguity neutrality, except that p_m is replaced by $\hat{p}_m(a^*)$, the ambiguity-adjusted weight on model m.

If we look at (6), it is clear that the ambiguity weighting emphasizes those models that predict low expected utilities, since ϕ' is a decreasing function (see also Lange and Treich 2008). Moreover, an increase in ambiguity aversion puts more weight on models with low expected utilities, and less on those with high expected utilities. This can be formalized using the concept of monotone likelihood ratios. Gollier and Gierlinger (2008) prove the following result:

Proposition 1 *Suppose there is a set $\mathcal{M}$ of possible models of cardinality M, and indexed by m. Without loss of generality, assume that the EU_m are ordered such that $EU_1 \leq \ldots \leq EU_M$. Let $\phi_2 = f(\phi_1)$, where f is increasing and concave, and let $(\hat{p}^1_m)_{m\in\mathcal{M}}$, $(\hat{p}^2_m)_{m\in\mathcal{M}}$ be the ambiguity-adjusted second-order probabilities associated with ϕ_1 and ϕ_2 respectively, as given by Eq. (6). Then $(\hat{p}^1_m)_{m\in\mathcal{M}}$ dominates $(\hat{p}^2_m)_{m\in\mathcal{M}}$ in the sense of the monotone likelihood ratio order, i.e. $\left(\hat{p}^2_m/\hat{p}^1_m\right)_{m\in\mathcal{M}}$ is decreasing in m.*

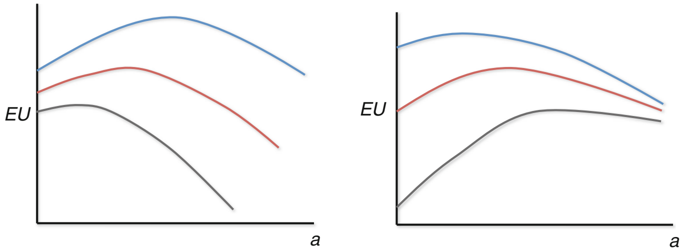

Fig. 2 Examples of comonotonic (*left*) and anti-comonotonic (*right*) relationships between the sequences $(EU_m(a))_{m\in\mathcal{M}}$ and $\left(\frac{dEU_m(a)}{da}\right)_{m\in\mathcal{M}}$

The proof is very simple—the ratio $\hat{p}_m^2/\hat{p}_m^1 \propto f'(\phi_1(EU_m))$, where the proportionality constant is independent of m. Since ϕ is increasing, f' is decreasing, and the EU_m are increasing in m, the factor on the right decreases when m increases. This provides a simple characterization of the effect of increased ambiguity aversion on the first-order condition. It is important to stress however that the weights $\left(\hat{p}_m\right)_{m\in\mathcal{M}}$ are endogenous to the optimization problem, as they depend on a^*. Thus translating how these weights change into statements about how optimal abatement changes when ambiguity aversion increases is a non-trivial task in general.

General results on the comparative statics of ambiguity aversion are hard to come by, and depend on the properties of the sequence of functions $(EU_m(a))_{m\in\mathcal{M}}$. The following proposition defines conditions on these functions that allow us to ascertain the effect of an arbitrary increase in ambiguity aversion on optimal abatement:[8]

Proposition 2 *Suppose that* $\frac{d^2EU_m}{da^2} < 0$ *for all m, and assume that for every fixed value of a, the sequences* $(EU_m(a))_{m\in\mathcal{M}}$ *and* $\left(\frac{dEU_m(a)}{da}\right)_{m\in\mathcal{M}}$ *are anti-comonotonic*[9] *(comonotonic) in m. Then an increase in ambiguity aversion increases (decreases) the optimal level of abatement for the objective* (2).

Proof See "Appendix". □

To understand the conditions on the functions $EU_m(a)$ in this proposition, consider the diagram in Fig. 2, in which abatement a is plotted horizontally and $EU_m(a)$ is plotted vertically for a model set containing 3 distinct models. Anti-comonotonicity means that for each value of a the models with low expected utilities have high derivatives of expected utility with respect to a—this case is represented in the right panel of Fig. 2. It is clear from the figure that when these conditions are satisfied an increase in abatement will reduce the spread of expected utilities across models, hence a rise in ambiguity aversion leads to more abatement. Conversely, if the sequences are comonotomic then a decrease in abatement will decrease the spread between expected utilities (represented in the left panel of Fig. 2) and so an increase in ambiguity aversion leads to a drop in abatement.

[8] Appendix A.2 of Lange and Treich (2008) investigates the comparative statics of ambiguity aversion in a simple parametric model. Our proposition considerably generalizes their analysis by abstracting from parametric assumptions.

[9] Two sequences are anti-comonotonic if one is increasing and the other is decreasing. They are comonotonic if they are both increasing, or both decreasing.

Note that an increase in ambiguity aversion always implies a policy change leading to a smaller spread of expected utilities. With anti-comonotonicity this means more abatement and with comonotonicity it means less.

Which (if any) of the conditions in this proposition holds in practice is an empirical question. The proposition provides us with sufficient, but not necessary, conditions—it is quite possible that neither condition is satisfied,[10] in which case we need to make further parametric assumptions in order to investigate the effect of ambiguity aversion on policy choice. To illustrate the application of the proposition however, consider the following example:

Example 1 Suppose that we have two models of the effect of greenhouse gases on climate:

- Model 1: Climate is not sensitive to anthropogenic emissions.
- Model 2: Climate is sensitive to anthropogenic emissions. Suppose that, conditional on this model being correct, there are two climate states, a high damage state H, and a low damage state L. Abatement increases the probability of state L materializing.

The DM attaches a default utility level u_0 to today's climate. Let $u_0 - d_H$ be the utility level in the high climate damages state, and $u_0 - d_L$ be the utility level in the low damages state, with $d_H > d_L$. The DM must decide the level of abatement, a. The utility cost of abatement is given by a function $\Lambda(a)$, and let $\pi(a)$ be the probability of the low-damages state occurring as a function of abatement. Clearly we need $\pi'(a) > 0$, $\Lambda'(a) > 0$. Proposition 2 allows us to obtain the following:

Proposition 3 *Suppose that* $\pi''(a) < 0$ *and* $\Lambda''(a) > 0$. *Then an increase in ambiguity aversion increases optimal abatement in Example* 1.

Proof The expected utilities obtained in models 1 and 2 are:

$$EU_1 = u_0 - \Lambda(a) \tag{7}$$

$$\begin{aligned} EU_2 &= \pi(a)(u_0 - d_L - \Lambda(a)) + (1 - \pi(a))(u_0 - d_H - \Lambda(a)) \\ &= u_0 - \Lambda(a) + \pi(a)(d_H - d_L) - d_H. \end{aligned} \tag{8}$$

Clearly, $EU_1 > EU_2$, since $\pi(a) \in [0, 1]$. In addition, simple differentiation shows that $\frac{dEU_1}{da} < \frac{dEU_2}{da}$. Thus the sequences (EU_1, EU_2) and $\left(\frac{dEU_1}{da}, \frac{dEU_2}{da}\right)$ are anti-comonotonic for all a. Finally, the conditions $\pi''(a) < 0$, $\Lambda''(a) > 0$ ensure that $\frac{d^2EU_m}{da^2} < 0$, $m = 1, 2$. Thus the conditions of Proposition 2 are satisfied, and the result is established.

□

At a very high level of abstraction we can ask which, if any, of the scenarios represented in Proposition 2 and Fig. 2, is likely to be relevant for understanding the effect of ambiguity on climate policy choice. There are three main sources of uncertainty in climate policy at the global scale—scientific uncertainty about the effect of increases in greenhouse gas concentrations on global climate, uncertainty about the damage function (i.e. how temperature changes translate into economic impacts), and technological uncertainty about the costs of abatement. Arguably, all of these uncertainties should be represented as ambiguous. At present the state of the empirical literature is that we have some credible probabilistic representations of scientific uncertainty (as in e.g. Fig. 1), while uncertainties in the damage function and technological change are at best represented by *ad hoc* assumptions based on

[10] For example, notice that the conditions of the proposition ensure that the curves $\{EU_m(a)\}$ never intersect. This is clearly a strong restriction if insisted upon globally.

loose extrapolations and intuition (see e.g. Nordhaus 2008 p. 126)—we simply do not have the data or the understanding to quantify these uncertainties probabilistically at this time. Thus, ironically, the least ambiguous uncertainty is the only one we can model credibly using formal tools. Nevertheless, it is clear that both scientific uncertainty and impacts uncertainty give rise to an exposure to ambiguity at *low* values of abatement, while technological uncertainty gives rise to an exposure to ambiguity at *high* values of abatement.

The key point to appreciate however is that the magnitude of the effects of these different uncertainties on aggregate welfare is very different—one would expect ambiguity in the science and impacts to give rise to much larger variations in expected welfare over a set of probabilistic models than what would arise from ambiguity in abatement costs. This is so because the downside of scientific and impacts uncertainty is both highly ambiguous (as Fig. 1 demonstrates) and deeply damaging to global welfare (Weitzman 2009, 2012). Even if aggressive abatement turns out to be very costly (e.g. about 10 % of global GDP in net present value terms for a 450ppm stabilization target—at the very high end of current estimates McJeon et al. 2011), these costs are likely to pale into insignificance when compared to the persistent global catastrophe that would occur if climate sensitivity is greater than say 6 °C.[11] Thus ambiguity in costs will translate into much smaller variations in aggregate welfare than ambiguity in the science and impacts. It thus seems likely that ambiguity is high for low levels of abatement, and that, at least for a reasonable range of abatement policies, it decreases as we increase abatement. Thus the anti-comonotonicity condition represented in the right panel of Fig. 2 is probably a reasonable, if abstract, first-order representation of the effect of ambiguity on climate policy.

3 Evaluating Dynamic Abatement Pathways Under Ambiguity

The preceding section abstracted the climate change abatement problem to a high level. Perhaps most importantly, the models examined thus far have all been atemporal. While static models are useful for gaining an intuition for the new effects that ambiguity aversion introduces into familiar problems, they are of limited use for deriving quantitative results about the effect of ambiguity on climate policy. The abatement problem is in its essentials a dynamic decision problem, in that it requires us to trade off near-term costs against long-term benefits. This section examines how scientific ambiguity affects the welfare analysis of alternative climate policies. We introduce the dynamic extension of the smooth ambiguity preference representation derived in Klibanoff et al. (2009) and apply it to the DICE integrated assessment model (Nordhaus 2008) to investigate how ambiguity about climate sensitivity, represented by the 20 distributions in Fig. 1, affects a measure of willingness to pay for abatement policy. As noted in the previous section, scientific ambiguity is only one of the channels through which ambiguity may affect welfare, however it is the only one we have credible (albeit ambiguous) probabilistic estimates of. It is also likely to be a dominant source of ambiguity because differences between estimates of the tails of the distribution for climate sensitivity have such large welfare consequences. Moreover, there is a large literature that focuses on the effects of uncertainty in climate sensitivity on climate policy (e.g. Weitzman 2009; Kelly and Kolstad 1999). Thus while our analysis is of necessity a partial treatment

[11] To put these numbers in perspective, real US GDP dropped by approximately 5 % (peak to trough) in the recent recession. Thus high end abatement cost estimates correspond to the cost of a single year of a recession about twice as bad—the rough equivalent of spending a year back in 1999 in GDP per capita terms. Average temperatures 6 °C higher than today's are substantially hotter than anything the earth has seen in the last 800,000 years (Jouzel et al. 2007).

of the full uncertainty around climate policy, it has the benefit of being rooted in credible empirical estimates, and allows a like-with-like comparison to the existing literature.

3.1 Dynamic Welfare Functions

The static ambiguity framework described in the previous section has been extended to the dynamic case in Klibanoff et al. (2009). They obtain a representation of preferences over time- and state-dependent acts, i.e. contingent plans that map the nodes of a decision tree into consumption streams. If we let $s^t = (x_1, \ldots, x_t) \in \Gamma$ denote a decision node, where $x_\tau \in \mathcal{X}_\tau$ is an observation at time τ and Γ is the set of all nodes, then a generic plan f maps s^t into consumption. Klibanoff et al. (2009) show that preferences over plans that satisfy consequentiality (any decision node is evaluated based only on those nodes that are reachable from it), dynamic consistency (plans made today are carried out tomorrow), and further axioms that are similar to those employed in the static representation result, can be represented by a function V which is recursively defined through the equation,

$$V_{s_t}(f) = U(f(s^t)) + \beta\phi^{-1}\left[\int_{\Theta}\phi\left(\int_{\mathcal{X}_{t+1}} V_{(s^t,x_{t+1})}(f)d\pi_\theta(x_{t+1};s^t)\right)dp(\theta|s^t)\right], \quad (9)$$

where $\theta \in \Theta$ indexes the set of alternative probability models, and $\beta \in [0, 1]$ is a discount factor. The first term is the current payoff of the plan f, while the second term is the continuation value of the plan, which is a nonlinear ϕ-weighted average over the set of models Θ for the evolution of the state variables x_t. We will use this representation to evaluate exogenous abatement policies under two simplifying assumptions about learning under ambiguity.

Introducing endogenous learning into the problem leads to an explosion in the dimensionality of the model's state space in our application. A full dynamic learning model that stays faithful to the nature of the uncertainty in Fig. 1 would require three levels of uncertainty—random noise in the state equations, parametric uncertainty about the value of climate sensitivity, and uncertainty over which of the priors in Fig. 1 is correct. Thus as new realizations of the noise are observed we would need to update both the parameters of each distribution for climate sensitivity in the standard Bayesian fashion, as well as the weights on models, in the manner of Bayesian model averaging (Hoeting et al. 1999). Both the parameters of the distributions, and the weights on models, thus become state variables of the dynamic programming problem, leading to a very high dimensional problem that will not yield to standard numerical optimal control methods (see for an example of the computational complexities involved in accounting for learning about climate sensitivity in a much simpler context Kelly and Kolstad 1999). To avoid these complexities we simply compute welfare for two sample exogenous learning scenarios—one in which ambiguity resolves after one time step, and one in which ambiguity persists unchanged for all time. These scenarios are not intended to be realistic—they are useful precisely because they are unrealistic extremes which can be thought of as bounding the space of learning scenarios. Since we will also be fixing the control variables in our numerical analysis, the (time-dependent) distributions of consumption that will enter our welfare analysis will be completely specified by a choice of controls and the learning scenario, both of which are exogenously specified.

Suppose that we have M probability distributions $\pi_m(S)$ for climate sensitivity, and that, given exogenous policy variables, consumption at time t when climate sensitivity is S is given by $c_t(S)$. Then if ambiguity resolves after one time step, the representation in (9) reduces to

$$V^{res} = U(c_0) + \beta\phi^{-1}\left[\sum_m p_m \phi\left(\sum_{t=1}^{T} \beta^{t-1} \int U(c_t(S))\pi_m(S)dS\right)\right] \tag{10}$$

where p_m is again the subjective weight on probability model m, T is the time horizon, and the superscript on V serves to remind us that ambiguity is immediately *resolved* under this welfare function. The structure of the representation is clear—at $t = 1$ all ambiguity is resolved, we learn which model is correct, and hence welfare reduces to discounted expected utility, the argument of ϕ in the round brackets. At $t = 0$ however we don't know which model is correct, so we must aggregate over the space of models using ϕ to represent our aversion to ambiguity, which in addition preserves the distinction between subjective (p_m) and conditionally objective ($\pi_m(S)$) uncertainty.

When ambiguity persists for all time the nonlinear recursive structure of the welfare function in (9) must be preserved at all time steps to reflect the presence of ambiguity. Unlike the resolved ambiguity case, the ambiguity preference ϕ always operates on the continuation value in this case, as it is always ambiguous. Examining the representation, and imposing constant exogenous distributions over S at all times, we find that in this case welfare can be calculated using the following recurrence relations:

$$\tilde{V}_{T+1} = 0 \tag{11}$$

$$\tilde{V}_{j-1} = \phi^{-1}\left(\sum_m p_m \phi\left(\int U(c_{j-1}(S))\pi_m(S)dS + \beta\tilde{V}_j\right)\right), \quad j = T+1\ldots 1 \tag{12}$$

$$V^{per} = U(c_0) + \beta\tilde{V}_1 \tag{13}$$

where the superscript on V reminds us that ambiguity *persists* under this welfare function. Note that we evaluate welfare in this case by starting in the terminal period and stepping forward to the first period, as the recursion in (9) requires. The recursive structure of this welfare function, and the fact that expectations are taken with respect to the same set of distributions $\{\pi_m(S)\}_{m=1\ldots M}$ at all times, reflects the persistence of ambiguity.

It will be convenient in what follows to have a means of representing the welfare difference between two policies in consumption units, since this makes welfare changes for different preferences directly comparable. Because we must potentially deal with non-marginal changes in welfare (because very high climate sensitivities could lead to very large damages), the measure we will use in order to convert welfare changes into consumption units is the Stationary Equivalent[12] (SE) (Weitzman 1976). The SE of a welfare function V is defined as the value of per capita consumption $c(V)$ which, when held constant over time, gives rise to welfare equivalent to V. Thus we define $c(V)$ through

$$\sum_{t=0} \beta^t P_t U(c(V)) = V, \tag{14}$$

where P_t is the population at time t. Define the fractional change in the SE induced by an abatement policy, relative to a business as usual (hereafter BAU) baseline, as Δ. A simple calculation shows that when U is of the constant relative risk aversion (CRRA) form [see (17) below], Δ is given by:

[12] Stern (2007) makes use of a 'balanced growth equivalent'. Relative changes in the balanced growth equivalent are identical to those for the stationary equivalent when the utility function is iso-elastic, as in our simulations below.

$$\Delta := \frac{c(V_{ABATE}) - c(V_{BAU})}{c(V_{BAU})} = \left(\frac{V_{ABATE}}{V_{BAU}}\right)^{\frac{1}{1-\eta}} - 1, \tag{15}$$

where η is the elasticity of marginal utility, and $ABATE$ denotes a generic abatement policy. All the results that follow use Δ to represent welfare differences between policies.

3.2 Application to the DICE Model

To make the discussion concrete, we will now use the DICE model (Nordhaus 2008) to analyze the empirical effect of ambiguity on the welfare analysis of climate policy, focusing on ambiguity over climate sensitivity. DICE is a well known integrated assessment model of the connections between economic activity and climate change. A standard Ramsey-Cass-Koopmans growth model with aggregate capital and labour inputs is linked to climate change through emissions of greenhouse gases, which cause global warming and, with a lag, reduce output by means of a reduced-form 'damage function' (more on this later). This damage function incorporates assumptions about adaptation to climate change, which can reduce output losses, leaving the representative agent with the choice of how much to invest in emissions abatement, as well, of course, as how much to save for investment in the composite capital good. The model includes many economic and climate parameters, all of which are at least to some extent uncertain. However for the reasons discussed above we will focus our attention on the climate sensitivity parameter, and how the ambiguity in this parameter represented in Fig. 1 affects the welfare benefits of abatement over a business as usual policy.

If we denote climate sensitivity by S, the exogenous savings rate by $\sigma(t)$, and exogenous abatement effort by $a(t)$, then we view DICE as the following function:

$$\text{DICE}\,(S; \sigma(t), a(t)) = c_t(S), \tag{16}$$

where $c_t(S)$ is an S-dependent stream of consumption per capita.[13] We can compute this function for a variety of values of S, holding $\sigma(t)$ and $a(t)$ constant, to see how the consumption stream depends on climate sensitivity. Specifying exogenous distributions for S at each point in time, and using the function $c_t(S)$, we thus induce an ambiguous set of distributions over consumption in each time period.

In our empirical results below, we pick specific exogenous values for the controls $\sigma(t)$ and $a(t)$. For simplicity, we assume that the savings rate $\sigma(t)$ is a constant 22 %, the default value in DICE. Abatement effort is represented in DICE by the emissions control rate, a number between 0 and 1, which controls the emissions intensity of gross economic output (i.e. before climate damages are incurred). When the control rate is $a(t)$, a fraction $1-a(t)$ of gross output contributes to emissions. Our two scenarios for the control rate are a 'Business as usual' scenario, and a scenario that limits the atmospheric concentration of CO_2 to twice its pre-industrial level (560 parts per million, hereafter referred to as the 2 CO_2 scenario).[14] The 2CO_2 abatement scenario has been prominent in recent international negotiations about

[13] Explicitly, $c_t(S) = (1 - \sigma(t)) * Y_t / P_t$, where Y_t is output, which is determined by previous savings and abatement decisions and the value of S, and P_t is the exogenous population size.

[14] Note that the control rates associated with the abatement scenario are designed so that they achieve the specified stabilization targets in an idealized run of DICE in which damages are zero for concentrations below the stabilization target, and rise sharply to very high values above the target. This is the method used by Nordhaus (2008) to generate controls that achieve a given stabilization target, however it should be born in mind that these controls will not in general achieve this target when they are used as inputs to DICE for alternative damage functions, or values of S. Since all we require for our purposes is a plausible choice of controls, we need not concern ourselves too much with whether they achieve a given stabilization target.

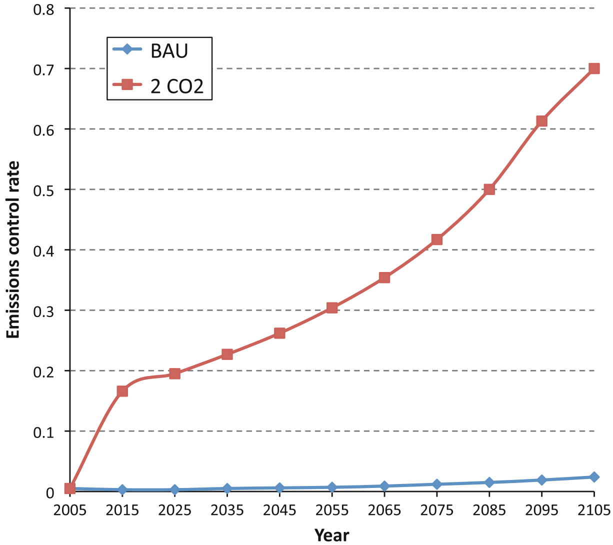

Fig. 3 Emissions control rates for our sample abatement scenarios

climate policy. The emissions control rates corresponding to our two scenarios are depicted in Fig. 3. They are each taken from Nordhaus (2008).

Since we have no means of objectively discriminating between the 20 alternative probability models in Fig. 1, we will invoke the principle of insufficient reason and assume equal subjective weight $p_m = 1/M$ on each of them.[15] We choose U and ϕ both to be isoelastic functions (constant relative risk and ambiguity aversion respectively):

$$U(c) = \frac{c^{1-\eta}}{1-\eta}, \quad \eta \neq 1 \tag{17}$$

$$\phi(U) = \begin{cases} \frac{U^{1-\xi}}{1-\xi} & \eta < 1 \\ \frac{-(-U)^{1+\xi}}{1+\xi} & \eta > 1 \end{cases}, \tag{18}$$

where η (ξ) is the coefficient of relative risk (ambiguity) aversion. Note that since ϕ operates on utility, which is measured in different units to consumption and will usually be estimated over a quite different range of absolute values, larger values of ξ are considered plausible, compared with η. Gollier and Gierlinger (2008) calibrated the above model to the experimental results reported in Camerer (1999) and found $\xi \in [5, 10]$ to be plausible values for the coefficient of relative ambiguity aversion. We will take the upper estimate of $\xi = 10$ as our representative case of an ambiguity-averse decision maker, but will conduct sensitivity analysis for a wider range of values where this is illuminating. See Ju and Miao (2012), for a more detailed discussion of calibration of parameters for a related preference representation.

Aside from performing sensitivity analysis over the two preference parameters η and ξ, we will also investigate how the effects of ambiguity interact with two other components of the model: the utility discount rate (also known as the pure rate of time preference), and the damage function. We consider two values for the utility discount rate—0.1 and 1.5 %

[15] Other choices for the weights p_m, which may reflect partial quantifications of model dependence and model performance, are of course possible. Current scientific techniques for achieving this are however tentative. See e.g. Knutti et al. (2010), Knutti (2010) for discussion.

per year. The lower value is justified on ethical grounds, and represents recent viewpoints by, for example, Stern (2008), Dasgupta (2008), and Heal (2009). The higher value is advocated by Nordhaus (2008). For the purposes of this paper we remain agnostic about which of these is more appropriate for welfare analysis of climate change, and investigate the effect of scientific ambiguity in both cases. The damage function in DICE has the form,

$$\Omega_{DICE}(T) = \frac{1}{1 + \alpha_1 T + \alpha_2 T^2}, \tag{19}$$

where T is the increase in global mean temperature above the pre-industrial level, and α_1 and α_2 are coefficients. The damage from warming as a fraction of GDP is $1 - \Omega_{DICE}(t)$, which with the default calibration of $\alpha_1 = 0$ and $\alpha_2 = 0.0028$, gives damages of 1.7 % of GDP for 2.5 °C warming and 6.5 % for 5 °C warming. An increase in the global mean temperature of 5 °C over the pre-industrial level is greater than the difference in temperature between the present day and the peak of the last ice age, the Last Glacial Maximum, and is thus expected to lead to biophysical changes of an unprecedented magnitude (in the context of human experience), occurring at an unprecedented rate (even in the context of geological history). Because of this, the scientific community is essentially reduced to speculation about the consequences of around 5 °C warming or more. The usual approach to calibrating the damage function is to make an estimate of the output loss accompanying 2.5 °C warming or thereabouts, and then to extrapolate to higher temperatures based on an assumed functional form that is essentially unsupported by data of any sort.

A number of scholars, including Weitzman (2012), consider predicted damages of 6.5 % of GDP given 5 °C warming to be remarkably low, and suggest that the damage function should be revisited (see also Ackerman et al. 2010). We thus also consider the damage function advocated in Weitzman (2012), and replace (19) with a damage function which replicates the DICE damage function's behaviour at better understood, lower temperatures, but which exhibits rapidly increasing damages for higher temperatures:

$$\Omega_{WEITZ}(T) = \frac{1}{1 + (\tilde{\alpha}_1 T)^2 + (\tilde{\alpha}_2 T)^\gamma}, \tag{20}$$

where $\tilde{\alpha}_1 = 0.049$, $\tilde{\alpha}_2 = 0.16$ and $\gamma = 6.75$. Like the function in (19), this choice gives damages of 1.7 % of GDP for 2.5 °C warming and still only 5.1 % of GDP for 3.5 °C warming, before increasing sharply to give damages of 9 % of GDP for 4 °C warming, 25 % for 5 °C warming, and so on. Our results below display the welfare effects of ambiguity in all four combinations of assumptions about the discount rate and damage function: low discount rate + DICE damages, high discount rate + DICE damages, low discount rate + WEITZ damages, and high discount rate + WEITZ damages.

As a base set of results, Fig. 4 plots Δ, the percentage change in the stationary equivalent under the $2CO_2$ abatement strategy relative to business as usual, as a function of the coefficient of relative ambiguity aversion (ξ), at $\eta = 2$.[16]

The figure shows that Δ increases as ξ increases. For the relatively flat DICE damages, Δ increases approximately linearly with ξ over the domain in the figure, and ambiguity aversion has a small effect on the magnitude of the welfare differences between the $2CO_2$ and BAU policies. For the steeper WEITZ damages however, there is a clear nonlinear dependence between Δ and ξ, and the magnitude of the ambiguity effect is much larger. For example if the discount rate is 0.1 %, setting $\xi = 10$ gives rise to about a 45 % difference between the $2CO_2$ and BAU policies, as opposed to only 20 % at $\xi = 0$.

[16] $\eta = 2$ is advocated in, for example, Nordhaus (2008), Weitzman (2007).

Figure 5 considers the relative importance of ambiguity preferences compared with two other important preferences, both represented by the elasticity of marginal utility (η). η plays two roles in the standard discounted utilitarian welfare function—it encodes both risk preferences and preferences for consumption smoothing over time.[17] While these two preferences can be disentangled (e.g.Kreps and Porteus 1978; Traeger 2009), the standard formulation, and almost all applied work in climate change economics, treats them with a single parameter. We thus investigate the dependence of welfare differences on η and ξ concurrently in Fig. 5, which plots Δ as a function of η for $\xi = 0$ and $\xi = 10$.

The figure shows that for the flat DICE damages, the effect of ξ on welfare is much smaller than that of η. Δ is much more sensitive to changes in η, represented by the vertical variation in the curves in the two left panels, than it is to changes in the value of ξ, represented by the difference between the black ambiguity neutral curve ($\xi = 0$) and the blue and red ambiguity averse curves ($\xi = 10$). The differences between ambiguity neutrality and ambiguity aversion are so small relative to the effect of η that all three curves lie nearly on top of one another in this case. The situation is however radically different under the steeper WEITZ damage function. In this case the difference between the ambiguity neutral (black) and ambiguity averse (blue and red) curves is comparable in magnitude to the effect of a change in η on welfare. Thus regardless of the value of the discount rate, accounting for ambiguity aversion is just as important as accounting for risk aversion and consumption smoothing under the WEITZ damage function, and has a negligible effect under the DICE damage function.

3.3 Discussion of Results

In this section we discuss the qualitative features of our main simulation results in Figs. 4 and 5, show how they can be understood in terms of simple intuition, and how they are related to theoretical work on the social discount rate under ambiguity.

At the coarsest level there are two relevant differences between the BAU and $2CO_2$ policies that are important for understanding the effect of ambiguity aversion on their welfare differences in Fig. 4—the level of average consumption under each policy, and the ambiguity in the distribution of future consumption under each policy. It is vital to understand that *both* these quantities determine the ranking of policies—for example, one would not prefer a policy with a very low level of future consumption that all models agree upon to a policy with a high, but ambiguous, level of future consumption unless one had a very extreme aversion to ambiguity. The difference between the average level of consumption in the $2CO_2$ and BAU policies is represented by the point $\xi = 0$ in Fig. 4. The fact that Δ is positive at $\xi = 0$ under all discount rates and damage functions shows that the $2CO_2$ policy is welfare improving, even in the absence of ambiguity aversion. Now it is also the case that the $2CO_2$ policy has less exposure to scientific ambiguity than the BAU policy. This is obvious, as the higher are emissions, the more CO_2 concentrations change, the more temperature rises (mediated by the climate sensitivity parameter), and the more consumption falls (mediated by the damage function). Thus ambiguity in climate sensitivity translates directly into ambiguity in future consumption. Since high emissions policies have a greater share of consumption that is determined by the climate components of the model, ambiguity in those components translates into more ambiguity in consumption than would arise under a more stringent abatement policy. The fact that the $2CO_2$ policy has both higher average consumption, and is less ambiguous than the BAU policy, means that we must observe a monotonically increasing dependence of

[17] The elasticity of inter temporal substitution is just $1/\eta$ in the discounted utilitarian framework with an iso-elastic utility function.

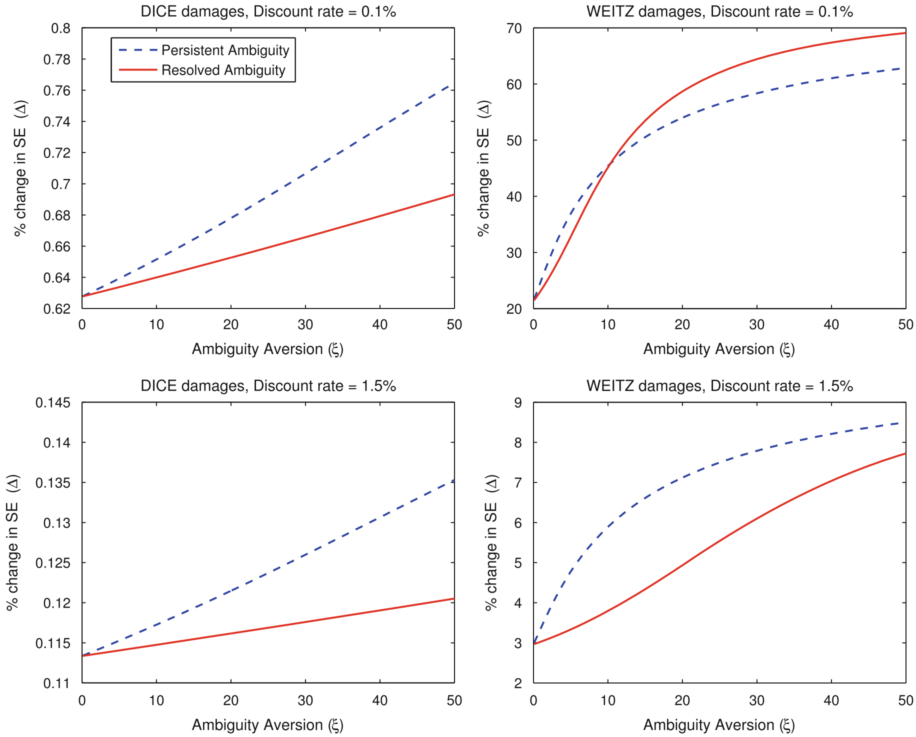

Fig. 4 Percentage change in SE of abatement relative to BAU for the 2 CO_2 (560 ppm) abatement pathway, as a function of the coefficient of relative ambiguity aversion (ξ). Estimates based on the assumption that ambiguity resolves after one 30 year time step are plotted in *red*, while estimates based on the assumption that ambiguity is perfectly persistent are plotted in *blue*. $\eta = 2$ in these simulations. Note the change of scale in the vertical axes. (Color figure online)

Δ on ξ—preference reversals are only possible if one policy has a low average consumption level but is not very ambiguous, while the other has a high average consumption level but is ambiguous.

Now consider the differences between the effect of ambiguity on welfare under the two damage functions in Fig. 4—i.e. compare the left hand to the right hand panels. Because damages are small for low temperatures under both damage functions, it is really the variation in the tails of the sensitivity distributions that is important for understanding the difference between the effect of ambiguity under these two damage functions. Variations in the distributions at low values of S translate into very small variations in consumption because at these values of S even the BAU policy does not achieve much warming. Variations in the tails of the distribution are quantitatively important however both because at high values of S large temperature changes are attained, and because the welfare effects of these temperature changes are magnified by the damage function. Now for the flat DICE damage function even the variation in the tails of the distributions for S does not affect consumption much—this gives rise to the approximately linear dependence of Δ on ξ in this case. One can think of this case as a first order Taylor approximation of the welfare function about $\xi = 0$ which holds well over the values of ξ considered in Fig. 4 because the inter-model spread in predicted future consumption is small. For the steep WEITZ damages however, the damage function amplifies the spread in the tails of the sensitivity distributions into substantial spreads in future

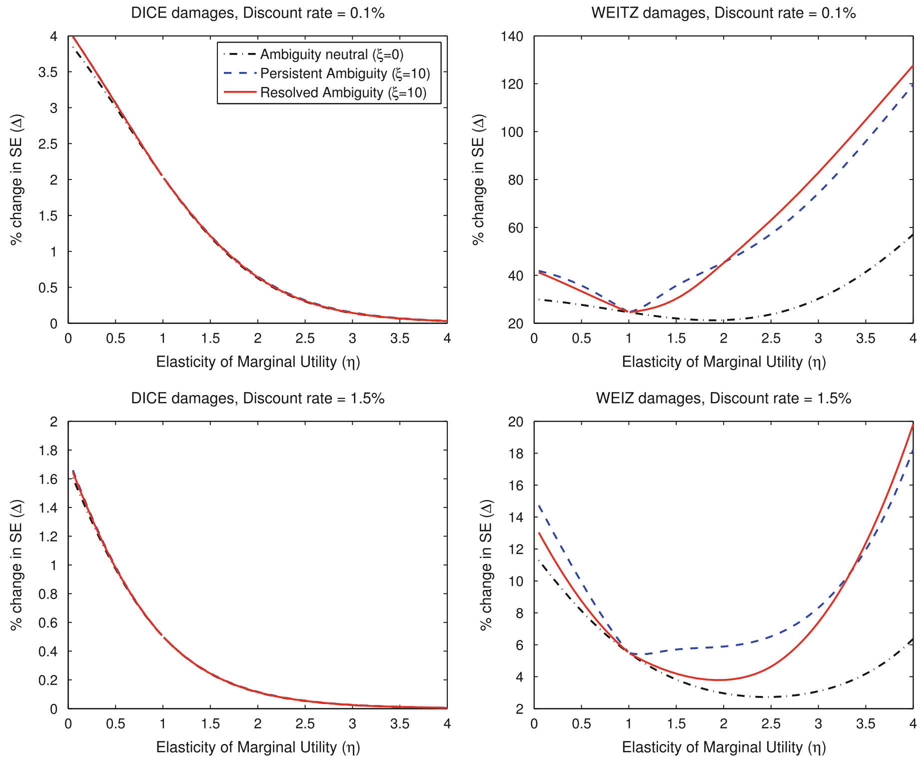

Fig. 5 Percentage change in SE of the 2 CO_2 abatement pathway relative to BAU, as a function of the coefficient of relative risk aversion (η). The *black line* plots the relationship when the coefficient of relative ambiguity aversion $\xi = 0$, while the *blue and red lines* plot the relationship when $\xi = 10$ for persistent and immediately resolved ambiguity respectively. (Color figure online)

consumption distributions—these are in turn amplified into large differences in expected utility by the sharp curvature of the utility function at low consumption values (Weitzman 2009). Thus in this case a linear approximation to the welfare function fails, and the nonlinear dependence of Δ on ξ is revealed. Notice that for the WEITZ damages Δ asymptotes to a constant value—this occurs since as ξ increases more and more weight is placed on models with low expected utility,[18] until for very large ξ these models dominate the welfare evaluation entirely and Δ reflects the welfare difference between the worst model under each policy.

In order to understand the qualitative features of Fig. 5 we contrast it with the expression for the certainty-equivalent social discount rate (ρ) under ambiguity aversion, derived in Traeger (2009) and Gollier and Gierlinger (2008). These derivations assume isoelastic forms for u and ϕ, and that consumption grows at an uncertain rate g which is normally distributed $g \sim \mathcal{N}(\theta, \sigma)$, where the mean growth rate θ is itself uncertain and distributed according to a second-order subjective probability distribution, $\theta \sim \mathcal{N}(\mu, \sigma_0)$. It can then be shown that

$$\rho = \delta + \eta\mu - \frac{\eta^2}{2}(\sigma^2 + \sigma_0^2) - \xi|1 - \eta^2|\frac{\sigma_0^2}{2}, \tag{21}$$

[18] For the resolved ambiguity welfare function, this is the model with the lowest discounted expected utility. The persistent ambiguity welfare function however picks out the model with the lowest expect utility at each point in time.

where δ is the utility discount rate. The first two terms in this expression are familiar from the standard Ramsey formula for the social discount rate under certainty, and capture pure time preference and inter-temporal substitution effects respectively. The third term is the standard correction due to risk aversion and the uncertainty in the growth rate (note that the variance of the growth rate is the variance of the composite distribution that arises from the combination of uncertainty about g and θ). The final term is a new addition due to ambiguity aversion.

Since differences in the expected utilities of abatement and BAU pathways only manifest significantly in the distant future (when damages are large), it is intuitive that the social discount rate may play a useful role in explaining our results. If ρ is large, we expect welfare differences between abatement and BAU policies to be small. Similarly, those values of η for which ρ is highly sensitive to ξ are also likely to be the values of η for which we see a significant effect of ambiguity aversion on the welfare difference measure Δ. Although the assumptions about the first- and second-order uncertainty on consumption growth rates used to derive (21) do not map exactly onto our empirical application, the expression nevertheless provides useful qualitative insights.

Focus first on the results for the DICE damage function, depicted in the left panels of Fig. 5. Empirically, we find that for our simulations with the DICE damage function, $\mu \gg \sigma > \sigma_0$ at each point in time. Neither growth rates, nor their variances, are constant in our simulations, with effective annual growth rates tending to fall from approximately 2 to 1 % over our 200-year simulation horizon. However the ratio of inter- and intra-model growth rate variances to the mean growth rate is smaller than 10^{-6} for every time step for the DICE damage function. With η of order 1, and ξ of order 10, it is clear that the last two terms in the expression for the discount rate are thus negligible, and all of the discounting work is done by the pure rate of time preference δ and the consumption smoothing term $\eta\mu$. As we increase η the discount rate thus increases monotonically (the quadratic terms are negligible), and the benefits of the $2CO_2$ abatement policy are monotonically discounted away. This is clearly reflected in the figure, which shows a strong monotonic decrease in the welfare benefits of abatement as η increases. The fact that $\mu \gg \sigma_0$ also explains why we observe such a small difference between the ambiguity neutral and ambiguity averse curves for the DICE damage function. The inter-model variation in the consumption distribution is simply too small to compete with the strong inter-temporal substitution effects.

Now consider the results of our simulations with the WEITZ damage function, depicted in the right panels of Fig. 5. These results are strikingly different. First, notice from the scale of the vertical axis that Δ is in general significantly larger, as we would have expected given the much greater damages at high temperatures that the abatement policy avoids under this damage function. Second, and most importantly, notice that Δ is now much more sensitive to ξ. Indeed, the difference between the ambiguity neutral and ambiguity averse curves is now easy to see, with the gap between abatement and BAU policies widening substantially once ambiguity aversion is accounted for. While $\Delta_{\xi=0}$ and $\Delta_{\xi=10}$ necessarily coincide at $\eta = 1$ [as we can see by consulting expression (21)], the ambiguity effect is large for most other values of η, and in particular, increases when η rises above unity. We can again explain this result by recourse to the expression for the social discount rate in (21). With high damages, the intra- and, critically, inter-model variance in consumption growth is now much larger in relation to mean consumption growth.[19] Thus the third and fourth terms in (21) are larger and work against the positive first and second terms more effectively, causing ρ to decline

[19] In the twenty-second century, the ratios of inter- and intra-model spreads in consumption growth to mean consumption growth increase by 1–2 orders of magnitude for the WEITZ damage function relative to their values for the DICE damage function.

significantly relative to its value for the DICE damage function. This gives rise to an ambiguity effect that can be very large for the higher values of η. The non-monotonic dependence of Δ on η is also explained by (21). Notice that the inter-temporal substitution terms in the discount rate are linear in η, while the risk and ambiguity terms are quadratic in η. For $\eta < 1$, a small increase in η increases the linear terms more than it decreases the quadratic terms, thus increasing the discount rate and leading to a decline in Δ. For $\eta > 1$ however a small increase in η decreases the quadratic terms more than it increases the linear terms and thus Δ starts to bend upwards if the uncertainty terms are sufficiently large.[20] Clearly the ambiguity effect also gets more weight in the welfare computation when $\eta > 1$, thus explaining the large divergence between the ambiguity averse and ambiguity neutral curves at high values of η.

4 Conclusions

This paper aimed to provide insight into how ambiguous knowledge, and aversion to ambiguity, affect the welfare analysis of climate change abatement policy. We have argued that our knowledge of the climate system is not of sufficient quality to be described by a unique probability distribution—there are many competing models of the response of climate to greenhouse gas emissions, and at present we do not have an objective means of discerning between them. Neglecting this model uncertainty is not a defensible strategy, as arbitrarily singling out one model for special treatment, or weighting the models and treating the result as if it were the output of a single objective model, fundamentally misrepresents the nature of our uncertainty. While some form of weighting procedure is likely to be unavoidable if one is to exploit the information in the model set, we would like to have a decision framework that recognizes the epistemic difference between conditionally objective model probabilities and arbitrarily chosen subjective weights on models, and, more generally, allows us to reflect a rational preference for policies with less ambiguous consequences. We have investigated one such framework—the smooth ambiguity model—which is easy to implement and interpret—two attractive features when it comes to policy applications.

Our analytical work in the static case developed intuition for the new effects that ambiguity aversion introduces into decision making. In particular, we derived sufficient conditions that enable us to determine how an increase in ambiguity aversion affects optimal policy choice without making any parametric assumptions. These conditions show the mechanism behind the new effects ambiguity aversion introduces—policies are pushed in the direction of decreasing spread in the set of expected utilities. Under the conditions of our proposition we need to know only very general qualitative information about how model expected utilities are ranked and depend on policy variables, but in general further parametric assumptions would be necessary to ascertain the effect of ambiguity aversion. Nevertheless we argued that these insights are relevant to understanding how ambiguity aversion might affect climate policy choice at a highly stylized level. Exposure to scientific and impacts ambiguity is high for high emissions policies, and exposure to cost ambiguity is high for low emissions policies, however the quantitative magnitude of these effects is likely to be very different, with scientific and impacts ambiguity giving rise to much greater variations in the distribution of future consumption than cost ambiguity. Thus, at a very coarse level, it would seem that accounting for ambiguity aversion is likely to push us towards more aggressive abatement policies.

[20] This follows trivially from the fact that $\frac{d}{d\eta}\eta = 1$, $\frac{d}{d\eta}\frac{\eta^2}{2} = \eta$.

In our dynamic application we attempted to answer the following questions: Is the ambiguity effect likely to be empirically important, and how does it compare to the effect of other preference parameters on welfare calculations? To investigate these questions we performed welfare calculations that account for scientific ambiguity, represented by 20 published estimates of the probability distribution for climate sensitivity. We showed that the relative welfare benefits of an abatement policy that stabilizes CO_2 concentrations at twice their preindustrial level over a business as usual policy increase as ambiguity aversion increases. This is a small effect if damages are flat, and a large and quantitatively important effect if damages are steep at high temperatures. Moreover, when damages are steep, the effects of ambiguity aversion are comparable in magnitude to the effects of the preferences encoded by the elasticity of marginal utility—risk aversion and inter-temporal consumption smoothing. This is true regardless of the value of the utility discount rate. Neglecting ambiguity thus runs the risk of understating, perhaps drastically, the welfare benefits of abatement.

While scientific uncertainty is only one channel through which the effects of ambiguity may be transmitted, it is the only one for which credible empirical quantifications of ambiguity exist, and thus the only one amenable to meaningful formal modeling at this time. Future work that accounts for ambiguity in impacts and abatement costs when empirical estimates become available will doubtless change the magnitude of our results, however we would be surprised if they change their direction, as the welfare effects of ambiguity in the damage function (which favor strong abatement) will likely dwarf the effects of cost ambiguity (which favors weak abatement). If anything, given the strong ambiguity effect we observe when damages are steep, it seems likely that a comprehensive account of ambiguity in all the model components will give rise to even larger 'ambiguity premia' than we have observed in our simulations.

There are some next steps that would be desirable extensions of our work. Fully optimizing stochastic control models with endogenous learning are likely to be computationally out of reach for some time, at least if we wish to stay true to the large number of distributions in Fig. 1.[21] As is so common in modeling, there is a stark trade-off between feasibility and fidelity to the data. Nevertheless, more limited empirical applications that focus on a smaller set of distributions may yield qualitatively useful insights about how learning and ambiguity interact in the determination of optimal policy choices. More robust calibrations of the smooth ambiguity model would also be desirable, and are currently being pursued in the experimental literature. When it comes to policy implications however, best practice will always be to present results for a reasonable range of preference parameters, reducing the importance of precise calibrations.

At a more fundamental level, there are some features of the smooth ambiguity model that one might wish to extend. While the separation between ambiguity preferences and ambiguous beliefs that the representation achieves is a desirable property, it may not always be justified. The very general representation obtained in Schmeidler (1989) provides an alternative, but suffers from problems of its own.[22] In addition, we have assumed that our model's state-space accurately describes the evolution of the climate-economy system, i.e. that there are no 'surprises'. This seems a strong assumption, given known inadequacies

[21] See the working papers by Lemoine and Traeger (2012), Cai et al. (2012) for the current computational state of the art, which permits endogenous learning in models with 3–5 state variables. For comparison, even a simplified representation of endogenous learning about the set of distributions in Fig. 1 would require $2 \times 20 + 19 = 59$ informational state variables.

[22] These include the fact that the representation theorem depends on an explicit 'uncertainty aversion' axiom. One would hope that such a behavioural constraint would be an optional special case of the representation (much as risk aversion is in expected utility theory), rather than a primitive requirement.

in our understanding of the climate system, not to mention economic aberrations. Decision theories that account for these deficiencies in our understanding have been proposed (e.g. Gilboa and Schmeidler 1995), although it is as yet unclear whether they can be usefully applied in the climate change context.

Reflective normative work should feel no shame in pursuing these leads, and challenging the privileged place that expected utility theory currently occupies. Our work suggests that this is no mere academic exercise. The decision tools we employ have as strong an effect on policy recommendations as the empirical inputs to our models. It is surely vital to make sure that the tools fit the task.

Appendix: Proof of Proposition 2

Start with our general objective function:

$$V(a) = \sum_m p_m \phi(EU_m(a)) \tag{22}$$

where $EU_m(a)$ is the expected utility in model m when abatement is a, and m indexes the set of models $\mathcal{M}$. To prove the result, begin by considering an ordered set of functions f_λ indexed by $\lambda \in \mathbb{R}$ such that when $\lambda_2 > \lambda_1$ we have that $f_{\lambda_2} = r \circ f_{\lambda_1}$ for some increasing and concave function r, and define f_{λ_0} to be the identity function.

The first order condition for the problem (2) when ambiguity preferences are given by $f_\lambda \circ \phi$ is:

$$V'_\lambda(a) := \sum_m p_m f'_\lambda(\phi(EU_m))\phi'(EU_m)\frac{dEU_m(a)}{da} = 0, \tag{23}$$

Note that ambiguity preferences $f_\lambda \circ \phi$ are always more ambiguity averse than ϕ when $\lambda > \lambda_0$, by definition. Let a_0 be the solution of $V'_{\lambda_0}(a_0) = 0$, i.e. the solution to the optimization problem when ambiguity preferences are given by ϕ. Now notice that $\frac{d^2 EU_m}{da^2} < 0$ implies that $\frac{d^2 V}{da^2} < 0$.[23] The concavity of $V(a)$ implies that, if we can find conditions under which $V'_\lambda(a_0) > 0$ for any $\lambda > \lambda_0$, then the solution of $V'_\lambda(a_\lambda) = 0$ must satisfy $a_\lambda > a_0$. Thus our strategy will be to show that the premises of the proposition imply that $V'_{\lambda_0}(a_0) = 0 \Rightarrow V'_\lambda(a_0) \geq 0$ when $\lambda > \lambda_0$.

We will make use of two lemmas to establish the result:

Lemma 1 *Define* $K_a(m,\lambda) := f'_\lambda(\phi(EU_m(a)))\phi'(EU_m(a))$. $K_a(m,\lambda)$ *is log-supermodular (log-submodular) when the sequence* $(EU_m(a))_{m\in\mathcal{M}}$ *is decreasing (increasing) in m, for any value of a.*

Proof Consider the log-supermodular case. By definition (see e.g., Gollier 2001, p. 100), $K_a(m,\lambda)$ is log-supermodular if for any $m_H > m_L$ and $\lambda_H > \lambda_L$ we have that

$$\frac{K_a(m_H,\lambda_H)}{K_a(m_L,\lambda_H)} \geq \frac{K_a(m_H,\lambda_L)}{K_a(m_L,\lambda_L)}. \tag{24}$$

[23] To see this, differentiate (2) twice to find $V''(a) = \sum_m p_m \left[\phi''(EU_m)\left(\frac{dEU_m}{da}\right)^2 + \phi'(EU_m)\frac{d^2 EU_m}{da^2}\right]$. Thus since $\phi'' < 0$ and $\phi' > 0$, $\frac{d^2 EU_m}{da^2} < 0$ implies that $V''(a) < 0$.

Substituting the definition of $K_a(m, \lambda)$ into the above inequality we find that it holds iff

$$\frac{f'_{\lambda_H}(\phi(EU_{m_H}))\phi'(EU_{m_H})}{f'_{\lambda_L}(\phi(EU_{m_H}))\phi'(EU_{m_H})} \geq \frac{f'_{\lambda_H}(\phi(EU_{m_L}))\phi'(EU_{m_L})}{f'_{\lambda_L}(\phi(EU_{m_L}))\phi'(EU_{m_L})} \tag{25}$$

Now since $\lambda_H > \lambda_L$ there exists an increasing and concave function r such that $f_{\lambda_H} = r \circ f_{\lambda_L}$, which implies that $f'_{\lambda_H} = (r' \circ f_{\lambda_L}) f'_{\lambda_L}$. Substituting this expression for f'_{λ_H} into the above inequality and canceling common factors we find that the inequality reduces to

$$r'(f_{\lambda_L}(\phi(EU_{m_H}))) \geq r'(f_{\lambda_L}(\phi(EU_{m_L}))). \tag{26}$$

Since ϕ and f_{λ_L} are increasing, and r' is decreasing, a sufficient condition for this inequality to hold is that $EU_{m_L} \geq EU_{m_H}$. Since the inequality must hold for all $m_L < m_H$, this implies that the sequence $\{EU_m\}$ must be decreasing in m. The log-submodular case follows analogously. □

The second lemma is described in detail in Gollier (2001, p. 102):

Lemma 2 *Let $g(m)$ be a function that crosses the m-axis singly from below (i.e. $\exists m_0$ such that, $\forall m$, $(m - m_0)g(m) \geq 0$). Consider a positive function $K(m, \lambda)$. Then the following condition holds:*

$$\mathbf{E}_m g(m)K(m, \lambda_1) = 0 \Rightarrow \forall \lambda_2 > \lambda_1,\ \mathbf{E}_m g(m)K(m, \lambda_2) \geq 0, \tag{27}$$

if and only if the function $K(m, \lambda)$ is log-supermodular, where $\mathbf{E}_m$ is the expectation operator over m, which is an arbitrarily distributed random variable. The same result obtains if the function $g(m)$ crosses singly from above, and the function $K(m, \lambda)$ is log-submodular.

We now combine these two lemmas. Consider the case in which $\left(\frac{dEU_m}{da}\right)_{m\in\mathcal{M}}$ is increasing in m, and $(EU_m)_{m\in\mathcal{M}}$ is decreasing in m. Define

$$g(m; a_0) := \left.\frac{dEU_m}{da}\right|_{a=a_0}. \tag{28}$$

Evaluating the first order condition (23) at $\lambda = \lambda_0$, we have that

$$\sum_m p_m \phi'(EU_m(a_0)) \left.\frac{dEU_m}{da}\right|_{a=a_0} = 0. \tag{29}$$

Since $\left(\frac{dEU_m}{da}\right)_{m\in\mathcal{M}}$ is increasing in m, (29) makes it clear that $g(m; a_0)$ crosses the horizontal axis singly from below. This is so since $\phi' > 0$ means that in order for (29) to be satisfied, we require some of the terms in the sequence $\left(\left.\frac{dEU_m}{da}\right|_{a=a_0}\right)_{m\in\mathcal{M}}$ to be negative, and some to be positive. Since the sequence is increasing, $g(m; a_0)$ must exhibit the single-crossing property.

Now since $\{EU_m\}$ is decreasing in m by assumption, by Lemma 1 we have that $K_{a_0}(m, \lambda)$ is a log-supermodular function. By the definition of $g(m; a_0)$ and λ_0,

$$\sum_m p_m g(m; a_0) K_{a_0}(m, \lambda_0) = V'_{\lambda_0}(a_0) = 0. \tag{30}$$

Now by Lemma 2, the fact that $g(m; a_0)$ exhibits the single crossing property, and the fact that $K_{a_0}(m, \lambda)$ is log-supermodular, we know that for any $\lambda > \lambda_0$,

$$\sum_m p_m g(m; a_0) K_{a_0}(m, \lambda) = V'_\lambda(a_0) \geq 0, \tag{31}$$

establishing the result.

To complete the proof, notice that when $(EU_m)_{m\in\mathcal{M}}$ is increasing and $\left(\frac{dEU_m}{da}\Big|_{a=a_0}\right)_{m\in\mathcal{M}}$ decreasing in m, we can apply the same reasoning and use Lemma 2 in the case of a function g that crosses singly from above, and a function h that is log-submodular.

Finally, this method of proof is easily extended to the case where the sequences $(EU_m)_{m\in\mathcal{M}}$ and $\left(\frac{dEU_m}{da}\Big|_{a=a_0}\right)_{m\in\mathcal{M}}$ are comonotonic, in which case ambiguity aversion has the opposite affect on optimal abatement. To do this, one simply defines $g(m; a_0) := -\frac{dEU_m}{da}\Big|_{a=a_0}$ in the step (28). Under the assumption that both sequences are decreasing in m for all a, all the following steps of the proof go through unchanged, and we are left with the conclusion that $-V'_{\lambda_0}(a_0) = 0 \Rightarrow -V'_\lambda(a_0) \geq 0$ when $\lambda > \lambda_0$. Thus by the concavity of $V(a)$, an increase in λ *decreases* the optimal value of a in this case.

References

Ackerman F, Stanton EA, Bueno R (2010) Fat tails, exponents, and extreme uncertainty: simulating catastrophe in DICE. Ecol Econ 69(8):1657–1665

Anscombe FJ, Aumann RJ (1963) A definition of subjective probability. Ann Math Stat 34(1):199–205

Arrow KJ, Hurwicz L (eds) (1977) An optimality criterion for decision-making under ignorance. In: Studies in resource allocation processes, p 482. Cambridge University Press, Cambridge

Athanassoglou S, Xepapadeas A (2012) Pollution control with uncertain stock dynamics: when, and how, to be precautious. J Environ Econ Manag 63(3):304–320

Bassett GW, Koenker R, Kordas G (2004) Pessimistic portfolio allocation and choquet expected utility. J Financ Econom 2(4):477–492

Binmore K. (2009) Rational decisions. Princeton University Press, Princeton, NJ

Bossaerts P, Ghirardato P, Guarnaschelli S, Zame WR (2010) Ambiguity in asset markets: theory and experiment. Rev Financ Stud 23(4):1325–1359

Cai Y, Judd KL, Lontzek TS (2012) DSICE: a dynamic stochastic integrated model of climate and economy. SSRN working paper

Camerer C (1999) Ambiguity-aversion and non-additive probability: experimental evidence, models and applications. In: Luini L (ed) Uncertain decisions: bridging theory and experiments.. Kluwer, Dordrecht pp 53–80

Dasgupta P (2008) Discounting climate change. J Risk Uncertain 37(2):141–169

Dow J, da Costa Werlang SR (1992) Uncertainty aversion, risk aversion, and the optimal choice of portfolio. Econometrica 60(1):197–204

Ellsberg D (1961) Risk, ambiguity, and the Savage axioms. Q J Econ 75(4):643–669

Frame DJ, Faull NE, Joshi MM, Allen MR (2007) Probabilistic climate forecasts and inductive problems. Philos Trans R Soc A Math Phys Eng Sci 365(1857):1971–1992

Ghirardato P, Maccheroni F, Marinacci M (2004) Differentiating ambiguity and ambiguity attitude. J Econ Theory 118(2):133–173

Gilboa I (2009) Theory of decision under uncertainty, 1st edn. Cambridge University Press, Cambridge, MA

Gilboa I, Postlewaite A, Schmeidler D (2009) Is it always rational to satisfy Savage's axioms?. Econ Philos 25(3):285–296

Gilboa I, Postlewaite AW, Schmeidler D (2008) Probability and uncertainty in economic modeling. J Econ Perspect 22(3):173–188

Gilboa I, Schmeidler D (1989) Maxmin expected utility with non-unique prior. J Math Econ 18(2):141–153

Gilboa I, Schmeidler D (1995) Case-based decision theory. Q J Econ 110(3):605–639

Gollier C (2001) The economics of risk and time. MIT Press, Cambridge

Gollier C (2009) Portfolio choices and asset prices: the comparative statics of ambiguity aversion. IDEI working paper 357. http://idei.fr/display.php?a=4812

Gollier C, Gierlinger J (2008) Socially efficient discounting under ambiguity aversion. IDEI working paper 561. http://idei.fr/display.php?a=9848

Gonzalez F (2008) Precautionary principle and robustness for a stock pollutant with multiplicative risk. Environ Resour Econ 41(1):25–46

Hansen LP, Sargent TJ (2007) Robustness. Princeton University Press, Princeton, NJ

Heal G (2009) Climate economics: a meta-review and some suggestions for future research. Rev Environ Econ Policy 3(1):4–21

Henry C, Henry M (2002) Formalization and application of the precautionary principle. Columbia University Department of Economics discussion paper series. http://www.columbia.edu/cu/economics/discpapr/DP0102-22.pdf

Herstein IN, Milnor J (1953) An axiomatic approach to measurable utility. Econometrica 21(2):291–297

Hoeting JA, Madigan D, Raftery AE, Volinsky CT (1999) Bayesian model averaging: a tutorial. Stat Sci 14(4):382–401

Hope C (2006) The marginal impact of CO2 from PAGE2002: an integrated assessment model incorporating the IPCC's five reasons for concern. Integr Assess 6(1):19–56

Jouzel J, Masson-Delmotte V, Cattani O, Dreyfus G, Falourd S, Hoffmann G, Minster B, Nouet J, Barnola JM, Chappellaz J, Fischer H, Gallet JC, Johnsen S, Leuenberger M, Loulergue L, Luethi D, Oerter H, Parrenin F, Raisbeck G, Raynaud D, Schilt A, Schwander J, Selmo E, Souchez R, Spahni R, Stauffer B, Steffensen JP, Stenni B, Stocker TF, Tison JL, Werner M, Wolff EW (2007) Orbital and millennial Antarctic climate variability over the past 800,000 years. Science 317(5839):793–796

Ju N, Miao J (2012) Ambiguity, learning, and asset returns. Econometrica 80(2):559–591

Karp L, Zhang J (2006) Regulation with anticipated learning about environmental damages. J Environ Econ Manag 51(3):259–279

Keller K, Bolker BM, Bradford DF (2004) Uncertain climate thresholds and optimal economic growth. J Environ Econ Manag 48(1):723–741

Kelly DL, Kolstad CD (1999) Bayesian learning, growth, and pollution. J Econ Dyn Control 23(4):491–518

Klibanoff P, Marinacci M, Mukerji S (2005) A smooth model of decision making under ambiguity. Econometrica 73(6):1849–1892

Klibanoff P, Marinacci M, Mukerji S (2009) Recursive smooth ambiguity preferences. J Econ Theory 144(3):930–976

Knight F (1921) Risk, uncertainty, and profit. Houghton Mifflin, New York

Knutti R (2010) The end of model democracy?. Clim Change 102(3):395–404

Knutti R, Furrer R, Tebaldi C, Cermak J, Meehl GA (2010) Challenges in combining projections from multiple climate models. J Clim 23(10):2739–2758

Kreps DM, Porteus EL (1978) Temporal resolution of uncertainty and dynamic choice theory. Econometrica 46(1):185–200

Lange A, Treich N (2008) Uncertainty, learning and ambiguity in economic models on climate policy: some classical results and new directions. Clim Change 89(1):7–21

Lemoine D, Traeger CP (2012) Tipping points and ambiguity in the Economics of climate change. NBER working paper no 18230

Lenton TM, Held H, Kriegler E, Hall JW, Lucht W, Rahmstorf S, Schellnhuber HJ (2008) Tipping elements in the Earth's climate system. Proc Natl Acad Sci 105(6):1786–1793

Maccheroni F, Marinacci M, Rustichini A (2006) Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica 74(6):1447–1498

Manne A, Richels R (1992) Buying greenhouse insurance: the economic costs of CO2 emission limits. The MIT Press, Cambridge, MA

McJeon HC, Clarke L, Kyle P, Wise M, Hackbarth A, Bryant BP, Lempert RJ (2011) Technology interactions among low-carbon energy technologies: what can we learn from a large number of scenarios?. Energy Econ 33(4):619–631

Meinshausen M, Meinshausen N, Hare W, Raper SCB, Frieler K, Knutti R, Frame DJ, Allen MR (2009) Greenhouse-gas emission targets for limiting global warming to 2C. Nature 458(7242):1158–1162

Nordhaus WD (2008) A question of balance. Yale University Press, New Haven, CT

Savage LJ (1954) The foundations of statistics. Wiley, London

Schmeidler D (1989) Subjective probability and expected utility without additivity. Econometrica 57(3):571–587

Slovic P, Tversky A (1974) Who accepts Savage's axiom?. Behav Sci 19(6):368–373

Smith LA (2002) What might we learn from climate forecasts? Proc Natl Acad Sci USA 99(Suppl 1):2487–2492

Smith LA (2007) Chaos: a very short introduction, vol 159. Oxford University Press, Oxford

Stainforth DA, Allen MR, Tredger ER, Smith LA (2007) Confidence, uncertainty and decision-support relevance in climate predictions. Philos Trans R Soc A Math Phys Eng Sci 365(1857):2145–2161

Stern N (2008) The economics of climate change. Am Econ Rev 98(2):1–37

Stern NH (2007) The economics of climate change: the Stern review. Cambridge University Press, Cambridge

Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. Philos Trans R Soc A Math Phys Eng Sci 365(1857):2053–2075

Tol R (1997) On the optimal control of carbon dioxide emissions: an application of FUND. Environ Model Assess 2(3):151–163

Traeger CP (2009) Recent developments in the intertemporal modeling of uncertainty. Annu Rev Resour Econ 1(1):261–286

von Neumann J, Morgenstern O (1944) Theory of games and economic behaviour. Princeton University Press, Princeton, NJ

Weitzman ML (1976) On the welfare significance of national product in a dynamic economy. Q J Econ 90(1):156–162

Weitzman ML (2007) A review of The Stern review on the economics of climate change. J Econ Lit 45(3): 703–724

Weitzman ML (2009) On modeling and interpreting the economics of catastrophic climate change. Rev Econ Stat 91(1):1–19

Weitzman ML (2012) GHG targets as insurance against catastrophic climate damages, ghg targets as insurance against catastrophic climate damages. J Public Econ Theory 14(2):221–244