Decision Making Under Model Uncertainty: Fréchet–Wasserstein Mean Preferences

Published Version:
Petracou, E.V., Xepapadeas, A., Yannacopoulos, A.N. (2022). Decision Making Under Model Uncertainty: Fréchet–Wasserstein Mean Preferences. MANAGEMENT SCIENCE, 68(2), 1195-1211 [10.1287/mnsc.2021.3961].

Availability:
This version is available at: https://hdl.handle.net/11585/863779 since: 2024-06-05

Published:
DOI: http://doi.org/10.1287/mnsc.2021.3961

Submitted to *Management Science*
manuscript MS-DEC-19-01009-REV

Decision Making under Model Uncertainty: Fréchet-Wasserstein Mean Preferences

Electra V. Petracou
Department of Geography, University of the Aegean, ipetr@geo.aegean.gr

Anastasios Xepapadeas
Department of European and International Economic Studies, Athens University of Economics and Business, and Department of Economics, University of Bologna, xepapad@aueb.gr

Athanasios N. Yannacopoulos
Department of Statistics and Stochastic Modelling and Applications Laboratory, Athens University of Economics and Business, ayannaco@aueb.gr

This paper contributes to the literature on decision making under multiple probability models by studying a class of variational preferences. These preferences are defined in terms of Fréchet mean utility functionals which are based on the Wasserstein metric in the space of probability models. In order to produce a measure which is the "closest" to all probability models in the given set, we find the barycenter of the set. We derive explicit expressions for the Fréchet-Wasserstein mean utility functionals and show that they can be expressed in terms of an expansion which provides a tractable link between risk aversion and ambiguity aversion. The proposed utility functionals are illustrated in terms of two applications. The first application allows us to define the social discount rate under model uncertainty. In the second application the functionals are used in risk securitization. The barycenter in this case can be interpreted as the model which maximizes the probability that different decision makers will agree upon, which could be useful for designing and pricing a catastrophe bond.

Key words: ambiguity aversion, Fre'chet mean preferences, Wasserstein metric, social discount rate, model uncertainty.
JEL Classification: D81
History: Revised version of MS-DEC-19-01009

1. Introduction

The classic paradigm of von Neumann-Morgenstern expected utility (or its subjective extension due to Savage) is the primary "industry standard". Concretely, if a decision maker (DM) faces a random variable (lottery) X and the probability measure Q can describe the distribution of X, then according to the expected utility functional, the utility of X is $U(X) = \mathbb{E}_Q[u(X)]$. This is a straightforward basis for decision rules; however, it is not clear what the basis would be if there are doubts as to whether Q is the right model to describe the distribution of X – a situation usually referred to as model uncertainty. What would happen if more than one probability measure for X exists? Should the DM pick one model out of this set, or use ex-ante information to construct a subjective prior probability over models (Marinacci 2015), or do something else?

Since the fundamental contribution of Frank Knight (1921) who first put forth the difference between risk and uncertainty, the literature has paid increasing attention to the effects on decision making of the absence of a single probability model for X. Hansen and Sargent's (2001) multiplier preferences model promoted the idea of introducing a reference or benchmark model for the random variable to be evaluated, which is surrounded by a cloud of alternative models (see also Strzalecki (2011) for the axiomatic foundation of multiplier preferences). More recently, Cerreia-Vioglio et al. (2011) introduced a general class of complete and transitive preferences that are monotone and convex, which they call "uncertainty averse" preferences. These preferences include as special cases the Maccheroni-Marinacci-Rustichini variational preferences model (Maccheroni et al. 2006), the seminal Gilboa-Schmeidler minimax utility model (Gilboa and Schmeidler 1989), the smooth ambiguity model of Klibanov et al. (2005) and the multiplier and constrained preferences models of Hansen and Sargent (2001) and Hansen et al. (2006). See also Strzalecki (2011, Figure 1) for relations between classes of preferences. Throughout this paper, the term "uncertainty averse" refers to the concept as introduced by Cerreia-Vioglio et al. (2011).

The present paper contributes to the literature on decision making under multiple probability models and uncertainty aversion by studying a class of variational preferences – and hence uncertainty averse preferences in the sense of Cerreia-Vioglio et al. (2011) – which are defined in terms of the Fréchet mean in the space of probability models focusing on the use of the Wasserstein metric and the Wasserstein barycenter. In Figure 1 we present the relations between the class of variational preferences and the preferences proposed in this paper. Such preferences are useful in situations in which the DM faces multiple probability models. These situations may appear naturally in areas such as finance, where an individual investor faces multiple models for relevant economic variables, e.g., stock returns; climate change, when there are multiple probability densities regarding the equilibrium climate sensitivity (Meinshausen et al. 2009, Rogelj et al. 2012, Heal and Millner 2014) or multiple subjective probability intervals for the emergence of major changes, i.e., tipping points (Kriegler et al. 2009); or the choice of the discount rate, where multiple models may be relevant for hidden variables (Gollier 2013).

The first contribution of this paper is to combine the set of plausible probability measures in order to construct another measure – the barycenter – which could be acceptable because of the

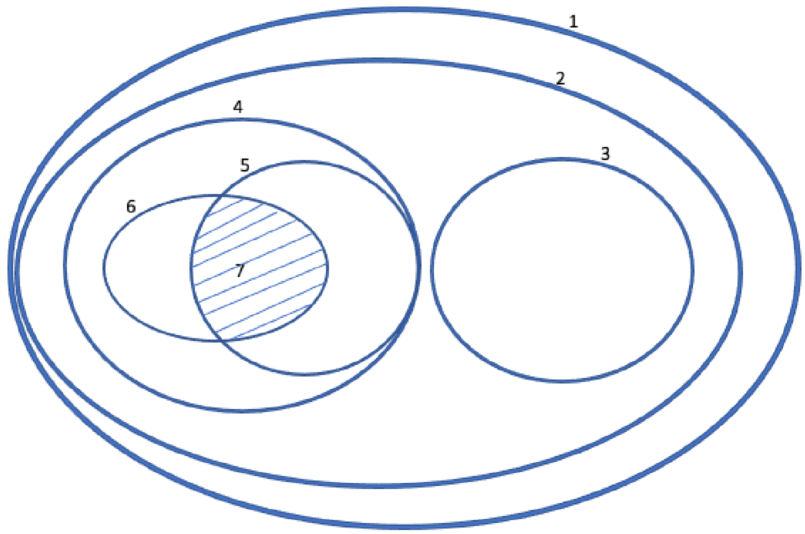

Figure 1 1: Uncertainty averse preferences (Cerreia-Vioglio et al. 2011), 2: Variational preferences (Maccheroni et al. 2006), 3: Hansen and Sargent multiplier preferences based on Kullback-Leibler divergence (Hansen and Sargent 2001), 4: Fréchet mean preferences (Definition 1a), 5: Fréchet multiplier preferences (Definition 1b), 6: Fréchet-Wasserstein mean preferences (Definition 1 and Section 2.5), 7: Fréchet-Wasserstein multiplier preferences (Definition 1 and Section 3).

desirable property that the total distance in the metric space of probability measures between the barycenter and all other probability measures is the minimum. Thus, in a sense, the barycenter is "as close as possible" to the measures provided by the experts or the beliefs of a DM or, to put it differently, it is the error minimizing choice.

Second, we introduce uncertainty aversion by augmenting the utility functional with an ambiguity penalty which is defined in terms of the Fréchet variance. This results in characterizing Fréchet mean preferences as variational preferences. Thus it provides a clear link between the use of the barycenter as a desirable probability measure when the problem is characterized by multiple probability measures and decision making under uncertainty aversion.

To address the problem of proper metrization of the set of probability measures arising in the construction of uncertainty averse utility functionals, we choose to metrize the set of probability measures using the (Monge-Kantorovich-) Wasserstein metric. For this class of uncertainty averse utility functionals, explicit representations are obtained in terms of appropriate quantile averaging of the set of models, a concept which has recently attracted attention as a model averaging tool (see, e.g., Lichtendahl et al. 2013). This is useful in providing tractable results in the attempt to answer questions related to applications.

The third contribution of the paper is to show that the Fréchet-Wasserstein mean utility functionals can be expressed in terms of an expansion with respect to an uncertainty aversion parameter.

The first term of this expansion is the barycentric expected utility, i.e., a von Neumann-Morgenstern expected utility calculated at the Wasserstein barycenter probability measure $Q_{\mathcal{B}}$, while all the higher order corrections can be interpreted as expected utility corrections calculated at the Wasserstein barycenter $Q_{\mathcal{B}}$. These corrections correspond to utility functions characterized by a higher risk aversion coefficient relative to the von Neumann expected utility of the first term of the expansion. This result connects the concepts of uncertainty aversion and risk aversion. Finally, we show that if many DMs with different models seek to reach consensus, the barycenter can be interpreted as the model which maximizes the probability of achieving consensus.

In terms of applications, we first use this version of variational preferences to calculate the social discount rate and show the differences in the discount rate estimates between risk and uncertainty aversion. Second, we use the interpretation of the barycenter as the model which maximizes the probability of consensus to suggest that the barycenter could be the consensus model which different agents can agree upon when considering the securitization of an extreme risk, such as when issuing a CAT bond. The proofs of the main results can be found in the Appendix at the end of the paper. All other proofs and technical details are available in the online Appendix.

2. Fréchet and Fréchet-Wasserstein mean preferences

2.1. Motivation

Consider a DM who acts as a planner and who asks a set of experts to provide a model describing a random variable X. The DM wishes to evaluate X, a task which requires determining the probability distribution of X. The experts are not unanimous concerning the distribution of X, hence, the DM faces n possible probabilistic models (priors) Q_i, with the set $\mathbb{M} := \{Q_i,\ i = 1, \cdots, n\}$ regarded as a reference set. The DM is not confident of the absolute validity of any of them, which constitutes the case of model uncertainty.

The discrepancy between the various probability distributions (models) Q_i can be quantified by considering them as elements of $\mathcal{P}$, the set of probability measures on some appropriate sample space, endowed with some metric d, chosen so that $d(Q_i, Q_j)$ provides an appropriate notion of distance between the probability measures $Q_i, Q_j \in \mathcal{P}$. Then, one way to choose a model Q^* using the set $\mathbb{M}$ is to consider the models in $\mathbb{M}$ as incomplete observations of the true model Q and try to interpolate between them, in analogy with a least squares fit in standard regression analysis. The resulting probability measure, which will be called the barycenter of the set of models $\mathbb{M}$, can then be considered as the average (or mean) model, in the sense that it provides the best estimate for an acceptable model compatible with $\mathbb{M}$.

2.2. Preliminaries and standing notation

Uncertainty here is modeled by a measurable space $(\Omega, \mathcal{B})$, where Ω is a set of possible states of nature and $\mathcal{B}$ is a σ-algebra on Ω containing more complex events. The DM faces a payoff (or loss) X, whose exact value depends on the exact state of nature $\omega \in \Omega$ that will be realized and is modeled as a random variable $X : \Omega \to \mathbb{R}$, on $(\Omega, \mathcal{B})$. However, each expert provides a different probability measure for the events on Ω, which in turn leads to a different probability measure on $\mathbb{R}$ that can be used to derive a distribution for X, leading to a set of feasible models $\mathbb{M} = \{Q_i,\ i = 1, \cdots, n\}$. This allows us to focus, if needed, on probability measures on $\mathbb{R}$.

In what follows we will use the notation $Q_i \in \mathcal{P}$ or $Q_i \in \mathcal{P}(\mathbb{R})$ to distinguish between the general case and the case where we restrict attention to probability measures on $\mathbb{R}$, respectively. Restricting attention to $\mathcal{P}(\mathbb{R})$ is convenient since all elements $Q \in \mathcal{P}(\mathbb{R})$ can be represented in terms of distribution functions $F : \mathbb{R} \to [0,1]$, defined by $F(x) = Q((-\infty, x])$ for every $x \in \mathbb{R}$, or quantile functions $F^{-1} : [0,1] \to \mathbb{R}$. The space of quantile functions will be denoted by $\mathcal{Q}$, and the set of strictly positive quantile functions by $\mathcal{Q}_{++}$. For most of this work we will restrict our attention to the case $\mathcal{P}(\mathbb{R})$.

The space of probability measures $\mathcal{P}$ will be endowed with a suitable metric d. The weighted average of the squared distances between $Q \in \mathcal{P}$, and all probability models Q_i contained in the reference set $\mathbb{M} \subset \mathcal{P}$, also known as the Fréchet mean, is given by

$$\mathsf{F}_{\mathbb{M}}(Q) = \sum_{i=1}^{n} w_i d^2(Q, Q_i),\ \ w = (w_1, \cdots, w_n) \in \Delta^{n-1},$$

where $\Delta^{n-1} := \{w = (w_1, \cdots, w_n)\ : w_i \in [0,1], \sum_{i=1}^{n} w_i = 1\}$ is the $(n-1)$-dimensional simplex.

The barycenter, or the Fréchet mean denoted by $Q_{\mathcal{B}}$, is the probability model which provides the minimum of the Fréchet function. The quantity $\mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}})$ is real valued and positive, and can be interpreted as the minimum square error induced if we approximate the models in $\mathbb{M}$ with $Q_{\mathcal{B}}$. The Fréchet variance of $\mathbb{M}$ is then defined as $\mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}})$. The larger the value of $\mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}})$, the larger the discrepancies in the models within the reference set of models $\mathbb{M}$. The weights w_i could provide a credibility weighting on each model, with the more credible ones – such as those supplied by the most authoritative expert or the most reliable measuring device – being assigned higher weight than the others. The choice $w_i = \frac{1}{n}$, $i = 1, \cdots, n$, corresponds to assigning equal credibility to all models in $\mathbb{M}$.

In this paper, from Section 2.5 onwards, we commit to the 2-Wasserstein metric (also denoted by W_2) and set $d = W_2$. For general definitions and properties of the Wasserstein metric see Section

EC.1 of the online Appendix. For probability measures $Q_i \in \mathcal{P}(\mathbb{R})$, this metric admits a convenient representation in terms of the quantile functions F_i^{-1} of the measures Q_i (where F_i is the distribution function) as

$$d(Q_1, Q_2) := \left(\int_0^1 (F_1^{-1}(s) - F_2^{-1}(s))^2 ds \right)^{1/2} = \|F_1^{-1} - F_2^{-1}\|_{L^2([0,1])}. \tag{1}$$

Then, for a set of models $\mathbb{M} = \{Q_1, \cdots, Q_n\} \subset \mathcal{P}(\mathbb{R})$, the Wasserstein barycenter can be obtained in terms of the corresponding quantiles as the weighted quantile average

$$F_{\mathcal{B}}^{-1} = \sum_{i=1}^{n} w_i F_i^{-1}.$$

We use this metric because:

(a) Unlike other popular distances such as the Kullback-Leibler divergence - which is not symmetric or may not satisfy the triangle inequality - it is a true metric, which is compatible with the weak* topology in the space of measures. Indeed, convergence with respect to the 2-Wasserstein metric is equivalent to weak convergence of measures plus convergence of the first two moments (see, e.g., Villani 2008, or Santambrogio 2015).

(b) It allows us to extend one of the most desirable properties of the Kullback-Leibler divergence - that of reducing robust decision problems within the exponential family of distributions to quadratic optimization problems - to any family. Furthermore, it allows the calculation in closed form of the resulting variational utility (see Theorem 1) and elucidates the connection between risk aversion and uncertainty aversion in terms of concrete perturbative expansions (see Theorem 2).

(c) It can be used to establish an upper bound for the difference in expected utility associated with using different probability measures for determining the expected utility associated with a random variable. By adapting a classical result (see e.g. Villani 2008 or Santambrogio 2015), it can be shown (see for example Section EC.2.1 in the online Appendix) that the error in the estimation of a random variable X, using expected utility under two different probability measures, e.g. $Q_i, Q_{\mathcal{B}}$, for the distribution of X, is bounded above as

$$|\mathbb{E}_{Q_i}(u(X)) - \mathbb{E}_{Q_{\mathcal{B}}}(u(X))| \leq C d(Q_i, Q_{\mathcal{B}}), \tag{2}$$

provided that the utility function is Lipschitz continuous, where C is the Lipschitz constant for u.

2.3. Barycentric expected utility

Once the barycenter $Q_{\mathcal{B}}$ for the set of models $\mathbb{M} \subset \mathcal{P}$ is derived, the utility associated with the random variable X could be defined, under standard risk aversion, in terms of a von Neumann-Morgenstern-Savage expected utility functional of the form

$$\mathcal{U}_{\mathcal{B}}(X) = \mathbb{E}_{Q_{\mathcal{B}}}[u(X)],$$

which we will call the barycentric expected utility. Note that even though this is an expected utility, it nevertheless acounts for the effect of model uncertainty since the barycenter $Q_{\mathcal{B}}$ is an average or composite model for the set of models $\mathbb{M}$. This is compatible with subjective expected utility if one assumes that Q_i are the opinions of different experts and $Q_{\mathcal{B}}$ is one's way to come up with one's subjective probability model (see also Proposition 3).

The next result provides an explicit form for the barycentric expected utility when we restrict our attention to $\mathcal{P}(\mathbb{R})$, i.e., to probability measures on $\mathbb{R}$ metrized by the 2-Wasserstein metric.

PROPOSITION 1. *Consider the set of models $\mathbb{M} = \{Q_1, \cdots, Q_n\} \subset \mathcal{P}(\mathbb{R})$, with each probability measure Q_i represented by the quantile function F_i^{-1}. For any choice of weights $w = (w_1, \cdots, w_n) \in \Delta^{n-1}$, the barycentric expected utility of a lottery $X : \Omega \to \mathbb{R}$ is given by*

$$\mathbb{E}_{Q_{\mathcal{B}}}[u(X)] = \int_0^1 u\left(\sum_{i=1}^n w_i F_i^{-1}(s)\right) ds.$$

Proof: The proof follows easily by the representation of the Wasserstein barycenter in terms of the quantile average and the change of variables $s = F_{\mathcal{B}}(x)$ in the resulting expectation. □.

2.4. Fréchet mean preferences

Barycentric expected utility does not take into account aversion to model uncertainty. To address such concerns we introduce the uncertainty aversion axiom, which within the framework of variational preferences is reflected by a penalty which is based on the Fréchet function and variance.

DEFINITION 1 (FRÉCHET MEAN UTILITY FUNCTIONALS). Consider a set of models $\mathbb{M} \subset \mathcal{P}$, the corresponding Fréchet function $\mathsf{F}_{\mathbb{M}}$ and barycenter $Q_{\mathcal{B}}$.

(a) For any real valued increasing convex function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$, the family of utility functionals

$$\mathcal{U}(X) := \min_{Q \in \mathcal{P}} \left(\mathbb{E}_Q[u(X)] + \phi(\mathsf{F}_{\mathbb{M}}(Q)) \right), \tag{3}$$

is called the family of Fréchet mean utilities.

(b) The family of utility functionals

$$\mathcal{U}_\theta(X) = \min_{Q \in \mathcal{P}} \left(\mathbb{E}_Q[u(X)] + \frac{\theta}{2} \left(\mathsf{F}_{\mathbb{M}}(Q) - \mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}}) \right) \right), \ \theta > 0, \tag{4}$$

are called Fréchet multiplier preferences. $\theta > 0$ is a parameter quantifying uncertainty aversion, called the multiplier. The constant $\mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}})$ (Fréchet variance) and the factor $\frac{1}{2}$ are included for normalization purposes.

The family of Fréchet mean utilities is well-posed for appropriate choice of the function ϕ. This can be shown by standard arguments using the direct method of the calculus of variations (see e.g. Kravvaritis and Yannacopoulos 2020). The function $Q \mapsto \phi(\mathsf{F}_{\mathbb{M}}(Q))$ plays the role of a penalty function in the space of models, which penalizes certain members of the models in $\mathbb{M}$ that are far apart from $Q_{\mathcal{B}}$ in our metric (i.e., the outliers). The minimization over the set of probability measures activates the penalty function and has as a consequence the selection of a probability measure on which the penalty function achieves moderate values, and X is evaluated on this in terms of an expected utility contribution plus a correction term depending on the penalty.

In the preferences defined in (3), the penalty function penalizes large dispersion of models in the sense of the Fréchet variance. This is a plausible choice for uncertainty averse agents, since large Fréchet variance in the set of models indicates to the agent that the situation at hand cannot be well modeled, potentially leading to poor results in the decision-making process. If the penalty function is sufficiently large, then the minimizer Q^* for the variational problem (3) will be closer to Q_B, where the minimum of $\mathsf{F}_{\mathbb{M}}$ is achieved. This implies that (3) will be similar to the barycentric expected utility $\mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$. For moderate contributions of the penalty function, Q^* will be located at some distance from $Q_{\mathcal{B}}$, leading to deviations of $\mathcal{U}(X)$ from $\mathbb{E}_{\mathcal{B}}[u(X)]$, hence one may express the utility functional $\mathcal{U}(X) = \mathbb{E}_{Q_{\mathcal{B}}}[u(X)] + \mathfrak{C}(X)$ where $\mathfrak{C}(X)$ is a correction term corresponding to model uncertainty. Moreover, since by definition $\mathcal{U}(X) \leq \mathbb{E}_{Q}[u(X)] + \phi(\mathsf{F}_{\mathbb{M}}(Q))$ for any $Q \in \mathcal{P}$, if ϕ is normalized such that $\phi(Q_{\mathcal{B}}) = 0$, then $\mathcal{U}(X) \leq \mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$. This can be considered as some form of pessimism.

The above arguments become very clear in the case of Fréchet multiplier preferences (see Definition 1). Note that $(\mathsf{F}_{\mathbb{M}}(Q) - \mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}})) \geq 0$ for all $Q \in \mathcal{P}(\Omega)$ and achieves the value 0 when $Q = Q_{\mathcal{B}}$. If θ is large ($\theta \to \infty$), then the dominant term in the variational problem (4) is the penalty term, hence the solution of the minimization problem will be achieved where the penalty term is minimized (i.e. at $Q_{\mathcal{B}}$), and the resulting value of the utility functional will be close to $\mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$. Hence in the limit as $\theta \to \infty$ the Fréchet multiplier preferences converge to the barycentric expected utility. If θ takes smaller values, the solution of the minimization problem in (4) moves from $Q_{\mathcal{B}}$ and the resulting utility functional will be the barycentric expected utility plus corrections. As stated above, $\mathcal{U}_\theta(X) \leq \mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$, for all $\theta > 0$. This pessimism effect can be intuitively understood as $\theta \to 0$, by the fact that the weak contribution of the penalty function allows the DM to adopt probability models which put high probability on the worst outcome. This intuitive approach is made rigorous in Theorem 2.

Remark 1. Under certain conditions on the function ϕ (e.g., convexity and lower semicontinuity for $Q \mapsto \phi(\mathsf{F}_{\mathbb{M}}(Q)))$, the corresponding Fréchet mean preferences belong to the class of variational preferences defined in Maccheroni et al. (2006) and satisfy axioms A.1-A.6 in the online Appendix EC.3 and most notably the weak certainty independence axiom A.2. Such preferences display ambiguity aversion effects with $Q \mapsto \phi(\mathsf{F}_{\mathbb{M}}(Q))$ acting as an ambiguity index. Details on this can be found in Maccheroni et al. (2006), Cerreia-Vioglio et al. (2011), and Cerreia-Vioglio et al. (2015) (see online Appendix EC.3 for a formal axiomatic definition).

2.5. Fréchet-Wasserstein mean utilities

From this section onwards we restrict our attention to probability measures on $\mathbb{R}$ and assume that $\mathcal{P}(\mathbb{R})$ is metrized in terms of the 2-Wasserstein metric. In this case the utility functionals introduced in Definition 1 will be called Fréchet-Wasserstein mean utilities (case (a)) or Fréchet-Wasserstein multiplier utilities (case (b)).

In what follows, we show that the Fréchet-Wasserstein mean utilities can be determined, in almost closed form, subject to the solution of two simple one-dimensional algebraic equations, involving the quantile functions that uniquely characterize the probability measures.

THEOREM 1. *Use the standing notation in Section 2.2, and let $F_{\mathcal{B}}^{-1} = \sum_{i=1}^n w_i F_i^{-1}$ be the quantile corresponding to the Wasserstein barycenter.*

Assume the existence of an open interval $I \subset \mathbb{R}_+$ such that for every $(s,\rho) \in [0,1] \times I$ the algebraic equation

$$\frac{1}{\rho} u'(z) + z = F_{\mathcal{B}}^{-1}(s), \;\; s \in [0,1], \;\; \rho > 0, \tag{5}$$

admits a solution z, denoted by $z =: g_\rho(s)$, such that the resulting function $s \mapsto g_\rho(s)$ satisfies $g_\rho \in \mathcal{Q}$, hence can be considered as the quantile F_ρ^{-1} of some distribution function F_ρ.

Define the function $\rho \mapsto S(\rho)$ by $S(\rho) = 2\phi'(\int_0^1 (F_\rho^{-1}(s) - F_{\mathcal{B}}^{-1}(s))^2 ds)$, and assume that the algebraic equation $\rho = S(\rho)$ admits a solution $\rho^* \in I$.

Then, under suitable smoothness and integrability conditions on u and ϕ, with ϕ convex and sufficiently steep (see the online Appendix, Section EC.2.2, Remarks EC.1, EC.2), a minimizer of (3) corresponds to the quantile $F_*^{-1} = F_{\rho^*}^{-1}$ and the utility functional $\mathcal{U}(X)$ can be represented in the form

$$\mathcal{U}(X) = \int_0^1 u(F_*^{-1}(s)) ds + \phi\left(\sum_{i=1}^n w_i \int_0^1 (F_*^{-1}(s) - F_i^{-1}(s))^2 ds\right). \tag{6}$$

Proof: See Appendix A. □

REMARK 2. In the case of Fréchet-Wasserstein multiplier preferences, the function $S(\rho) = \theta$ for all $\rho \in I$, hence Theorem 1 simplifies in that $\rho^* = \theta$, and we only need to solve one algebraic equation, namely (5) for this single value of $\rho = \theta$ (see also the online Appendix, Section EC.2.2, Remark EC.2). Moreover, in this case we can obtain a perturbative expansion of F_θ^{-1} as well as $\mathcal{U}_\theta$ in terms of the uncertainty aversion parameter θ, which elucidates the connection between risk and uncertainty (see Section 3).

Remark 3. If we cannot find solutions of (5), then the solution of the variational problem leading to the definition of the utility functional $\mathcal{U}$ is still feasible in terms of non-interior solutions. This issue is technical and is discussed in the online Appendix (see Section EC.2.3).

It is interesting to note that the quantile F_*^{-1} for the probability measure for which the minimum in (3) is achieved, which is determined by condition (5), depends on: (a) the utility function u adopted by the DM, (b) the penalty function ϕ (or the uncertainty aversion parameter θ), and (c) the Wasserstein barycenter $F_{\mathcal{B}}^{-1}$ corresponding to the set of models $\mathbb{M}$.

In fact, F_*^{-1} can be considered as a nonlinear transformation of $F_{\mathcal{B}}^{-1}$; for example, the solution of (5) can expressed as $F_*^{-1} = \Psi_{\rho^*}^{-1}(\rho^* F_{\mathcal{B}}^{-1})$ where Ψ_ρ^{-1} is the inverse of the function $x \mapsto \Psi_\rho(x) = u'(x) + \rho x$, $x \in \mathbb{R}$, $\rho > 0$. This is a nonlinear quantile distortion effect, since u and ϕ are (strictly) increasing functions $\rho^* \geq 0$, and therefore $F_*^{-1}(s) = F_{\mathcal{B}}^{-1}(s) - \frac{1}{\rho^*}u'(F_*^{-1}(s)) \leq F_{\mathcal{B}}^{-1}(s)$ for any $s \in [0,1]$. The fact that $F_*^{-1} \leq F_{\mathcal{B}}^{-1}$ has an interesting economic interpretation, meaning that F_*^{-1} underestimates X at all confidence levels relative to the Wasserstein barycenter. This effect can be understood as a precautionary effect, an effect which is commonly observed in models of choice under uncertainty. Moreover, the dependence of F_*^{-1} on the penalty function as well as the utility function allows us to derive quantitative connections between risk aversion, which is related to the form of the utility function u adopted by the DM, and uncertainty aversion, which is related to the choice of penalty function ϕ. Moreover, if the penalty function ϕ is scaled so that $\phi(\mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}})) = 0$, then $\mathcal{U}(X) \leq \mathcal{U}_{\mathcal{B}}(X) := \mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$.

Condition (5) can be interpreted as a first-order condition for the variational problem defining the utility functional. Its solution provides the quantile function $F_*^{-1} := F_{\rho^*}^{-1}$ corresponding to the probability measure for which the minimum in (3) is achieved. The first-order condition can have an intuitive interpretation as follows: Consider any quantile function F^{-1} for X and any significance level s. Then, $F^{-1}(s)$ can be interpreted as an estimate for the value of X at significance level s. In this respect, the quantile function F_*^{-1} determined by condition (5) is the one for which the marginal utility matches the marginal penalty of diverting from the barycenter $F_{\mathcal{B}}^{-1}$.

3. Fréchet-Wasserstein multiplier preferences, effective risk aversion and marginal utility

We now consider the class of Fréchet-Wasserstein multiplier preferences, i.e., the choice $\phi(x) = \frac{\theta}{2}(x - \mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}}))$, where $\mathsf{F}_{\mathbb{M}}$ is the Fréchet function for $\mathcal{P}(\mathbb{R})$ metrized by the 2-Wasserstein metric. Without loss of generality we restrict our attention to random variables X achieving positive values and to utility functions u satisfying the following assumption.

Assumption 1. *The utility function $u : \mathbb{R}_+ \to \mathbb{R}$ is C^2 and satisfies the standard Inada conditions $u'(0) = \lim_{x\to 0} u'(x) = +\infty$ and $u'(\infty) = \lim_{x\to\infty} u'(x) = 0$, with $|u''|$ decreasing.*

The above assumption on u requires the introduction of a further assumption on the set of models $\mathbb{M} \subset \mathcal{P}(\mathbb{R})$, which since the uncertain quantity of interest is assumed to be positive, consists of quantiles $F_i^{-1} : [0,1] \to \mathbb{R}_+$.

Assumption 2. *For every $Q_i \in \mathbb{M}$, $Q_j(X \in [0, K_j]) = 0$ for some $K_j > 0$.*

This assumption is not too restrictive as X is not necessarily a relative gain or loss, but the outcome in some broad sense.

If Assumption 2 holds, then the Wasserstein barycenter quantile also satisfies $\inf_{s\in[0,1]} F_{\mathcal{B}}^{-1}(s) > 0$. Clearly, if at least one of the models in $\mathbb{M}$ satisfies the optimism condition of Assumption 2, then the corresponding Wasserstein barycenter also shares this property. However, as all the models in $\mathbb{M}$ satisfy Assumption 2, it is reasonable to restrict our attention to minimizers of problem (3) for multiplier preferences, to quantile functions which are strictly bounded below by a positive constant (say the minimum of the constants K_j in Assumption 2). In the next section we will show that such a solution exists and provide a characterization for it.

3.1. An approximate closed form expression for Fréchet-Wasserstein multiplier preferences and effective risk aversion

In what follows we use the standing notation in Section 2.2 and the framework and further notation of Theorem 1.

Theorem 2. *Under Assumptions 1 and 2, there exists a positive constant θ_c depending on w_i and K_i (the exact dependence is given in the proof, in Section EC.2.3), such that for $\theta > \theta_c$:*

(a) A minimizer Q_θ of problem (4) in $\mathcal{P}(\mathbb{R})$, is expressed in terms of the quantile $F_\theta^{-1} \in \mathcal{Q}_{++}$, which is the solution of (5) for $\rho = \theta$, and

$$\mathcal{U}_\theta(X) = \int_0^1 u(F_\theta^{-1}(s))ds + \frac{\theta}{2}\int_0^1 (F_\theta^{-1}(s) - F_{\mathcal{B}}^{-1}(s))^2 ds = \int_0^1 u(F_\theta^{-1}(s))ds + \frac{1}{2\theta}\int_0^1 (u'(F_\theta^{-1}(s)))^2 ds. \tag{7}$$

It holds that $F_\theta^{-1} \le F_{\mathcal{B}}^{-1}$ and $\mathcal{U}_\theta(X) \le \mathcal{U}_{\mathcal{B}}(X) := \mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$ for all $\theta > \theta_c$, i.e., the DM underestimates X at all confidence levels as compared to the Wasserstein barycenter, whereas $F_\theta^{-1} \to F_{\mathcal{B}}^{-1}$ and $\mathcal{U}_\theta(X) \to \mathcal{U}_\beta(X) = \mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$ as $\theta \to \infty$. If u' is a decreasing function we get the sharper estimate $F_\theta^{-1} \le F_{\mathcal{B}}^{-1} - \frac{1}{\theta}u'(F_{\mathcal{B}}^{-1})$ for all $\theta > \theta_c$.

(b) Assuming that $u \in C^3$, the following perturbative expansion for sufficiently large θ holds:

$$\begin{aligned}
&F_\theta^{-1} = F_{\mathcal{B}}^{-1} - \frac{1}{\theta}u'(F_{\mathcal{B}}^{-1}) + \frac{1}{\theta^2}u''(F_{\mathcal{B}}^{-1})u'(F_{\mathcal{B}}^{-1}) + O\left(\frac{1}{\theta^3}\right),\\
&\mathcal{U}_\theta(X) = \mathbb{E}_{Q_{\mathcal{B}}}[V_\theta(X)],\\
&V_\theta(x) := u(x)\left(1 - \frac{1}{2\theta}\frac{(u'(x))^2}{u(x)} + \frac{1}{2\theta^2}\frac{(u'(x))^2 u''(x)}{u(x)}\right) + O\left(\frac{1}{\theta^3}\right).
\end{aligned} \tag{8}$$

Proof: See Appendix B. □

The expansion in (8) looks like an expected utility representation, but note that the "utility function" V_θ is not in general an increasing function of x unless θ is large enough – a fact which may be attributed to the observation that (8) arises as a perturbative expansion of a utility functional which presents uncertainty aversion effects. Assuming higher differentiability for u, we can define an absolute and a relative risk aversion coefficient for V_θ, by $\rho_a(x) = -\frac{V_\theta''(x)}{V_\theta'(x)}$ and $\rho_r(x) = -x\frac{V_\theta''(x)}{V_\theta'(x)}$ respectively, and show that V_θ has the correct behaviour for a utility function as long as $u''' > 0$ and $u'''' < 0$. Substituting the exact expression for V_θ, we can obtain a perturbative expansion for ρ_a and ρ_r in powers of $\frac{1}{\theta}$, valid for sufficiently large θ, which provides information on the way in which the effective risk aversion coefficient for V_θ is affected by the uncertainty aversion parameter θ.

While this calculation is feasible (yet tedious) in terms of u and its derivatives, it is possibly more informative to provide the exact form of the "effective utility function" V_θ for the case of CRRA utilities of the form

$$u_\gamma(x) = \begin{cases} \ln x, & \text{for } \gamma = 1, \\ \frac{x^{1-\gamma}}{1-\gamma}, & \text{for } \gamma > 1. \end{cases}$$

In this case, after some elementary calculations, we obtain - upon defining $\gamma_1 = 1 + 2\gamma$, $\gamma_2 = 2 + 3\gamma$ - that

$$V_\theta(x) = u_\gamma(x) - \frac{1-\gamma_1}{2\theta} u_{\gamma_1}(x) - \frac{\gamma(1-\gamma_2)}{2\theta^2} u_{\gamma_2}(x) + O\left(\frac{1}{\theta^3}\right), \tag{9}$$

with the corresponding uncertainty averse utility functional expressed up to second order in the expansion as $\mathcal{U}_\theta(X) = \mathbb{E}_{Q_\mathcal{B}}[V_\theta(X)]$. The leading order in the expansion for $\mathcal{U}_\theta(X)$ is the barycentric expected utility, while the corrections can be interpreted again as barycentric expected utilities albeit corresponding to other members of the CRRA family, but importantly with larger risk aversion coefficients. This observation connects uncertainty aversion with risk aversion in a quantitative manner, and indicates that an uncertainty averse agent will act as a risk averse agent with a higher risk aversion coefficient. It is also clear that $\mathcal{U}_\theta(X) \leq E_{Q_\mathcal{B}}[u(X)]$, thus effective risk aversion leads to a utility functional for X reduced as compared to the barycentric expected utility. Moreover, the effects of ambiguity aversion become weaker for high wealth levels.

A similar result can be obtained for CARA utility functions[1] $u(x) = \frac{1}{\lambda}(1 - e^{-\lambda x})$, as

$$V_\theta(x) = \frac{1}{\lambda}(1 - e^{-\lambda x}) - \frac{\lambda}{2\theta} e^{-2\lambda x} - \frac{\lambda^2}{2\theta^2} e^{-3\lambda x} + O\left(\frac{1}{\theta^3}\right). \tag{10}$$

It is interesting to note that the expansion in (9) in the case where $\gamma = 1$ corresponds to a family of utility functions proposed by Bell and Fishburn (2001, p. 604) while the expansion in (10)

[1] Such functions do not satisfy $\lim_{x\to 0} u'(x) = \infty$ but this is not essential for our arguments, see Remark EC.5.

corresponds to the family of sumex utility functions (see Bell 1988, 1995). These utility functions for the leading order correction satisfy either the strong one switch or the one-switch property (Bell 1988, Bell and Fishburn 2001). The utility function expansion in (10) to higher order than the first satisfies the n-switch property (Bell 1988).[2]

3.2. Marginal utility

Since in some of the potential applications of the Fréchet-Wasserstein multiplier preferences it will be necessary to explicitly use marginal utility, in this subsection a convenient representation for this quantity is provided. In this context marginal utility is the change in the expected utility of an individual from a small deterministic change in his/her initial endowment when the individual is uncertainty averse.

Definition 2 (Marginal ambiguity-averse utility). Assume that an agent considers the random endowment X and let $\mathbb{M}$ be a relevant set of models concerning the distribution of X. Let $\epsilon > 0$ be a non-random infinitesimal endowment and let $\mathcal{U}_\theta(X)$, $\mathcal{U}_\theta(X+\epsilon)$ be the Fréchet-Wasserstein multiplier utilities corresponding to X and $X+\epsilon$ respectively as in Definition 1. The marginal utility of X is defined as

$$\mathcal{M}_\theta(X) := \lim_{\epsilon \to 0} \frac{1}{\epsilon}(\mathcal{U}_\theta(X+\epsilon) - \mathcal{U}_\theta(X)),$$

provided that the limit exists.

Proposition 2, which uses the standing notation in Section 2.2, and the framework and further notation of Theorems 1 and 2, provides the representation of the marginal utility under Fréchet-Wasserstein multiplier preferences.

Proposition 2. *Under the same assumptions as in Theorem 2, for $\theta > \theta_c$ it holds that*

(a) The marginal utility $\mathcal{M}_\theta$ is represented as

$$\mathcal{M}_\theta(X) = \int_0^1 u'(F_\theta^{-1}(s))ds = \mathbb{E}_{Q_\theta}[u'(X)]. \tag{11}$$

Morever, it holds that

$$\mathcal{M}_\theta(X) \geq \mathcal{M}_\mathcal{B}(X) = \mathbb{E}_{Q_\mathcal{B}}[u'(X)], \ \forall \theta > \theta_c,$$

while $\mathcal{M}_\theta(X) \to \mathcal{M}_\mathcal{B}(X)$ for $\theta \to \infty$.

(b) For large θ, and assuming sufficient smoothness for u, marginal utility admits the expansion

$$\mathcal{M}_\theta(X) = \mathbb{E}_{Q_\mathcal{B}}[u'(X)\mathcal{C}_\theta(X)], \tag{12}$$

where $C_\theta(X)$ is a correction factor of the form

$$\mathcal{C}_\theta(X) = 1 - \frac{1}{\theta}u''(X) + \frac{1}{\theta^2}\left[(u''(X))^2 + \frac{1}{2}u'''(X)u'(X)\right] + O\left(\frac{1}{\theta^3}\right). \tag{13}$$

[2] The authors wish to thank the Department Editor Manel Baucells for bringing this fact to their attention.

Proof: See Appendix C. □

Remark 4. The statement of Proposition 2(b) admits a very intuitive interpretation. Recall that by Theorem 2(b) we know that $F_\theta^{-1} \le F_{\mathcal{B}}^{-1}$ for all $\theta > 0$, which means that the uncertainty averse agent using the probability model with quantile function F_θ^{-1} provides statistical estimates for X, which at all confidence levels are lower than the corresponding estimates provided by the model related to the Wasserstein barycenter of $\mathbb{M}$. However, by Proposition 2(a), the marginal utility of the uncertainty averse agent admits a representation under the model with quantile function F_θ^{-1}. Since F_θ^{-1} underestimates $F_{\mathcal{B}}^{-1}$, the marginal utility corresponding to model F_θ^{-1} will be higher than the marginal utility corresponding to model $F_{\mathcal{B}}^{-1}$ (since u' is a decreasing function so that $u'(F_\theta^{-1}) \ge u'(F_{\mathcal{B}}^{-1})$). Note furthermore that while $\mathcal{M}(X) \ge \mathcal{M}_{\mathcal{B}}(X)$ for all $\theta > 0$, this does not necessarily hold for the approximations of $\mathcal{M}_\theta$ in terms of expansion (12), since in general we do not know the sign of the term u'''. However, for families such as the exponential or the CRRA family, $u''' > 0$, so the general property $\mathcal{M}(X) \ge \mathcal{M}_{\mathcal{B}}(X)$ which is valid for all θ is satisfied by the expansion as well. An example of marginal utility in the CRRA family is presented in the online Appendix, Section EC.2.4.

4. Applications

Fréchet mean preferences can be used as a tool for decision making in various applications. Their use will be more important in cases where multiple plausible models are available for the phenomenon under study, such as determining the social discount rate (Section 4.1), or in consensus decision making when multiple agents, each with a different viewpoint or model, must reach a commonly acceptable decision such as for the design and pricing of catastrophe bonds (Section 4.2).

4.1. The social discount rate under Fréchet-Wasserstein multiplier preferences

The social discount rate (SDR) is one of the most fundamental but also controversial parameters in cost-benefit analysis. In the area of climate change, for example, the discussion between Stern and Nordhaus (Stern 2007, Nordhaus 2007)[3] about the choice of the discount rate revealed the importance of this parameter for the design of climate policy. Given the uncertainty associated with valuation problems in which the SDR is used to discount future cost and benefits, the approach developed in this paper could be useful when the SDR is used under conditions of deep uncertainty.

4.1.1. Uncertainty aversion, Fréchet-Wasserstein multiplier preferences and the Ramsey discounting formula In the absence of model uncertainty the SDR is determined by the classical consumption-based Ramsey discounting formula

$$r(t) = \delta - \frac{1}{t} \ln \frac{\mathbb{E}[u'(C(t))]}{u'(C(0))},$$

[3] For the choice of the SDR, see also for example Gollier (2002), Weitzman (2007), Dasgupta (2008) and Heal (2009).

where δ is the utility discount rate, $C(t)$ is consumption at time t (which is a random variable) and $C(0)$ is today's consumption. Intuitively, if expected marginal utility in the future is higher, then the future is discounted less. This formula provides a term structure for r and is a crucial parameter in standard cost-benefit analysis (Gollier 2013).

Assume now that a regulator seeks to calculate the SDR in the presence of model uncertainty. The regulator has a standard CRRA utility function and has its own estimates of the utility discount rate δ and the elasticity of marginal utility or, equivalently, the coefficient of relative risk aversion γ. The regulator, however, does not have a model for the evolution of uncertain future consumption and thus asks experts. The experts provide different models, i.e., probability measures, for the stochastic future consumption, but the regulator is not able to consider a specific model as the most credible or as the benchmark model. The Fréchet mean preferences developed in this paper, and in particular the Fréchet-Wasserstein multiplier preferences, can be used to provide a satisfactory approach to estimating the SDR.

To bring this problem into our general framework, we assume that for any fixed $t > 0$, the random variable $X = C(t)$ is the unknown consumption at this instant in time, and that there is a set of models $\mathbb{M}_t$ of probability measures concerning the distribution of $C(t)$. This set of models is described in terms of the quantiles $F_{t,i}^{-1}$, $i = 1, 2, \cdots, n$, with Wasserstein barycenter $Q_{t,\mathcal{B}}$, represented by the quantile function $F_{t,\mathcal{B}}^{-1} = \sum_{i=1}^{n} w_i F_{t,i}^{-1}$.

We can distinguish two cases:

1. The regulator has von Neumann preferences (see Section 2.3), specified by the barycentric expected utility using the Wasserstein barycenter $Q_{t,\mathcal{B}}$, so that the SDR is

$$r_{\mathcal{B}}(t) = \delta - \frac{1}{t} \ln \frac{\mathbb{E}_{Q_{t,\mathcal{B}}}[u'(C(t))]}{u'(C(0))}. \tag{14}$$

2. The regulator is uncertainty averse with an aversion towards model dispersion which is quantified by Fréchet-Wasserstein multiplier preferences for some multiplier θ (see Definition 1). For any fixed $t > 0$, a direct application of Theorem 2 for $X = C(t)$ allows the calculation of the utility functional $\mathcal{U}_\theta(C(t))$, whereas by repeating the arguments that led to the derivation of (14), we can see that the relevant SDR formula now assumes the form

$$r(t) = \delta - \frac{1}{t} \ln \frac{[\mathcal{M}(C(t))]}{u'(C(0))} = \delta - \frac{1}{t} \ln \frac{\mathbb{E}_{Q_t^*}[u'(C(t)]}{u'(C(0))}, \tag{15}$$

where the standard expected marginal utility is now replaced by $\mathcal{M}(C(t))$ (see Definition 2 and Proposition 2) and Q_t^* is the probability measure corresponding to the quantile function $(F_t^*)^{-1}$.

With regard to formula (15), it should be noted that:

(a) This seemingly simple formula takes model uncertainty fully into consideration since the effects of uncertainty are included in the minimizing quantile $(F_t^*)^{-1}$.

(b) In the limit as $\theta \to \infty$, $r(t) \to r_{\mathcal{B}}(t)$, and the barycentric SDR $r_{\mathcal{B}}$ provided by (14) is obtained.

(c) Since by Proposition 2 it holds that $\mathbb{E}_{Q_t^*}[u'(C(t)] > \mathbb{E}_{Q_{t,\mathcal{B}}}[u'(C(t)]$, and keeping in mind that $u'(C(0)) > 0$, we conclude that

$$r(t) < r_{\mathcal{B}}(t), \;\; t \in \mathbb{R}_+, \;\; \theta > 0, \tag{16}$$

which implies that the effect of uncertainty aversion is to decrease the SDR relative to the SDR obtained under risk aversion with expected utility defined in terms of the Wasserstein barycenter model $Q_{\mathcal{B}}$. This can be regarded as a second-order precautionary effect.

(d) The perturbative expansions obtained in Section 3 can be used to analytically approximate the SDR using formula (15). Such approximations, which can provide information on the dependence of the SDR on various parameters of interest (such as θ or in the case of CRRA utilities the risk aversion coefficient γ), are provided in the online Appendix EC.2.5. As shown there, the effect of uncertainty aversion, at least to first order in $\frac{1}{\theta}$, is to decrease the SDR as compared to the barycentric one.[4] The relevant comparison for general θ follows from Proposition 2. Moreover, in the general case and without the need to resort to any approximation, one can directly calculate the SDR using formula (15) numerically, using the algorithm presented in the online Appendix (Section EC.2.5 and in particular Section EC.2.5.5).

4.1.2. Numerical experiments In this section we present some numerical experiments which clarify our approach in estimating the SDR under multiple models and uncertainty aversion.

The case of a single model Following Gollier (2013, Ch. 4), we assume that the consumption process $C(t)$ follows a single factor (autoregressive) model of the form

$$\begin{aligned} C(t+1) &= C(t)\exp(x(t)), \\ x(t+1) &= \mu + y(t) + \varepsilon_x(t), \\ y(t) &= \phi y(t-1) + \varepsilon_y(t), \end{aligned} \tag{17}$$

where $\varepsilon_x(t), \varepsilon_y(t)$ are independent and serially independent with $\mathbb{E}[\varepsilon_x(t)] = \mathbb{E}[\varepsilon_y(t)] = 0$ and $Var(\varepsilon_x(t)) = \sigma_x^2$, $Var(\varepsilon_y(t)) = \sigma_y^2$, y_{-1} is some initial state, and $\phi \in [0,1]$ is a parameter representing the degree of persistency (mean reversion) of y. The choice $\phi = 0$ reduces the model to a standard random walk model which is a discretization of a Wiener process. The case where $\phi \neq 0$ corresponds to a discretization of an Ornstein-Uhlenbeck process. Typically, $\{y(t)\}$ is an unobserved stochastic factor, which has an effect on the observed growth rate $\{x(t)\}$ of the consumption process $\{C(t)\}$.

[4] Gierlinger and Gollier (2009) study the SDR under ambiguity in the context of smooth ambiguity preferences (Klibanoff et al. 2005). They identify an ambiguity prudence effect which tends to reduce the SDR and a pessimism effect which has an ambiguous impact. However Gierlinger and Gollier (2017) indicate that, under variational preferences, pessimism reduces the SDR.

A straightforward induction procedure shows that, given ϕ and y_{-1}, the stochastic consumption process $\{C(t)\}$ is lognormally distributed and in particular

$$\ln C(t) - \ln C(0) \sim N(\mu_t, \sigma_t^2),$$

where

$$\mu_t = \mu\, t + y_{-1}\frac{1-\phi^t}{1-\phi},$$
$$\sigma_t^2 = \frac{\sigma_y^2}{(1-\phi)^2}\left[t - 2\phi\frac{\phi^t - 1}{\phi - 1} + \phi^2\frac{\phi^{2t}-1}{\phi^2-1}\right] + \sigma_x^2 t.$$

When all the parameters and the distributions of noise terms concerning model (17) are fully known, i.e., when we are in a world of a single model, the Ramsey formula can be used to produce a term structure for the SDR (Gollier 2013). Using the general class of CRRA utilities, Gollier produces an analytic formula for the term structure of the discount rate as

$$r(t) = \delta + \gamma\frac{1}{t}\mu_t - \frac{1}{2}\gamma^2\frac{1}{t}\sigma_t^2.$$

Bansal and Yaron (2004) calibrated the factor model for consumption to data from the period 1929-1998 using annual data from the USA, producing estimates for the monthly mean return and volatilities of $\mu = 0.0015$, $\sigma_x = 0.0078$, $\sigma_y = 0.00034$, and estimated the reversion parameter as $\phi = 0.979$. Using these parameter values, Gollier implemented formula (17) to produce a term structure which is increasing or decreasing depending on the sign of y_{-1}. In particular Gollier used a range of values for $y_{-1} \in [-0.001, 0.001]$ for his numerical experiments for the term structure. In the case where $\phi = 0$, the term structure is flat.

Multiple models without uncertainty aversion Even if we trust the autoregressive model for the evolution of consumption, there are parameters related to the hidden variables included in the model, the value of which can be doubted. For the sake of illustration consider the two parameters ϕ and y_{-1}. Different estimations or opinions regarding the parameters $p = (\phi, y_{-1})$ produce different parameters μ_t and σ_t^2, therefore different models for the distribution of $C(t)$. The important question that arises is how the emergence of multiple models affects the SDR relative to the single model case (17).

We study this case when each model in the set $\mathbb{M}_t$ is a lognormal model of the type presented in (17). Each has parameters with values $\mu = 0.0015$, $\sigma_x = 0.0078$, $\sigma_y = 0.00034$ (see Bansal and Yaron 2004), but has a different choice for the parameters ϕ or y_{-1} or both. It is quite natural to allow some uncertainty for these parameters and in particular for y_{-1} as these are the ones whose estimation is more delicate. For the needs of the experiments presented here, we have generated the set of models $\mathbb{M}_t$ by selecting the parameters ϕ and y_{-1} corresponding to each model in $\mathbb{M}_t$ by uniform sampling from the intervals $[0.977, 0.98]$ and $[-0.001, 0.001]$ respectively.[5] Then, for the given set of models $\mathbb{M}_t$, and assuming no uncertainty aversion for the regulator, we calculate the term structure of barycentric SDR $r_{\mathcal{B}}(t)$.

The results for a typical calculation of this type are displayed in Figure 2, where 10 models were considered as comprising the model set $\mathbb{M}_t$. Each model was used to calculate the term structure for r over the next 50 years (blue lines), while $r_{\mathcal{B}}$ (red line) was calculated by assigning equal weight $w_i = 1/10$ to each model in $\mathbb{M}$ and using the algorithm in section EC.2.5.5. Note that the difference in the shape of the term structure for the individual realizations of the models in Figure 2 (blue lines) arises from the different values of the parameter y_{-1} used in each simulation. As pointed out in Gollier (2013), the shape of the term structure curves depends on the sign of y_{-1}, with $y_{-1} > 0$ and $y_{-1} < 0$ leading to convex and concave curves respectively. In particular, Gollier notes that if the recent growth rate is exactly at its historical mean ($y_{-1} = 0$), the yield curve is decreasing with the slope describing the precautionary effect of the increasing annualized variance of future log consumption due to the persistence of shocks. During a downturn ($y_{-1} < 0$), the yield curve is upwards sloping, with the shape mostly expressing an accelerating wealth effect generated by rising growth expectations, while when the economy is booming ($y_{-1} > 0$), the yield curve is decreasing because of diminishing expectations. This shows that different views concerning y_{-1} may lead to different yield curves, and this is the effect of the presence of various models (blue lines). On the other hand, the red curve is the average model (barycentric model) which takes into account all models, properly weighted through their Wasserstein distance.

Multiple models under uncertainty aversion Finally, assuming uncertainty aversion and Fréchet multiplier preferences, we use expansion (EC.16) for suitably truncated lognormals, to assess the effects of the uncertainty aversion parameter θ on the term structure of the SDR $r(t)$. In Figure 3 we display the results of two numerical experiments. In the first panel we assume that the barycentric model $Q_{\mathcal{B}}$ corresponds to a log-normal distribution derived from the Gollier model (17) for the choice of parameters $\mu = 0.0015$, $\sigma_x = 0.0078$, $\sigma_y = 0.00034$ (Bansal and Yaron 2004) and $\phi = 0$. The red line corresponds to the term structure as described by the Wasserstein barycenter model $r_{\mathcal{B}}(t)$ while the blue line corresponds to the term structure when the uncertainty aversion parameter is $\theta = 10$. Note that $r_{\mathcal{B}}(t)$ produces a flat term structure, as expected, since in the case when $\phi = 0$, model (17) is a random walk model and $r_{\mathcal{B}}$ is defined in terms of expected utility (Gollier 2013). On the other hand, the uncertainty averse discount rate r displays a term

[5] The reason for choosing such an interval of values for ϕ is that ϕ is estimated with a statistical error and the range of the interval corresponds to possible values of the statistical estimator $\hat{\phi}$ for ϕ, while y_{-1} is not specified at all as it corresponds to a hidden variable in the model.

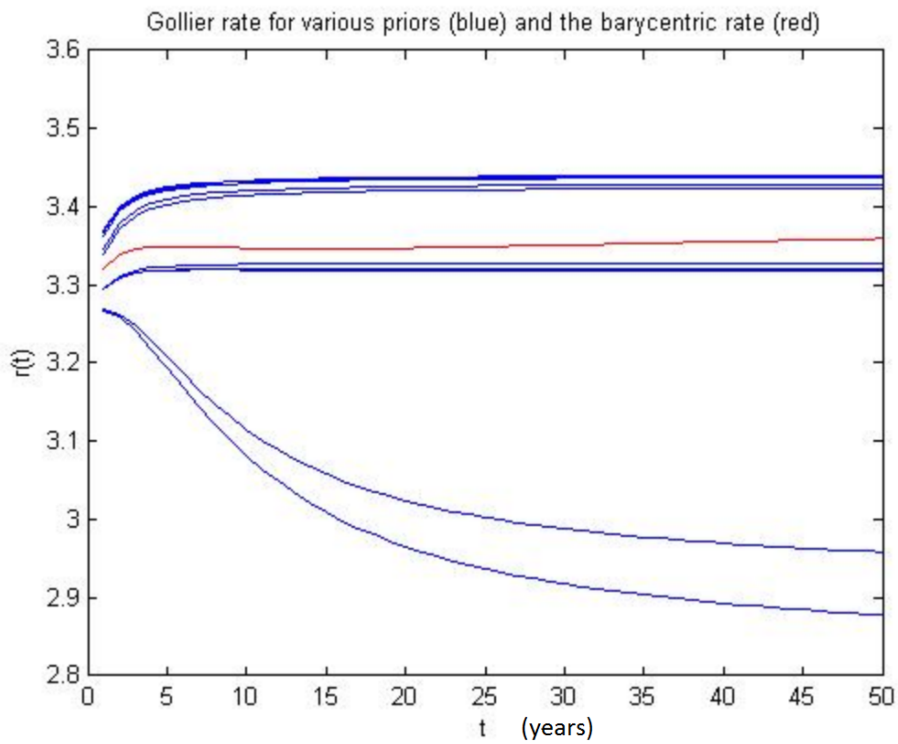

Figure 2 **The individual term structure curves (blue) and the barycentric term structure (red).**

structure and, furthermore, $r(t) \leq r_{\mathcal{B}}(t)$ as expected from (16). The discrepancy between $r_{\mathcal{B}}$ and r appears for small t, while over longer horizons the two discount rates approach each other. In the second panel, the parameter values are the same, with the exception of ϕ which is now chosen to be $\phi = 0.979$, thus rendering model (17) a mean reverting model. The set of models $\mathbb{M}$ consists of probabilistic models for $C(t)$ drawn from the general class of stochastic models (17), where each model in $\mathbb{M}$ is generated by keeping the parameters $\mu = 0.0015$, $\sigma_x = 0.0078$, $\sigma_y = 0.00034$, $\phi = 0.979$ in (17) fixed, and varying the parameter y_{-1} in the range $[-0.001, 0.001]$, using equal weights for all scenarios. In this case, $r_{\mathcal{B}}$ (the red line) displays term structure as well, as expected (see Gollier 2013), and so does the uncertainty averse discount rate r (drawn for $\theta = 10$ as the blue line). Note that $r(t) \leq r_{\mathcal{B}}(t)$, the difference being more prominent over short horizons, while becoming rather weak over longer horizons. Finally, by expansion (EC.15), $r_{\mathcal{B}}(t) - r(t)$ is decreasing with increasing θ for large enough values of θ.

4.2. The Wasserstein barycenter in consensus group decision making: Application to CAT bond pricing

As a final application, we consider the use of the Wasserstein barycenter in group decision making and show that it can be used in the formulation of proposals that are most likely to lead to consensus of a group of DMs with varying models concerning an unknown risk. In particular we

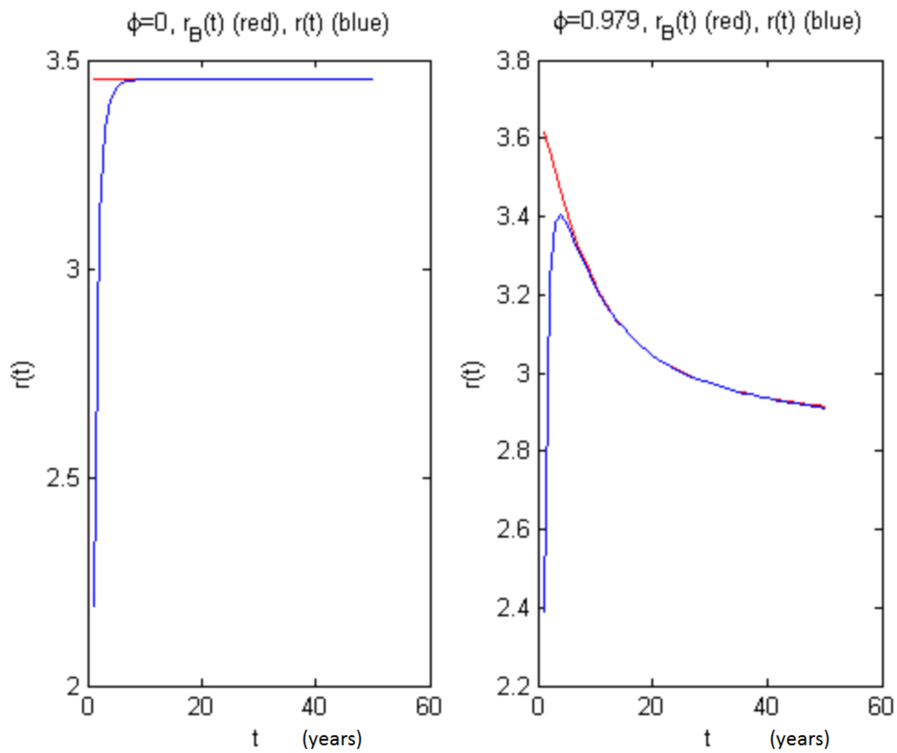

Figure 3 **Comparison of the term structure predicted by the barycentric model $r_{\mathcal{B}}(t)$ (red line) and the uncertainty averse Fréchet-Wasserstein model $r(t)$ (blue line) for two different choices of the parameters in the Gollier model (17).**

show that the use of the Wasserstein barycenter as a commonly acceptable probability model upon which a decision is made is the one which maximizes the probability of reaching consensus.

While the above framework can be encountered in a wide variety of situations, we choose as a motivating example the issuance and pricing of a CAT bond.

CAT bonds are risk sharing instruments, used by a group of firms in the management of extreme risks such as natural and climate-change related disaster risks (e.g., earthquakes, hurricanes, floods, fires) and cybersecurity risks. CAT bonds are issued by insurance and reinsurance firms, corporations, government bodies and others. CAT bonds have attracted interest in both the academic and the more applied case study oriented literature. Various pricing methods were proposed, including actuarial type pricing methods, stochastic pricing models based on the Poisson or the doubly stochastic Poisson process, utility-based pricing methods or econometric-based methods (see e.g. Burnecki and Kukla 2003, Froot 2007, Froot and O'Connell 2008, Cummins and Barrieu 2013, Galeotti et al. 2013, Edesess 2015). The important effects of ambiguity on the spreads of CAT bonds have been noted quite early in the literature (see e.g. Bantwal and Kunreuther 2000). Another important aspect of CAT bonds is the involvement of multiple agents in their design and pricing (see e.g. Edesess 2015), a fact that affects the calculation of the expected loss, which according to econometric studies is one of the key drivers in CAT bonds prices (see e.g. Galeotti et al. 2013). The application in this section focuses on these two last aspects, and in particular on the determination of a model for the risk which maximizes the probability of acceptance by all parties involved, and on the resulting pricing of the CAT bond. For a detailed review on CAT bonds, including their design and mechanics, see Cummins and Barrieu (2013) and the references therein.

As a consequence of their nature, more than one agent is involved in the issuance of a CAT bond, such as insurers; reinsurers; corporations; pension funds; structuring agents who assist the issuer in selecting trigger type and are involved in placing the bond with investors (investment banks or brokers); modelling agents who estimate the risk based on models and simulations (e.g., Risk Management Solution, Inc, or Eqcat); ratings agencies and others (see for example Edesess 2015).

All, or the majority, of these different actors must agree on some common characteristics concerning the contract structure of the CAT bond, which are related to a common agreement concerning the estimation of the extreme risk. Since extreme risks are by nature rare events, the lack of sufficient historical data places them within the realm of model uncertainty, as it is not possible on the basis of statistical evidence to single out a unique probabilistic model for the random variable L corresponding to the risk. On the other hand, in order for all parties to agree upon the issue and the actual contract terms, a commonly agreed model for the distribution of the extreme risk must be adopted. The agreement is a necessity, as the issuance of a risk sharing instrument is of mutual benefit to all parties. Since in principle each agent involved may have a different prior for the risk, the valuation has to be effected by a commonly agreed probability model for the risk, or at least by a model which the agents involved have the maximum probability of agreeing upon. The issue of identifying such a model of common acceptance is important for the construction and pricing of the CAT bond.

Before proceeding further, it is useful to present more details concerning the structure and function of CAT bonds. This type of instrument has become very popular in recent years, as a vehicle for transferring extreme risks from insurers and re-insurers to investors. It constitutes a tool that enables: (a) extreme risks to be covered more efficiently, providing solvency to those involved in the insurance business; and (b) attractive investment opportunities with potentially high returns to be provided to investors, which are largely uncorrelated to other market indices, hence offering at the same time a useful hedging tool. The basic structure of a CAT bond is as follows. A sponsor or group of sponsors, typically a reinsurer, contacts a special purpose vehicle (SPV) in order to enter an alternative reinsurance contract which will guarantee solvency in case of occurrence of extreme losses. The sponsor, at the cost of some premium ρ, will receive insurance coverage up to some level h, in the case of extreme losses. The SPV for its own coverage, and in order to guarantee the possibility of covering amount h for the sponsor, issues a CAT bond, which is a standard defaultable bond, with the default triggered by the event of extreme losses of the sponsor. Several payoff structures are possible: if the amount h is issued in bonds, then (a) in the absence of a triggering event, the bond provides coupons to the investors corresponding to interest $r+\rho$, where r is a standard interest rate (e.g., LIBOR) and a principal h; while (b) in the presence of the triggering event, coupons are reduced to $(r+\rho)(1-d_1)$ and the principal is reduced to $h(1-d_2)$, for suitable d_1, d_2. Other payoff structures are possible, and there exists a variety of CAT-based derivatives such as CAT swaps which provide a multitude of opportunities for risk sharing and efficient risk management.

However, the success of such instruments, especially in the primary market, crucially depends on the choice of the premium ρ, which in turn is related to the spread of the CAT bond. Numerous theoretical and empirical studies have shown that the most important quantity is the expected loss $EL = \mathbb{E}[G(L)]$ where L is the random variable corresponding to the catastrophic risk and $G : \mathbb{R}_+ \to \mathbb{R}_+$ is an appropriate function related to the cover agreement between the sponsor(s) and the SPV.[6] One commonly used model for the spread is a linear model of the form $\rho = c_1 + c_2 EL$ (Galeotti et al. 2013) for appropriate constants c_1, c_2 which are determined by linear regression and may incorporate geographic or seasonal effects. Other models are based upon utility pricing arguments and result in nonlinear models of the form $\rho = EL + \gamma(PFL)^{\alpha}(CEL)^{\beta}$, where $PFL = P(L > a)$ is the probability of first loss and $CEL = \frac{\mathbb{E}[L_{(a,a+h)}|L>a]}{h} = \frac{\mathbb{E}[L_{(a,a+h)}]}{hP(L>a)}$ is the conditional expected loss.[7]

4.2.1. Consensus achievement and pricing of CAT bonds The above discussion clearly indicates the need for agreement on the probability $P(L \geq x)$, no matter which pricing methodology is adopted for the CAT bond. This requires the development of a scheme which allows the parties involved in the design of the CAT bond to reach consensus concerning the probability model for the extreme risk. Even though each agent may have a different prior concerning the probability $P(L \geq x)$, it is in their common interest to agree on a common model that will favor the issuance of the bond, hence each agent will be willing to change her/his initial prior and accept a new probability model for L as long as the uneasiness caused by this change is not too high. It is reasonable to assume that this uneasiness is an increasing function of the distance $d(Q, Q_i)$ between the prior Q_i of agent i, concerning the risk, and the adopted probability measure Q. Put differently, the probability p_i of agent i accepting the probability measure Q can be expressed as $p_i = \varphi_i(d^2(Q, Q_i))$, where $\varphi_i : \mathbb{R}_+ \to [0,1]$ is a decreasing function characterizing the strength of each agent's belief in

[6] For example, a typical agreement in such type of contracts is the cover of a particular tranche of the catastrophic risk, e.g. $G(L) = L_{(a,a+h)} = 0\,\mathbf{1}_{L\leq a} + (L-a)\,\mathbf{1}_{a<L\leq a+h} + h\,\mathbf{1}_{L>a+h}$, for appropriately chosen a and h.

[7] A different type of popular models for pricing CAT bonds is models using the Wang distortion operator (Wang 2000, Galeotti et al. 2013) in which the probability of loss $P(L \geq x)$ plays an important role, as the premium is expressed in terms of $\rho = \frac{1}{h}\int_a^{a+h} \mathfrak{g}(P(L \geq x))dx$ where $\mathfrak{g} : [0,1] \to [0,1]$ is a nonlinear function such that $\mathfrak{g}(0) = 0$, $\mathfrak{g}(1) = 1$, $\mathfrak{g}' \geq 0$, $\lim_{s\to 0} \mathfrak{g}'(s) = \infty$ and $\mathfrak{g}'' \leq 0$.

the prior and her/his willingness to move.[8] If the agents are independent, the probability of all of them agreeing with the mediator's proposal is equal to $p = p_1 \cdots p_n$. The choice of the probability measure Q that satisfies as many agents as possible can then be expressed as the problem of choosing Q so as to maximize probability p. Since the problem of maximizing p is equivalent to the problem of maximizing $\ln(p)$ under the above assumptions, it can be seen that the probability measure which will maximize the probability of attaining a consensus is the solution to

$$\max_{Q \in \mathcal{P}} \sum_{i=1}^{n} \ln(\varphi_i(d^2(Q, Q_i)). \tag{18}$$

We adopt the metrization of the space of probability measures in terms of the Wasserstein metric $d(Q, Q_i) = W_2(Q, Q_i)$ and consider the corresponding problem (18). We will show that the probability measure Q which maximizes the probability of all agents agreeing to it, is the Fréchet barycenter of the set of models, with a choice of weights which corresponds to the functions φ_i.

PROPOSITION 3. *Assume that the models Q_i, $i = 1, \cdots, n$, for risk L are expressed in terms of the probability distributions F_i and the corresponding quantiles F_i^{-1} and define the quantities $M_{ij} := \int_0^1 F_i^{-1}(s) F_j^{-1}(s) ds$, $i, j = 1, \cdots, n$. Moreover, assume sufficient smoothness and integrability conditions for the decreasing functions $\varphi_i : \mathbb{R}_+ \to [0,1]$ and let $\Phi_i := -\ln(\varphi_i)$, $i = 1, \cdots, n$, be increasing and convex.*

A probability measure $Q \in \mathcal{P}(\mathbb{R})$ that maximizes the probability of agreement of all agents coincides with the Wasserstein barycenter $Q_{\mathcal{B}}$ represented by the distribution function F_B given by the quantile average

$$F_B^{-1} = \sum_{i=1}^{n} w_i F_i^{-1},$$

where the weights $w = (w_1, \cdots, w_n)$ are solutions of the set of algebraic equations

$$w_i = \frac{\Phi_i'(\Lambda_i(w))}{\sum_{j=1}^{n} \Phi_j'(\Lambda_j(w))}, \ \ i = 1, \cdots, n, \tag{19}$$

with

$$\Lambda_i(w) := M_{ii} - 2\sum_{\ell=1}^{n} M_{i\ell} w_\ell + \sum_{\ell=1}^{n}\sum_{k=1}^{n} M_{\ell k} w_\ell w_k, \ \ i = 1, \cdots, n. \tag{20}$$

Proof: See Appendix D. □

Proposition 3 shows that the probability measure for L that maximizes the probability of agreement of all agents corresponds to the Wasserstein barycenter with a particular choice of weights,

[8] We express ϕ_i in terms of $d^2(Q, Q_i)$ for ease, as $d^2(Q, Q_i)$ displays smoothness properties which make the analysis more explicit; clearly, upon redefining ϕ_i, this probability could be expressed in terms of $d(Q, Q_i)$.

which are endogenously determined in terms of the elasticities of the functions φ_i which model the rigidity of the various agents to their priors. In some sense these reflect the bargaining power or authority of each agent in the group.[9] If agents are symmetric with $\varphi_i(z) = \exp(c_i - c\,z)$, $i = 1, \cdots, n$ for c_i, c appropriate constants and with c_i possibly varying from agent to agent but c being the same for all agents, the resulting weights will be $1/n$. This interpretation of the Wasserstein barycenter provides a further argument in favor of its use as a decision-making tool under model uncertainty. Other schemes for the choice of the weights w_i could be adopted in place of the one in Proposition 3. For example, the weights can be assigned to the models in terms of a learning scheme, where more weight is assigned to models which perform better in predictions of the risk, whereas models which underperform in this task are penalized (e.g., Papayiannis and Yannacopoulos 2018).

We are now ready to proceed with the pricing of the CAT bond. A class of suitable models for the extreme risk is the class of generalized extreme value (GEV) distributions, described by the probability distribution functions

$$F_i(x) = \begin{cases} \exp\left(-\left(1+\xi_i\left(\frac{x-\mu_i}{\sigma_i}\right)\right)^{-1/\xi_i}\right), & \xi_i \neq 0, \\ \exp\left(-e^{-\frac{(x-\mu_i)}{\sigma_i}}\right), & \xi_i = 0, \end{cases}$$

which can be inverted to explicitly obtain the quantile functions

$$F_i^{-1}(s) = \begin{cases} \mu_i + \frac{\sigma_i}{\xi_i}[(-\ln s)^{-\xi_i} - 1], & \xi_i \neq 0, \\ \mu_i - \sigma_i \ln(-\ln s), & \xi_i = 0, \end{cases}$$

where $\mu_i \in \mathbb{R}$ is a location parameter, $\sigma_i > 0$ is a scale parameter and $\xi_i \in \mathbb{R}$ is a shape parameter. The support of the distribution depends on the scale parameter, with

$$\operatorname{supp}(F_i) = \begin{cases} [\mu_i - \frac{\sigma_i}{\xi_i}, +\infty), & \xi_i > 0, \\ (-\infty, \infty), & \xi_i = 0, \\ (-\infty, \mu_i - \frac{\sigma_i}{\xi_i}), & \xi_i < 0. \end{cases}$$

Note that the above distributions do not necessarily have the same support.

Given a set of weights $w = (w_1, \cdots, w_n)$ determined in the context of Proposition 3, the corresponding Wasserstein barycenter is $F_{\mathcal{B}}^{-1} = \sum_{i=1}^n w_i F_i^{-1}$. In the case where all models correspond to the same shape parameter $\xi_i = \xi$, the Wasserstein barycenter corresponds to a member of the GEV family with $\mu_B = \sum_{i=1}^n w_i \mu_i$, $\sigma_B = \sum_{i=1}^n w_i \sigma_i$ and $\xi_B = \xi$. In the general case, the quantile function $F_{\mathcal{B}}^{-1}$ is explicitly known and its inversion is therefore an easy numerical task (though not feasible

[9] Note that the problem studied in Proposition 3 is formally similar to (and in fact inspired by) a Nash bargaining game in the space of probability models (measures). In such an interpretation, each prior is associated with the preferred point of each agent in bargaining space, whereas the disutility which arises from shifting from the preferred point is a function of the distance between the initial and the final points. In our interpretation, this disutility is associated with the probability of accepting the new probability model.

in closed form). The family of GEV distributions, which encompasses in one family the three types of extreme value distributions (Gumbel, Weibull and Fréchet), has been successfully used in the literature to model the distribution of extreme risks such as earthquakes and floods.

The premium will be determined in terms of the quantity $\mathbb{E}[L_{(a,a+h]}]$ which will now be calculated under the Wasserstein barycenter, in terms of

$$EL = \mathbb{E}_{Q_{\mathcal{B}}}[L_{(a,a+h]}] = \int_0^h S_B(a+y)dy = h - \int_0^h F_B(a+y) = h - \int_a^{a+h} F_B(y)dy, \tag{21}$$

or through the equivalent representation

$$EL = \mathbb{E}_{Q_{\mathcal{B}}}[L_{(a,a+h]}] = \int_0^1 G(F_B^{-1}(s))ds = \int_{F_B(a)}^{F_B(a+h)} (F_B^{-1}(s) - a)ds + h(1 - F_B(a+h)). \tag{22}$$

Both representations are easily computed numerically for the Wasserstein barycenter. For certain special cases, such as when all the agents have models with the same shape parameter $\xi_i = \xi$ (which is a reasonable assumption for certain types of extreme risks which are modelled by the Gumbel type), the calculation can be performed analytically. In such cases,

$$F_{\mathcal{B}}^{-1}(s) = \begin{cases} \mu_B + \frac{\sigma_B}{\xi}[(-\ln s)^{-\xi} - 1], & \xi \neq 0, \\ \mu_B - \sigma_B \ln(-\ln s), & \xi = 0, \end{cases} \;\; \text{and} \;\; F_B(x) = \begin{cases} \exp\left(-\left(1 + \xi\left(\frac{x-\mu_B}{\sigma_B}\right)\right)^{-1/\xi}\right), & \xi \neq 0, \\ \exp\left(-e^{-\frac{(x-\mu_B)}{\sigma_B}}\right), & \xi = 0, \end{cases}$$

so that EL can be approximated in terms of the exponential integral function E_1 or an appropriate series expansion.

5. Concluding Remarks

We study decision making under model uncertainty which is characterized by multiple probability measures, or models, associated with a random variable. We use as the basis for decision making in such a context a probability measure which is the Fréchet mean, or the barycenter, of the set of probability measures. By introducing ambiguity aversion, we show that Fréchet mean preferences are variational preferences, and provide their representations in terms of utility functionals.

The Fréchet-Wasserstein mean utility functionals developed in this paper could provide another representation of variational preferences when the framework of a benchmark model which determines a Kullback-Leibler entropy ball cannot be used. Using the Wasserstein metric provides a tractable representation of utility functionals which allows us to characterize changes in expected utility when different models are used, and allows for differential treatment of models according to a weight vector. It also overcomes some technical issues associated with the Kullback-Leibler entropy. Moreover, it allows for use of expansions in the ambiguity aversion parameter to derive a tractable quantitative relation between the degree of relative (or absolute) risk aversion and the parameter reflecting ambiguity aversion.

The tractability of the Fréchet-Wasserstein mean or multiplier preferences could be useful as a decision-making framework under uncertainty and uncertainty (or ambiguity) aversion for problems encountered in the economics of uncertainty. We thus consider two applications. In the first we show that the Fréchet-Wasserstein multiplier utility functional can be used to define and calculate, through an easy-to-apply algorithm, the SDR and its term structure under model uncertainty and ambiguity aversion, which could be useful in cost-benefit analysis. In the second we show that the Wasserstein barycenter could be a useful tool for achieving consensus among agents involved in the issuance of a CAT bond. The proposed framework could also be helpful in the study of many issues which are characterized by model uncertainty and aversion to ambiguity, such as climate policy, finance, insurance, real options or alternative investments.

Acknowledgments

We would like to thank an Associate Editor, the Department Editor Manel Baucells and two anonymous referees for valuable comments and suggestions on an earlier draft of this paper. The first and third authors acknowledge useful and enjoyable discussions with Professor N. A. Yannacopoulos. We further wish to thank Professor G. Papayiannis for suggesting the term Fréchet preferences and Professor S. Koukoulas for useful discussions and support. We are indebted to Joan Stefan for her invaluable assistance with technical editing.

Appendix. Proofs of main results

Appendix A: Proof of Theorem 1

Let $\mathcal{Q} \subset L^2([0,1])$ be the set of quantile functions. Here for simplicity we restrict attention to continuous square integrable quantile functions, i.e., we consider $\mathcal{Q}$ as a convex subset of $L^2([0,1])$ and let $\langle \cdot, \cdot \rangle$ be its standard inner product.

We express the Fréchet function in terms of quantiles as the functional $\mathsf{F}_{\mathbb{M}} : \mathcal{Q} \to \mathbb{R}$ defined by

$$\mathsf{F}_{\mathbb{M}}(g) = \sum_{i=1}^{n} w_i \int_0^1 |g(s) - F_i^{-1}(s)|^2 ds, \ \forall g \in \mathcal{Q}, \tag{23}$$

and using the same arguments as in the proof of Proposition 1 we rephrase (3) as the variational problem

$$\mathcal{U}(X) = \min_{g \in \mathcal{Q}} J(g) := \min_{g \in \mathcal{Q}} \left[\int_0^1 u(g(s)) ds + \phi(\mathsf{F}_{\mathbb{M}}(g)) \right]. \tag{24}$$

We relax the problem in $L^2([0,1])$, and assuming $u, \phi \in C^1$ we calculate the Gâteaux derivative of J, as

$$\begin{aligned} \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} (J(g+\epsilon \bar{g}) - J(g)) =: \langle DJ(g), \bar{g} \rangle &= \langle u'(g) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g)) \sum_{i=1}^{n} w_i (g - F_i^{-1}), \bar{g} \rangle \\ &= \langle u'(g) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g))(g - F_{\mathcal{B}}^{-1}), \bar{g} \rangle, \end{aligned}$$

where we used the Lebesgue dominated convergence theorem to pass to the limit and the definition of the Wasserstein barycenter.

The local minimizer g^* is an element of $\mathcal{Q}$ that satisfies the variational inequality

$$\langle DJ(g^*), g^* - \bar{g}\rangle = \langle u'(g^*) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g^*))(g^* - F_{\mathcal{B}}^{-1}), g^* - \bar{g}\rangle \leq 0, \ \forall \bar{g} \in \mathcal{Q}. \tag{25}$$

A candidate interior solution of this is a $g^* \in \mathcal{Q}$ such that

$$u'(g^*(s)) + 2\phi'\big(\mathsf{F}_{\mathbb{M}}(g^*)\big)\big(g^*(s) - F_{\mathcal{B}}^{-1}(s)\big) = 0, \ \ a.e. \ s \in [0,1], \tag{26}$$

which in turn can be constructed in terms of the solution $g_\rho(s)$ of (5) for a value of $\rho = \rho^*$ such that $\rho^* = 2\phi'(\mathsf{F}_{\mathbb{M}}(g_{\rho^*}))$. If the corresponding function g_{ρ^*} is a quantile then $F_*^{-1} = g_{\rho^*}$ is an admissible solution for (3) and the representation (6) follows. For further comments on the proof, regarding conditions under which the solution of (26) is a quantile as well as the connection of this property with the second-order conditions see the online Appendix, Section EC.2.2, Remarks EC.1 and EC.2.

Appendix B: Proof of Theorem 2

The proof requires some technical estimates which are provided in the online Appendix, Section EC.2.3. Here we sketch the main ideas of the proof.

(a) Applying Theorem 1, for $\phi(x) = \frac{\theta}{2}(x - \mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}}))$, we conclude that a minimizer Q_θ can be expressed in terms of the solution of (5) for $\rho = \theta$ if this solution represents a quantile. For the solvability of (5) for all $s \in [0,1]$, the function F_B^{-1} must be bounded away from 0, a fact which is guaranteed by Assumption 2. On the other hand, problems with the solvability of (5) may arise from the term $\frac{1}{\theta}u'(g(s))$ if θ is too small. This requires restricting $\theta > \theta_c$, with the critical value θ_c, depending on the constants of Assumption 2 and the utility function, arising from a careful balancing of the terms in (5) so that a suitable solution exists. By further restricting θ, the resulting solution has monotonicity properties that allow it to be a quantile function. These arguments are stated in detail in Section EC.2.3.

The first representation in (7) follows by direct substitution of the minimizer in the functional while the second by observing that by (5), $F_{\mathcal{B}}(s)^{-1} - F_\theta^{-1}(s) = \frac{1}{\theta}u'(F_\theta^{-1}(s)) \geq 0$, for all $s \in [0,1]$. From (5) it is clear that $F_\theta^{-1} \to F_{\mathcal{B}}^{-1}$ as $\theta \to \infty$ and from that (with a suitable application of Lebesgue's dominated convergence theorem) the convergence of the utility functional follows. The bound $\mathcal{U}_\theta(X) \leq \mathcal{U}_{\mathcal{B}}(X)$ for all $\theta > 0$ follows by the variational definition of $\mathcal{U}_\theta(X)$ from which it follows that $\mathcal{U}_\theta(X) \leq \mathbb{E}_{Q_{\mathcal{B}}}[u(X)] + \phi(Q_{\mathcal{B}})$ and the fact that for multiplier preferences $\phi(Q_{\mathcal{B}}) = 0$.

(b) We assume an expansion of the solution of (5) in terms of $F_\theta^{-1}(s) = g_0 + \frac{1}{\theta}g_1(s) + \frac{1}{\theta^2}g_2(s) + ...$ for $g_0, g_1, g_2, \cdots$, to be determined. Substituting this ansatz in (5), by Taylor expansion of $u'(F_\theta^{-1}(s))$ in terms of the small parameter $\frac{1}{\theta}$, and matching orders we obtain to second order the expression for F_θ^{-1} in (8). Then, the expansion for $\mathcal{U}_\theta$ in (8) follows by direct substitution of the expansion for F_θ^{-1} into (7). Assuming higher regularity of u, one may continue the expansion to higher orders.

Appendix C: Proof of Proposition 2

(a) Working as in the proof of Theorem 2, in terms of the quantiles, and setting for notational simplicity $g^*(\cdot, \epsilon) = F_{\theta, X+\epsilon}^{-1}(\cdot)$ for the quantile corresponding to the minimizer of problem (4) for the random variable $X + \epsilon$, we have the representation

$$\mathcal{U}_\theta(X + \epsilon) = \int_0^1 u(g^*(s,\epsilon) + \epsilon)ds + \frac{\theta}{2}\int_0^1 (g^*(s,\epsilon) - F_{\mathcal{B}}^{-1}(s))^2 ds, \tag{27}$$

where $g^*(\cdot,\epsilon)$ is the solution to the variational inequality

$$\int_0^1 (u'(g^*(s,\epsilon)+\epsilon)+\theta(g^*(s,\epsilon)-F_{\mathcal{B}}^{-1}(s))(g^*(s)-\bar{g}(s))ds \leq 0, \;\; \forall \bar{g} \in \mathcal{Q}. \tag{28}$$

Since $\theta > \theta_c$, we consider interior solutions, so that $g^*(s,\epsilon)$ is the solution of the parametric algebraic equation

$$\frac{1}{\theta}u'(g^*(s,\epsilon)+\epsilon)+g^*(s,\epsilon)=F_{\mathcal{B}}^{-1}(s), \;\; s\in[0,1]. \tag{29}$$

Clearly, by continuity arguments, $g^*(s,0)=F_\theta^{-1}(s)$, where $g^*(s,0)$ solves

$$\frac{1}{\theta}u'(g^*(s,0))+g^*(s,0)=F_{\mathcal{B}}^{-1}(s), \;\; s\in[0,1],$$

whereas under suitable smoothness assumptions on u it can be shown that $g^*(s,\epsilon)$ is differentiable with respect to ϵ. The marginal utility is given by the derivative of $\mathcal{U}_\theta(X+\epsilon)$ with respect to ϵ, calculated at $\epsilon=0$,

$$\mathcal{M}_\theta(X) := \frac{d}{d\epsilon}\mathcal{U}_\theta(X+\epsilon)\bigg|_{\epsilon=0}.$$

Using Lebesgue's dominated convergence theorem to differentiate under the integral sign in (27), we can express

$$\frac{d}{d\epsilon}\mathcal{U}_\theta(X+\epsilon)=\int_0^1\left\{u'(g^*(s,\epsilon)+\epsilon)+\left(u'(g^*(s,\epsilon)+\epsilon)+\theta\Big(g^*(s,\epsilon)-F_{\mathcal{B}}^{-1}(s)\Big)\right)\frac{\partial}{\partial\epsilon}g^*(s,\epsilon)\right\}ds,$$

which using (29) simplifies to

$$\frac{d}{d\epsilon}\mathcal{U}_\theta(X+\epsilon)=\int_0^1 u'(g^*(s,\epsilon)+\epsilon)ds.$$

Passing to the limit as $\epsilon\to 0$, using once more Lebesgue's dominated convergence and continuity of u', we conclude that

$$\mathcal{M}_\theta(X)=\int_0^1 u'(g^*(s,0))ds=\int_0^1 u'(F_\theta^{-1}(s))ds.$$

If Q_θ is the probability measure minimizing problem (4) for the random variable X – corresponding to the quantile F_θ^{-1} – an equivalent representation is

$$\mathcal{M}_\theta(X)=\mathbb{E}_{Q_\theta}[u'(X)].$$

Since $F_\theta \leq F_{\mathcal{B}}^{-1}$ and u' is a decreasing function, it follows that $u'(F_\theta^{-1})\geq u'(F_{\mathcal{B}}^{-1})$ hence $\mathcal{M}_\theta(X)\geq\mathcal{M}_{\mathcal{B}}(X)$.

(b) The claims follow using a perturbative expansion as in the proof of Theorem 2.

Appendix D: Proof of Proposition 3

We express problem (18) in terms of quantile functions (assumed for simplicity to be continuous and $L^2([0,1])$) as

$$\max_{Q\in\mathcal{P}(\mathbb{R})}\sum_{i=1}^n \ln\left(\varphi_i(d^2(Q,Q_i))\right)=\min_{g\in\mathcal{Q}} I(g) := \min_{g\in\mathcal{Q}}\sum_{i=1}^n \Phi_i\left(\int_0^1 |g(s)-F_i^{-1}(s)|^2 ds\right).$$

We relax the problem in $L^2([0,1])$ and - using the smoothness of Φ_i and the Lebesgue dominated convergence theorem to pass to the limit - calculate the Gâteaux derivative DI of the functional I as

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{1}{\epsilon}(I(g+\epsilon\bar{g})-I(g))=\langle DI(g),\bar{g}\rangle &= 2\sum_{i=1}^n \Phi_i'\left(\int_0^1 |g(s)-g_i(s)|^2 ds\right)\int_0^1 (g_i(s)-g(s))\bar{g}(s)ds \\ &= 2\langle\sum_{i=1}^n \Phi_i'\left(\int_0^1 |g(s)-F_i^{-1}(s)|^2 ds\right)(F_i^{-1}-g),\bar{g}\rangle,\end{aligned}$$

where by $\langle\cdot,\cdot\rangle$ we denote the inner product in $L^2([0,1])$.

By the assumptions on φ_i – in particular since ϕ_i are assumed log-concave – a candidate for a local extremum is the solution g^* of the first-order condition $DI(g^*)=0$, if it corresponds to a quantile function. The first-order condition yields

$$\sum_{i=1}^n \Phi_i'\left(\int_0^1 |g(s)-F_i^{-1}(s)|^2 ds\right)(g-F_i^{-1})=0. \tag{30}$$

Let g be any solution of (30). Defining the real numbers $\bar{w}_i=\Phi_i'\left(\int_0^1 |g(s)-F_i^{-1}(s)|^2 ds\right)$ for $i=1,\cdots,n$, we express (30) as

$$\sum_{i=1}^n \bar{w}_i(g-F_i^{-1})=0,$$

which is readily solved to yield

$$g=\sum_{i=1}^n w_i F_i^{-1}, \text{ where } w_i=\frac{\bar{w}_i}{\bar{w}_1+\cdots+\bar{w}_n}. \tag{31}$$

By the properties of Φ, we have that $w=(w_1,\cdots,w_n)\in\Delta^{n-1}$.

Hence, any solution of (30) is a quantile average of the form (31) for an appropriate choice of $w\in\Delta^{n-1}$. It remains to determine the choice of w. For this we recall the definition of $\bar{w}=(\bar{w}_1,\cdots,\bar{w}_n)$, substitute g as in (31) and obtain the set of algebraic equations

$$\bar{w}_i=\Phi_i'\left(\int_0^1 |\sum_{j=1}^n w_j F_j^{-1}(s)-F_i^{-1}(s)|^2 ds\right),\ i=1,\cdots,n, \tag{32}$$

which if solved will provide us with the value of $\bar{w}_i$ (or equivalently w_i) for $i=1,\cdots,n$.

It is convenient to express the system (32) in terms of $w\in\Delta^{n-1}$ only. By a straightforward calculation

$$\begin{aligned}\Lambda_i(w)&:=\int_0^1 |\sum_{j=1}^n w_j F_j^{-1}(s)-F_i^{-1}(s)|^2=\int_0^1 \big(\sum_{j=1}^n w_j F_j^{-1}(s)-F_i^{-1}(s)\big)\big(\sum_{\ell=1}^n w_\ell F_\ell^{-1}(s)-F_i^{-1}(s)\big)ds\\&=\sum_{j=1}^n\sum_{\ell=1}^n w_j w_\ell\int_0^1 F_j^{-1}(s)F_\ell^{-1}(s)ds-2\sum_{j=1}^n w_j\int_0^1 F_j^{-1}(s)F_i^{-1}(s)ds+\int_0^1 (F_i^{-1}(s))^2 ds,\end{aligned}$$

which reduces to (20). Then, using the definition of $w_i=\frac{\bar{w}_i}{\sum_{j=1}^n \bar{w}_j}$, (32) and (20) we deduce (19). The existence of a solution to (19) can be obtained by a fixed point argument using e.g., the Brouwer fixed point theorem.

References

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

Bantwal, V.J., H.C. Kunreuther. 2000. A Cat bond premium puzzle? *J Psychol. Financ. Markets* **1** 76–91.

Bell, D.E. 1988. One switch utility functions and a measure of risk. *Management Sci.* **34**(12) 1416–1424.

Bell, D.E. 1995. Risk, return and utility. *Management Sci.* **41**(1) 23–30.

Bell, D.E., C. Fishburn. 2001. Strong one switch utility. *Management Sci.* **47**(4) 601–604.

Burnecki, K., G. Kukla. 2003. Pricing of zero coupon and coupon CAT bonds. *Applicationes Mathematicae* **30** 315–334.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, L. Montrucchio. 2011. Uncertainty averse preferences. *J. Econ. Theory* **146**(4) 1275–1330.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, A. Rustichini. 2015. The structure of variational preferences. *J. Math. Econ.* **57** 12–19.

Cummins, J.D., P. Barrieu. 2013. *Innovations in Insurance Markets: Hybrid and Securitized Risk Transfer Solutions* Ch. 20 in G. Dionne (ed), Handbook of insurance, Springer.

Dasgupta, P. 2008. Discounting climate change. *J. Risk Uncertainty* **37(2-3)** 141–169.

Edesess, M. 2015. Catastrophe bonds: An important new financial instrument. *Alternative Investment Analyst Review* **Fall 2015**, 6–11.

Froot, K.A. 2007. Risk management, capital budgeting and capital structure policy for insurers and reinsurers. *J. Risk Ins.* **74**, 273–299.

Froot, K.A., P. O'Connell. 2008. On the pricing of intermediated risks: Theory and application to catastrophe reinsurance. *J Banking Financ.* **32** 69–85.

Galeotti, M., M. Gürtler, C. Winkelvos. 2013. Accuracy of premium calculation models for CAT bonds: An empirical analysis. *J. Risk Ins.* **80**(2) 401–421.

Gierlinger, J., C. Gollier. 2009. Socially efficient discounting under ambiguity aversion, IDEI Working paper 561, Toulouse, France.

Gierlinger, J., C. Gollier. 2017. Do interest rates decline when there is ambiguity about growth?, Working Paper.

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

Gollier, C. 2002. Discounting an uncertain future. *J. Public Econ.* **85** 149–166.

Gollier, C. 2013. *Pricing the Planet's Future: The Economics of Discounting in an Uncertain World*. Princeton University Press, Princeton, NJ.

Hansen, L. P., T. J. Sargent. 2001. Robust control and model uncertainty. *Am. Econ. Rev.* **91**(2) 60–66.

Hansen, L., T. Sargent, G. Turmuhambetova, G. N. Williams. 2006. Robust control and model misspecification. *J. Econ. Theory* **128**, 45–90.

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

Heal, G., A. Millner. 2014. Uncertainty and decision making in climate change economics. *Rev. Environ. Econ. Policy* **8**(1) 120–137.

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

Knight, F. H. 1921. *Risk, Uncertainty and Profit*. AM Kelley, New York 1964.

Kravvaritis, D. C., A. N. Yannacopoulos. 2020. *Variational Methods in Nonlinear Analysis: with Applications in Optimization and Partial Differential Equations*. Walter de Gruyter, Berlin.

Kriegler, E., J. Hall, H. Helda, R. Dawson, H. J. Schellnhuber. 2009. Imprecise probability assessment of tipping points in the climate system. *P. Natl. Acad. Sci. USA* **106**(13) 5041–5046.

Lichtendahl Jr, K. C., Y. Grushka-Cockayne, R. L. Winkler. 2013. Is it better to average probabilities of quantiles? *Management Sci.* **59**(7) 1594–1611.

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

Marinacci, M. 2015. Model uncertainty. *J. of European Econ. Assoc.* **1**3(6) 998–1076.

Meinshausen, M., N. Meinshausen, W. Hare, S. C. Raper, K. Frieler, R. Knutti, D. J. Frame, M. R. Allen. 2009. Greenhouse-gas emission targets for limiting global warming to 2 ^{o}C. *Nature* **458**(7242) 1158–1163.

Nordhaus, W.D. 2007. A review of the Stern review on the economics of climate change. *J. Econ. Lit.* **45**(3) 686–702.

Papayiannis, G. I., A. N. Yannacopoulos. 2018. A learning algorithm for source aggregation. *Math. Method. App. Sci.* **41**(3) 1033–1039.

Rogelj, J., M. Meinshausen, R. Knutti. 2012. Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nat. Clim. Change* **2**(4) 248.

Santambrogio, F. 2015. *Optimal Transport for Applied Mathematicians*. Birkhäuser, NY.

Stern, N. 2007. *The Economics of Climate Change: The Stern Review*. Cambridge University Press, Cambridge, UK.

Strzalecki, T. 2011. Axiomatic foundations of multiplier preferences. *Econometrica* **79**(1) 47–73.

Villani, C. 2008. *Optimal Transport: Old and New*, Grundlehren der mathematischen Wissenschaften, Volume 338. Springer Science & Business Media, Berlin.

Wang, S. S. 2000. A class of distortion operators for financial and insurance risks. *J. Risk Insur.* **67** 15–36.

Weitzman, M. 2007. A review of the Stern review on the economics of climate change. *J. Econ. Lit.* **45**(3) 703–724.

Online Appendix: Proofs and review of uncertainty averse preferences

Appendix EC.1: The Wasserstein metric in the space of probability measures

The definition of the Fréchet mean in the space of probability measures on Ω, denoted by $\mathcal{P}(\Omega)$, requires the choice of an appropriate metric d for this space. A good candidate for this task is the Wasserstein metric. To minimize technicalities, we will focus here on the particular case where $\Omega = \mathbb{R}^d$, $d \geq 1$, endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$. For a general discussion of the Wasserstein metric for measures defined in more general probability spaces as well as its various applications, see for example Santambrogio (2015) or Villani (2008).

For any two probability measures Q_1, Q_2 on $\mathcal{B}(\mathbb{R}^d)$, we can define the p-Wasserstein distance between then as follows:

Consider a $2d$-dimensional random variable $Z = (Z_1, Z_2)$, where $Z_i \in \mathbb{R}^d$, $i = 1, 2$, distributed in terms of a probability measure Π on $\mathcal{B}(\mathbb{R}^{2d})$, with marginals Q_1 and Q_2, i.e., such that $\Pi(A \times \mathbb{R}^d) = Q_1(A)$ and $\Pi(\mathbb{R}^d \times A) = Q_2(A)$ for every $A \in \mathcal{B}(\mathbb{R}^d)$. Let us denote by $\mathcal{T}$ the set of all such measures on $\mathcal{B}(\mathbb{R}^{2d})$, called transportation plans. The p-Wasserstein distance between these two measures is defined as

$$W_p(Q_1, Q_2) = \left\{ \inf_{\Pi \in \mathcal{T}} \mathbb{E}_\Pi \left[|Z_1 - Z_2|^p \right] \right\}^{1/p}, \quad p \geq 1.$$

This defines a metric in the space of probability measures compatible with the weak* topology (i.e., the convergence in the duality with C_b, the space of bounded continuous functions; see Santambrogio 2015).[10] The space of probability measures $\mathcal{P}(\mathbb{R})$ endowed with this metric is denoted by $\mathcal{P}_p(\mathbb{R}^d)$.

In the case where $d = 1$, i.e., when we consider any two probability measures Q_1, Q_2 on $\mathbb{R}$, each one of which may serve as an alternative probabilistic model for the lottery X, in the sense that for any set $A \in \mathcal{B}(\mathbb{R})$, $Q_1(A)$ and $Q_2(A)$ can be interpreted as two alternative assessments of the probability $Pr(X \in A)$ for the random variable X, the p-Wasserstein distance admits a closed form representation in terms of

$$W_p(Q_1, Q_2) = \left\{ \int_0^1 (F_1^{-1}(s) - F_2^{-1})^p \right\}^{1/p}, \quad p \geq 1,$$

[10] This is the topology which is related to the weak* convergence in the space of probability measures. Recall that a sequence of probability measures $\{Q_n\}$ on $\mathbb{R}$ converges to a probability measure Q on $\mathbb{R}$ in the weak* sense if $\int_{\mathbb{R}} \phi(x) dQ_n(x) \to \int_{\mathbb{R}} \phi(x) dQ$, for any bounded continuous function $\phi : \mathbb{R} \to \mathbb{R}$. This is a very important notion of convergence because of its useful compactness properties as well as its connection with simulation and Monte-Carlo methods.

where F_i^{-1} are the quantile functions corresponding to the distribution functions characterizing the measures Q_i, $i=1,2$.

In this work we focus on the case $\Omega=\mathbb{R}$ and $p=2$, i.e., on $\mathcal{P}(\mathbb{R})$ metrized by the 2-Wasserstein metric, that allows the determination of the corresponding Fréchet mean utility functionals in closed form. However, these utilities are well-posed for any choice of $p>1$, as well as for general probability measures for vector valued lotteries or factors $Q\in\mathcal{P}(\mathbb{R}^d)$.

Appendix EC.2: Proofs of Statements: Remarks, comments and extensions

Notation To ease notation in this section we will denote all quantile functions by g, and by $\mathcal{Q}$ the set of quantile functions.

EC.2.1. Proof of the error bound (2)

Let Q_1 and Q_2 be the probability laws of agents 1 and 2. These are probability measures on $\mathcal{B}(\mathbb{R})$, expressed in terms of distribution functions F_i, $i=1,2$. Let Π be any transportation plan between these two measures, i.e., a probability measure on $\mathcal{B}(\mathbb{R}\times\mathbb{R})$ with marginals Q_1 and Q_2 or equivalently with the property $\Pi(A\times\mathbb{R})=Q_1(A)$ and $\Pi(\mathbb{R}\times A)=Q_2(A)$ for every $A\in\mathcal{B}(\mathbb{R})$. Assuming without loss of generality continuous distributions, a transportation plan Π can be expressed in terms of a joint density $f:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, with the property that

$$\int_{\mathbb{R}} f(x_1,x_2)dx_1 = f_2(x_2),\ \ \forall x_2\in\mathbb{R},$$
$$\int_{\mathbb{R}} f(x_1,x_2)dx_2 = f_1(x_1),\ \ \forall x_1\in\mathbb{R},$$

where f_i are the probability densities of the measures Q_i, $i=1,2$. We can easily see by applying Fubini that

$$\mathbb{E}_{Q_1}[u(X)] = \int_{\mathbb{R}} u(x_1)f_1(x_1)dx_1 = \int_{\mathbb{R}\times\mathbb{R}} u(x_1)f(x_1,x_2)dx_1\,dx_2,$$
$$\mathbb{E}_{Q_2}[u(X)] = \int_{\mathbb{R}} u(x_2)f_2(x_2)dx_2 = \int_{\mathbb{R}\times\mathbb{R}} u(x_2)f(x_1,x_2)dx_1\,dx_2,$$

so that

$$\begin{aligned}|U_1(X)-U_2(X)| &= \left|\int_{\mathbb{R}\times\mathbb{R}} (u(x_1)-u(x_2))f(x_1,x_2)dx_1\,dx_2\right| \\ &\le \int_{\mathbb{R}\times\mathbb{R}} |u(x_1)-u(x_2)|f(x_1,x_2)dx_1\,dx_2 \le C\int_{\mathbb{R}\times\mathbb{R}} |x_1-x_2|f(x_1,x_2)dx_1\,dx_2,\end{aligned}$$

where C is the Lipschitz constant of the utility function u. This implies that

$$\begin{aligned}|U_1(X)-U_2(X)|^2 &\le C^2\left(\int_{\mathbb{R}\times\mathbb{R}} |x_1-x_2|f(x_1,x_2)dx_1\,dx_2\right)^2 \\ &\le C^2\int_{\mathbb{R}\times\mathbb{R}} |x_1-x_2|^2 f(x_1,x_2)dx_1\,dx_2,\end{aligned}$$

where we use Jensen's inequality for the convex function $\phi(x) = x^2$. The above estimate is true for any density f corresponding to any transportation plan Π compatible with the marginals Q_1, Q_2. It therefore holds also for the infimum of this quantity over the set of such transportation plans $\mathcal{T}$,

$$|U_1(X) - U_2(X)|^2 \leq C^2 \inf_{f \in \mathcal{T}} \int_{\mathbb{R}\times\mathbb{R}} |x_1 - x_2|^2 f(x_1, x_2) dx_1\, dx_2 =: C^2\, W_2^2(Q_1, Q_2).$$

This completes the proof.

EC.2.2. Remarks on the proof of Theorem 1

Remark EC.1 (Second-order conditions). It is interesting to look at the second-order condition for the optimization problem for the functional J. We calculate the second derivative $D^2J(g)$ which is considered as an operator acting from $L^2([0,1])$ onto itself defined as

$$\langle D^2J(g)g_2, g_1\rangle = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\langle DJ(g+\epsilon g_2) - DJ(g), g_1\rangle,\ \ \forall\, g_1, g_2 \in L^2([0,1]).$$

A critical point g^* of J corresponds to a local minimum in $L^2([0,1])$ if

$$\langle D^2J(g^*)h, h\rangle \geq 0,\ \forall\, h \in L^2([0,1]). \tag{EC.1}$$

Using the stated smoothness and integrability conditions on u and ϕ we may calculate the second derivative, and in particular the quadratic form in (EC.1).

Since for any $g, h \in L^2([0,1])$,

$$\mathsf{F}_{\mathbb{M}}(g+\epsilon h) = \mathsf{F}_{\mathbb{M}}(g) + 2\epsilon\langle g - g_{\mathcal{B}}, h\rangle,$$

we have upon expanding that

$$\begin{aligned} DJ(g+\epsilon h) &= u'(g+\epsilon h) + \phi'(\mathsf{F}_{\mathbb{M}}(g+\epsilon h))(g+\epsilon h - g_{\mathcal{B}}) \\ &= DJ(g) + \epsilon\big(u''(g) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g))\big)h + 4\epsilon\phi''(\mathsf{F}_{\mathbb{M}}(g))\langle g - g_{\mathcal{B}}, h\rangle\,(g - g_{\mathcal{B}}) + O(\epsilon^2). \end{aligned}$$

Hence, carefully passing to the limit we obtain that

$$\begin{aligned} \langle D^2J(g)h, h\rangle &= \big\langle\big(u''(g) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g))\big)h, h\big\rangle + 4\phi''(\mathsf{F}_{\mathbb{M}}(g))(\langle g - g_{\mathcal{B}}, h\rangle)^2 \\ &\geq \langle\big(u''(g) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g))\big)h, h\rangle = \int_0^1 \big(u''(g(s)) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g))\big)h(s)^2 ds,\ \ \forall\, g, h \in L^2([0,1]), \end{aligned}$$

where we also used the convexity of ϕ.

The second-order condition for an internal minimum g^* then becomes

$$\langle D^2J(g^*)h, h\rangle \geq \int_0^1 \big(u''(g^*(s)) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g^*))\big)h(s)^2 ds \geq 0,\ \forall\, h \in L^2([0,1]),$$

which is satisfied if

$$u''(g^*(s)) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g^*)) \geq 0, \;\; a.e. \; s \in [0,1]. \tag{EC.2}$$

Note that g^* is the solution of the first-order condition

$$u'(g^*(s)) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g^*))(g^*(s) - g_B(s)) = 0, \;\; s \in [0,1]. \tag{EC.3}$$

Differentiating (EC.3) with respect to s we obtain that

$$\left(u''(g^*(s)) + 2\phi'(\mathsf{F}_{\mathbb{M}}(g^*))\right)\frac{dg^*}{ds}(s) = 2\phi'(\mathsf{F}_{\mathbb{M}}(g^*))\frac{dg_{\mathcal{B}}}{ds}(s) \geq 0,$$

since $g_{\mathcal{B}}$ is a quantile function, therefore, we observe that the second-order condition (EC.2) is satisfied as long as g^* corresponds to a quantile. In fact the second-order condition (EC.2) and the condition required for a solution of the first order condition (EC.3) to be a quantile coincide.

In condition (EC.2), the first term is negative for a typical utility function, hence the condition is satisfied as long as $\phi'(\mathsf{F}_{\mathbb{M}}(g^*)) \geq \sup_{s\in[0,1]} |u''(g^*(s))|$, i.e., if ϕ is steep enough. The latter is easily satisfied if suitable assumptions on $\mathbb{M}$ and u are imposed (see, e.g., Assumptions 1, 2 and the arguments in the proof of Theorem 2). For Fréchet-Wasserstein multiplier preferences the second-order condition simplifies to $\theta > \sup_{s\in[0,1]} |u''(g^*(s))|$, which is satisfied by the choice of $\theta > \theta_c$ (see proof of Theorem 2).

Remark EC.2. (a) The right hand side of the auxiliary equation

$$u'(x) + \rho x = \rho z, \;\; \rho > 0, \tag{EC.4}$$

is considered as varying, since it depends on $g_{\mathcal{B}}(s)$. Observe that if $\rho \to \infty$, then $\Psi_\rho(z) \simeq z$, hence the solution of (26) is expected to be close to $g_{\mathcal{B}}$, i.e. $g^*(s) \simeq g_{\mathcal{B}}(s)$. For smaller values of ρ, the solution g^* will be a distorted quantile, the distortion coming from the effects of the $\frac{1}{2\rho}u'$ term.

(b) The existence or not of a solution to (26) for any $s \in [0,1]$ depends on the value of $\inf_{x\in\mathbb{R}} v_\rho(x) =: \Gamma(\rho)$. Adopting this notation it is clear that if $z < \Gamma(\rho)$, then no solution for (EC.4) exists, therefore, if s is such that $g_{\mathcal{B}}(s) < \Gamma(\rho)$, then $u'(g^*(s)) + 2\phi'(A(g^*))(g^*(s) - g_{\mathcal{B}}(s)) > 0$. Since $g_{\mathcal{B}}$ is an increasing function, we conclude that as long as $s \in [0, s^*)$ where $g_{\mathcal{B}}(s^*) = \Gamma(\rho)$, then $u'(g^*(s)) + 2\phi'(A(g^*))(g^*(s) - g_{\mathcal{B}}(s)) > 0$ for every $g^* \in \mathcal{Q}$, and an interior solution is not possible. The conditions under which g^* is a quantile depend on the choice of penalty function. For conditions under which g^* is a quantile in the special case where $\phi(x) = \frac{\theta}{2}(x - V_{\mathbb{M}})$, see Theorem 2 and its proof in Section EC.2.3.

(c) If $u'(x) \to \infty$ as $x \to 0$, then more than one solution of (26) can exist. Care has to be taken to choose the appropriate one, i.e., the one that provides the smallest value for the function to be minimized. That may depend on both the choice of u and of the penalty function.

EC.2.3. Full proof of Theorem 2

(a) We need to introduce the following quantities:

$$K_B = \sum_{i=1}^{n} w_i K_i,$$
$$K^* = \min_{i=1,\cdots,n} K_i.$$

Furthermore, let $g^*(s) = g_\theta(s)$ be the solution of the parametrized algebraic equation

$$\frac{1}{\theta} u'(g^*(s)) + g^*(s) = g_{\mathcal{B}}(s), \;\; s \in [0,1],$$

and define the function

$$V_\theta(x) := u(x)\left(1 - \frac{1}{2\theta}\frac{(u')^2}{u(x)} + \frac{1}{2\theta^2}\frac{(u')^2 u''(x)}{u(x)}\right) + O\left(\frac{1}{\theta^3}\right).$$

We now proceed with the proof.

In this case $\phi'(x) = \frac{\theta}{2}$, and the variational inequality (25) simplifies to

$$\int_0^1 (u'(g^*(s)) + \theta(g^*(s) - g_{\mathcal{B}}(s))(g^*(s) - \bar{g}(s))ds \leq 0, \;\; \forall \bar{g} \in \mathcal{Q}. \tag{EC.5}$$

Under Assumption 2 it is reasonable to define the set of admissible quantile functions as

$$\mathcal{Q}_{K^*} = \{g \in L^2([0,1]) \; : \; g \text{ (right) continuous increasing } g \geq K^* := \min_{i=1,\cdots,n} K_i\},$$

which is a convex set, and replace $\mathcal{Q}$ by $\mathcal{Q}_{K^*}$ in (EC.5). For simplicity we restrict attention to continuous quantiles. An interior solution for (EC.5) is a solution of the parametric algebraic equation

$$\frac{1}{\theta} u'(g^*(s)) + g^*(s) = g_{\mathcal{B}}(s), \;\; g^*(s) \geq K^* \;\; a.e.\; s \in [0,1]. \tag{EC.6}$$

Equation (EC.6) clarifies the need to introduce Assumption 2. If $g_{\mathcal{B}}(s) \to 0$ as $s \to 0$, then (EC.6) is impossible as its left hand side is always strictly greater than 0. Therefore, in order to guarantee internal solutions, we need to make sure that $g_{\mathcal{B}}$ does not touch 0 as $s \to 0$, and this is guaranteed by Assumption 2. By the same argument, we cannot allow g^* to touch 0 (unless we drop the assumption that $u'(0) = \infty$) and this leads to the introduction of the set of admissible quantile functions $\mathcal{Q}_{K^*}$.

Moreover, the parameter $\theta > 0$ plays a role as well in the above considerations. Clearly, as $\theta \to \infty$, the first term becomes insignificant and $g^*(s) \to g_{\mathcal{B}}(s)$ for every $s \in [0,1]$. However, for very small values of θ, i.e., for $\theta \to 0$, the left hand side will tend to infinity for any $s \in [0,1]$, pushing $g^*(s)$ to 0 and the equation (EC.6) will be impossible. Hence, a lower value for θ must be selected, i.e., a θ_c such that for $\theta > \theta_c$ one may find admissible internal solutions. This choice should be such as to keep a balance between the two terms on the left hand side of (EC.6).

In order to estimate the critical θ_c, we need to ensure two conditions: (i) that (EC.6) admits a solution $g^*(s)$ for every choice of right hand side $g_{\mathcal{B}}(s) \geq K_{\mathcal{B}}$, and (ii) the resulting function $g^*(s) \in \mathcal{Q}_{K^*}$.

We start with (ii). Assuming a solution exists upon differentiating (EC.6) with respect to s, we have that

$$\left(\frac{1}{\theta}u''(g(s))+1\right)\frac{dg}{ds}(s) = \frac{dg_{\mathcal{B}}}{ds}(s),$$

and since $g_{\mathcal{B}}$ is increasing, g will be increasing as long as $\frac{1}{\theta}u''(g(s))+1 \geq 0$ which, since $u'' \leq 0$, reduces to

$$|u''(g(s))| \leq \theta. \tag{EC.7}$$

Since $x \mapsto |u''(x)|$ is decreasing and $g(s) > K^*$, we have that $|u''(g(s))| < |u''(K^\star)|$; hence (EC.7) holds as long as

$$|u''(K^*)| \leq \theta, \tag{EC.8}$$

is satisfied.

We now consider (i). We want to choose θ so that the equation $\psi(x) := \frac{1}{\theta}u'(x) + x = z$ has a solution $x > K^*$ for any $z > K_{\mathcal{B}}$. This will be true if

$$\min_{y \geq K^*} \psi(y) = \min_{y \geq K^*} \left(\frac{1}{\theta}u'(y) + y\right) < K_{\mathcal{B}}. \tag{EC.9}$$

For the range of values for θ given by condition (EC.8), the function $x \mapsto \psi(x)$ is increasing, hence $\min_{y \geq K^*} \psi(y) = \psi(K^*)$. Therefore (EC.9) becomes

$$\frac{1}{\theta}u'(K^*) + K^* < K_{\mathcal{B}},$$

which upon rearrangement yields

$$\theta > \frac{u'(K^*)}{K_{\mathcal{B}} - K^*}. \tag{EC.10}$$

Combining (EC.8) with (EC.10), we can characterize the critical θ as

$$\theta_c = \max\left\{|u''(K^*)|, \frac{u'(K^*)}{K_{\mathcal{B}} - K^*}\right\}. \tag{EC.11}$$

Expression (7) follows by substitution of $g^* = g_\theta$ in the functional to be minimized and by simple algebraic manipulation of the penalty term.

For the comparison of $\mathcal{U}_\theta(X)$ and $\mathcal{U}_\beta(X) = \mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$, note that

$$\mathcal{U}_\theta(X) = \min_{Q \in \mathcal{P}_2(\mathbb{R})} \left[\mathbb{E}_Q[u(X)] + \frac{\theta}{2}\big(\mathsf{F}_{\mathbb{M}}(Q) - \mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}})\big)\right] \leq \mathbb{E}_{Q_{\mathcal{B}}}[u(X)] + \frac{\theta}{2}\big(\mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}}) - \mathsf{F}_{\mathbb{M}}(Q_{\mathcal{B}})\big) = \mathcal{U}_{\mathcal{B}}(X).$$

Moreover, one may prove monotonicity properties of $\mathcal{U}_\theta$ with respect to θ.

For the sharper upper bound on $g^* = g_\theta$, rewrite the algebraic equation as

$$g_\theta(s) = g_{\mathcal{B}}(s) - \frac{1}{\theta}u'(g_\theta(s)) \leq g_{\mathcal{B}}(s) - \frac{1}{\theta}u'(g_{\mathcal{B}}(s)),$$

where we used the estimate $g_\theta \leq g_{\mathcal{B}}$ and the fact that u' is a decreasing function.

(b) We now consider the expansion of the utility functional $\mathcal{U}$ for large θ. Using the notation $\epsilon = \frac{1}{\theta}$, we look for a solution of (EC.6) in the form $g^* = g_0 + \epsilon g_1 + \epsilon^2 g_2 + O(\epsilon^3)$. Substituting into (EC.6) and assuming smoothness of u, by Taylor expansion and matching orders of ϵ, we obtain the approximate solution

$$g^* = g_{\mathcal{B}} - \epsilon u'(g_{\mathcal{B}}) + \epsilon^2 u''(g_{\mathcal{B}})u'(g_{\mathcal{B}}) + O(\epsilon^3).$$

Substituting into (7), after some Taylor expansions we obtain the expansion for $\mathcal{U}_\theta$ in terms of $\epsilon = \frac{1}{\theta}$, as

$$\mathcal{U}_\theta(X) = \int_0^1 u(g_{\mathcal{B}}(s))ds - \frac{\epsilon}{2}\int_0^1 (u'(g_{\mathcal{B}}(s)))^2 ds + \frac{\epsilon^2}{2}\int_0^1 (u'(g_{\mathcal{B}}(s)))^2 u''(g_{\mathcal{B}}(s))ds + O(\epsilon^3),$$

reconfirming the fact that $g^* \leq g_{\mathcal{B}}$ and that $\mathcal{U}_\theta(X) \leq \int_0^1 u(g_{\mathcal{B}}(s))ds = \mathbb{E}_{Q_{\mathcal{B}}}[u(X)]$. □

Remark EC.3 (Error estimates for the expansion). The validity of the perturbative expansion relies on the dependence of g on $\epsilon := \frac{1}{\theta}$, and in particular on the smoothness of g as a function of the parameter ϵ, denoted for emphasis as $g(s;\epsilon)$. As the above expansion is a Taylor expansion of the function $\epsilon \mapsto g(s;\epsilon)$ around the point $\epsilon = 0$, with the function g defined implicitly in terms of

$$\epsilon u'(g(s;\epsilon)) + g(s,\epsilon) = g_{\mathcal{B}}(s), \tag{EC.12}$$

the validity of the expansion depends on the smoothness of the function g with respect to ϵ, which in turn, by (EC.12), depends on the smoothness of the utility function $x \mapsto u(x)$. Indeed, a straightforward differentiation of (EC.12) shows that, as long as $u \in C^2$, we have $\frac{\partial g}{\partial \epsilon} = -\frac{u'(g)}{1+\epsilon u''(g)}$, so that iterating the argument, we have that in general g has one derivative less than u. This observation indicates that in order to be able to perform the expansion up to order n in ϵ, we need a utility function which is C^{n+1}. Standard results related to the error induced by the Taylor approximation can be used to show that upon defining

$$R_1(\epsilon) := \sup_{s\in[0,1]} |g_\theta(s) - g_{\mathcal{B}}(s) + \epsilon u'(g_{\mathcal{B}}(s)|$$
$$R_2(\epsilon) := \sup_{s\in[0,1]} |g_\theta(s) - g_{\mathcal{B}}(s) + \epsilon u'(g_{\mathcal{B}}(s) - \epsilon^2 u''(g_{\mathcal{B}}(s))u'(g_B(s))|,$$

the error in approximation of g_θ by keeping the first- or the second-order terms with respect to ϵ in the expansion respectively is

$$R_n(\epsilon) \le \frac{C_n}{(n+1)!}\epsilon^{n+1}, \text{ for } \epsilon \le \epsilon_0,\ n = 1, 2,$$

as long as

$$\sup_{s\in[0,1]} \left| \frac{\partial^n g(s;\epsilon)}{\partial \epsilon^n} \right| \le C_n, \text{ for } \epsilon < \epsilon_0 \tag{EC.13}$$

where in fact the error bound holds for any admissible order n in the expansion, higher than 1 or 2. Condition (EC.13) depends on the choice of utility function u as well as the support of the random variable X or the nature of $g_{\mathcal{B}}$. For example, to consider the error bound for the first-order approximation, we need to find some estimate for $\frac{\partial^2 g(s;\epsilon)}{\partial \epsilon^2}$. After a simple calculation we see that

$$\begin{aligned}
\frac{\partial}{\partial \epsilon} g(s;\epsilon) &= -\frac{u'(g(s;\epsilon))}{1+\epsilon u''(g(s;\epsilon))}, \\
\frac{\partial^2}{\partial \epsilon^2} g(s;\epsilon) &= -\frac{\left(2u''(g(s;\epsilon))\frac{\partial}{\partial \epsilon}g(s;\epsilon) + \epsilon u'''(g(s;\epsilon)\left(\frac{\partial^2}{\partial \epsilon^2}g(s;\epsilon)\right)^2\right)}{1+\epsilon u''(g(s;\epsilon))},
\end{aligned}$$

so upon specification of u, estimates for the derivatives can be obtained.

Remark EC.4 (Non-interior solutions). In the case where $\theta < \theta_c$, when (EC.6) does not admit a solution, then construction of the utility functional is slightly more complicated and requires the solution of the variational inequality (EC.5), which can be done using numerical techniques or a suitable perturbation expansion.

Remark EC.5 (Relaxing Assumptions 1 and 2). If we relax Assumption 1 so that the marginal utility is finite at 0, then we are allowed to set $K^* = 0$, removing the restriction that g is strictly positive. This also allows us to relax Assumption 2 on the set of models $\mathbb{M}$.

EC.2.4. Marginal utility in the CRRA family

Assuming now that we operate within the CRRA family, $u(x) = \frac{1}{1-\gamma_0}x^{1-\gamma_0}$, we can calculate the expansion for $\mathcal{M}(X)$ in the form

$$\begin{aligned}
\mathcal{M}_\theta(X) = \mathcal{M}_{\mathcal{B}}(X,\gamma_0) &+ \frac{\gamma_0}{\theta}\mathcal{M}_{\mathcal{B}}(X, 1+2\gamma_0) \\
&+ \frac{\gamma_0(3\gamma_0+1)}{2\theta^2}\mathcal{M}_{\mathcal{B}}(X, 2+3\gamma_0),
\end{aligned} \tag{EC.14}$$

where

$$\mathcal{M}_{\mathcal{B}}(X,\gamma) := \mathbb{E}_{Q_{\mathcal{B}}}[u'(X)] = \mathbb{E}_{Q_{\mathcal{B}}}[(X)^{-\gamma}]$$

is the marginal utility of X for an agent whose preferences are characterized by an expected utility function, under the probability measure which is related to the Wasserstein barycenter $Q_{\mathcal{B}}$ of the model set $\mathbb{M}$, with CRRA utility with risk aversion coefficient γ. This leads to an interesting interpetation of the marginal utility for an uncertainty averse agent, as a linear combination of $\mathcal{M}_{\mathcal{B}}(X,\gamma)$ with increasing values of risk aversion coefficients (i.e., γ_0, $1+2\gamma_0$ and $2+3\gamma_0$) with these terms becoming increasingly important as θ decreases. This interpretation provides a clear link between risk aversion and uncertainty aversion. A similar expression and interpretation holds for the exponential family.

EC.2.5. Analytic approximations for the social discount rate (SDR)

EC.2.5.1. A general expansion for the SDR In particular the following expansion holds:

$$\begin{aligned} r(t) = r_{\mathcal{B}}(t) &- \frac{1}{\theta}\frac{1}{tu'(C(0))}\mathbb{E}_{Q_{t,\mathcal{B}}}[u'''(C(t))] \\ &- \frac{1}{\theta^2}\frac{1}{tu'(C(0))}\mathbb{E}_{Q_{t,\mathcal{B}}}\left[[u'(C(t))\left[(u''(C(t)))^2+\frac{1}{2}u'''(C(t))u'(C(t))\right]\right]+O\left(\frac{1}{\theta^3}\right). \end{aligned} \tag{EC.15}$$

We can see that, for utility functions for which $u'''>0$ (e.g., members of the exponential and the CRRA family), relation (16) is also satisfied by the expansion (EC.15). Since the term $u'(C(0))$ simply contributes a scaling factor to the formula, without loss of generality we set $C(0)=1$ as the numeraire for consumption to simplify the notation.

EC.2.5.2. The SDR with CRRA utility Assuming now that we operate within the CRRA family $u(x)=\frac{1}{1-\gamma_0}x^{1-\gamma_0}$, upon substitution into the Ramsey formula under uncertainty (15) where

$$\mathcal{M}_{\mathcal{B}}(X,\gamma_0) := \mathbb{E}_{Q_{\mathcal{B}}}[u'(X)] = \mathbb{E}_{Q_{\mathcal{B}}}[(X)^{-\gamma_0}],$$

we obtain an expansion for the SDR for large enough θ in terms of

$$r(t)=\delta-\frac{1}{t}\ln\left(\mathcal{M}_{\mathcal{B}}(C(t),\gamma_0)+\frac{\gamma_0}{\theta}\mathcal{M}_{\mathcal{B}}(C(t),1+2\gamma_0)+\frac{\gamma_0(3\gamma_0+1)}{2\theta^2}\mathcal{M}_{\mathcal{B}}(C(t),2+3\gamma_0)\right),$$

which can further be approximated as

$$\begin{aligned} r(t) = r_{\mathcal{B}}(t) &- \frac{1}{t}\frac{\gamma_0}{\theta}\frac{\mathcal{M}_{\mathcal{B}}(C(t),1+2\gamma_0)}{\mathcal{M}_{\mathcal{B}}(C(t),\gamma_0)} \\ &- \frac{1}{t}\frac{\gamma_0}{\theta^2}\left\{(3\gamma_0+1)\frac{\mathcal{M}_{\mathcal{B}}(C(t),2+3\gamma_0)}{\mathcal{M}_{\mathcal{B}}(C(t),\gamma_0)}-\gamma_0\left(\frac{\mathcal{M}_{\mathcal{B}}(C(t),1+2\gamma_0)}{\mathcal{M}_{\mathcal{B}}(C(t),\gamma_0)}\right)^2\right\}+O\left(\frac{1}{\theta^3}\right). \end{aligned}$$

Here $r_{\mathcal{B}}(t)$ is the SDR under the probability measure corresponding to the Wasserstein barycenter of the set of models $\mathbb{M}_t$. The remaining terms are corrections due to uncertainty aversion, which are expressed in terms of nonlinear functions of the expected marginal utilities under $Q_{\mathcal{B}}$. These terms correspond to CRRA utilities but with a higher risk aversion coefficient. Hence, the effect of uncertainty aversion, at least to first order in $\frac{1}{\theta}$, is to decrease the SDR as compared to the barycentric one. The relevant comparison for general θ follows from Proposition 2.

We further consider two specific examples for the probability models concerning the random variable $C(t)$ which are considered in reference set $\mathbb{M}$, and the corresponding Wasserstein barycenter model.

EC.2.5.3. Location scale families: Under the additional assumption that all agents have models according to which $C(t)$ follows a location scale family, each one with different parameters $(\mu_{t,i}, \sigma_{t,i})$, i.e., $F_{t,i}^{-1}(s) = \mu_{t,i} + \sigma_{t,i} F^{-1}(s)$, the Wasserstein barycenter for M_t is of the form $F_{t,B}^{-1}(s) = \mu_{t,B} + \sigma_{t,B} F^{-1}(s)$, where $\mu_{t,B} = \sum_{i=1}^n w_i \mu_{t,i}$ and $\sigma_{t,i} = \sum_{i=1}^n w_i \sigma_{t,i}$ so that

$$\mathcal{M}_{\mathcal{B}}(C(t), \gamma) = \mathbb{E}_{Q_{t,\mathcal{B}}}[(C(t))^{-\gamma}] = \int_0^1 (\mu_{t,B} + \sigma_{t,B} F^{-1}(s))^{-\gamma} ds.$$

The calculation of the SDR $r(t)$ is then reduced to the calculation of a class of parametric integrals of the form $\int_0^1 (\mu + \sigma F^{-1}(s))^{-\gamma} ds$ which, if not possible analytically in closed form, can always be performed numerically, either using quadrature techniques or Monte-Carlo integration. The second choice is very easy to implement, since knowledge of the quantile allows for creation of the relative sample on which the integral can easily be estimated, using the method of inversion. This integral is well-posed as long as $\inf_{s \in [0,1]} \mu_{t,B} + \sigma_{t,B} F^{-1}(s) \geq K_* > 0$, which is always true if Assumption 2 holds and at least one of the $K_i > 0$.

EC.2.5.4. Log normal consumption distribution In the case where consumption follows the log normal distribution, that is for any $t > 0$ under the Wasserstein barycenter measure $Q_{t,\mathcal{B}}$, $\ln C(t) \sim N(\mu_{t,B}, \sigma_{t,B}^2)$ where $\mu_{t,B}$ and $\sigma_{t,B}^2$ depend on t, an expression for the uncertainty averse SDR can be obtained in closed form.

In order to perform the calculation for the quantities $\mathcal{M}_{\mathcal{B}}(C(t), \gamma)$, we express

$$\mathcal{M}_{\mathcal{B}}(C(t), \gamma) = \mathbb{E}_{Q_{t,\mathcal{B}}}[(C(t))^{-\gamma}] = \mathbb{E}_{Q_{t,\mathcal{B}}}[\exp(-\gamma \ln C(t))],$$

and using the formula for the exponential moments of the normal distribution we find, after some algebra, that

$$r(t) = r_{\mathcal{B}}(t) - \frac{\gamma_0}{\theta} \frac{1}{t} \exp(A_1(t)) - \frac{\gamma_0}{\theta^2} \frac{1}{t} \{(3\gamma_0 + 1) \exp(A_2(t)) - \gamma_0 \exp(2(A_1(t))\} + O\left(\frac{1}{\theta^3}\right), \quad \text{(EC.16)}$$

where

$$r_{\mathcal{B}}(t) = \delta + \gamma_0 \frac{1}{t} \mu_{t,B} - \frac{\gamma_0^2}{2} \frac{1}{t} \sigma_{t,B}^2,$$

and $A_1 = a_1 - a_0$, $A_2 = a_2 - a_0$ with

$$a_i(t) = -\gamma_i \left(\mu_{t,B} - \frac{1}{2} \gamma_i \sigma_{t,B}^2 \right), \quad i = 0, 1, 2.$$

EC.2.5.5. Numerical estimation of the SDR The SDR can be calculated numerically using the following step-by-step algorithmic procedure.

Algorithm [Calculation of $r(t)$]

A. Suppose that we have n different models concerning the distribution of $C(t)$. Thus, for each $t \in \mathbb{R}_+$ we have a set of $\mathbb{M}_t = \{Q_{t,i},\ i = 1, \cdots, n\}$ of n distinct probability measures $Q_{t,i}$ that can model the random variable $C(t)$. These measures can be characterized through the distribution functions $F_{t,i}$ or their inverses, the quantile functions $g_{t,i}$.

B. If each of the models is assigned a weight w_i, the Wasserstein barycenter is of the form

$$g_{t,\mathcal{B}}(s) = \sum_{i=1}^{N} w_i g_{t,i}(s),\ \ s \in (0,1).$$

C. If a Fréchet means utility functional is used to evaluate $C(t)$ and the corresponding marginal utility, then we can calculate the minimizing quantile function by solving for each $s \in [0,1]$ the algebraic equation

$$u'(x) + \theta x - \theta g_{t,\mathcal{B}}(s) = 0,$$

and then set $g_t^*(s) = x$. This step is very easy to handle numerically, as it is an algebraic equation for a single variable, and in the cases of CRRA utilities with $\gamma = 1, 2$ can be obtained analytically (see previous section).

D. Having obtained the minimizing quantile g_t^*, the marginal utility $\mathcal{M}(C(t))$ is calculated by calculating the integral

$$\mathcal{M}(C(t)) = \int_0^1 u'(g_t^*(s))ds.$$

This can be performed numerically, using a quadrature procedure.

E. The SDR can then be calculated by

$$r(t) = \delta - \frac{1}{t} \ln \frac{\mathcal{M}(C(t))}{u'(C(0))}.$$

Appendix EC.3: A brief review of uncertainty averse preferences

Cerreia-Vioglio et al. (2011) introduced a general class of preferences called uncertainty averse preferences. These preferences take ambiguity into consideration and can be considered as generalizations of certain important paradigms in decision making under uncertainty in the sense that they include as special cases the Maccheroni-Marinacci-Rustichini variational preferences model (Maccheroni et al. 2006) and because of that the Gilboa-Schmeidler minimax utility model (Gilboa and Schmeidler 1989), the Hansen and Sargent multiplier preferences model (see Hansen and Sargent 2001).

The uncertainty averse preferences require the introduction of the standard Anscombe-Aumann set-up in decision theory, in which $\mathfrak{F}$ is a set of uncertain acts $f : \mathcal{S} \to \mathbb{X}$, where $\mathcal{S}$ is a state space and $\mathbb{X}$ is a convex outcome space. By Δ we denote the set of all probability measures on $\mathcal{S}$. As is standard in this theory, the preferences $\succeq$ are on the acts. According to Cerreia-Vioglio et al. (2011), the preference relation $\succeq$ is uncertainty averse if and only if there exists a utility function $u : \mathbb{X} \to \mathbb{R}$ and a quasi-convex function $G : u(\mathbb{X}) \times \Delta \to (-\infty, \infty]$, increasing in the first variable such that the preference functional

$$U(f) := \min_{P \in \Delta} G\bigg(\int u(f) dP, P \bigg), \ \forall f \in \mathfrak{F} \tag{EC.17}$$

represents the preference relation. The outcome space can be considered as a lottery space (i.e., a space of random variables), and the meaning of an act is that, depending on the state of the world s, the DM faces a different lottery (which is in fact a random variable with a given probability distribution).[11]

Depending on the choice of G we may recover all the aforementioned preference models. For example, if $G(t, P) = t + \delta_{P_0}(P)$, where δ_{P_0} is the Dirac measure on Δ (i.e., a delta measure on a space of measures), then $U(f) = \int u(f) dP_0 = \mathbb{E}_{P_0}[u(f)]$ is a subjective expected utility function, over the subjective probability measure P_0. If, on the other hand, $G(t, P) = t + c(P)$ where $c : \Delta \to \mathbb{R}$ is a convex function, then the preference functional becomes the variational preferences functional $U(f) = \min_{P \in \Delta}[\int u(f) dP + c(P)]$. Depending on the choice for c, we obtain a variety of preferences:

(i) If $c(P) = \mathbb{I}_C(P)$ where $C \subset \Delta$ is closed and convex, and $\mathbb{I}_C$ is the indicator function in the sense of convex analysis, then we recover the Gilboa-Schmeidler minmax utility $U(f) = \min_{P \in C}[\int u(f) dP]$. In the particular case where $C = \{P_0\}$, i.e., where the set C consists of a single element, we recover the subjective expected utility case.

(ii) If $c(P) = \theta\, KL(P||P_0)$ where $KL(\cdot||P_0)$ is the Kullback-Leibler divergence from a given reference measure, the $U(f) = \min_{P \in \Delta}[\int u(f) dP + \theta KL(P||P_0)]$, the multiplier preferences introduced by Hansen and Sargent.

(iii) If $c(P) = \mathbb{I}_C(P)$ where $C = \{P \in \Delta : KL(P||P_0) < H_0\}$, then we obtain the constraint preferences of Hansen and Sargent. It is well-known (through a simple application of Lagrange multiplier theory) that for particular values of θ the corresponding variational preferences coincide with the constraint preferences.

If on the other hand we use non-separable forms for G, i.e., if $G(t, P) = t + \min_{Q \in \Gamma(P)} I_t(Q||Q_0)$ where $I_t(\cdot\, ||Q_0)$ is some statistical distance function that generalizes relative entropy and $\Gamma(P)$ is a suitable set of second-order probabilities, then the model (EC.17) reduces to the smooth preferences model of Klibanoff, Marinacci and Mukerjii where $U(f) = \phi^{-1}(\int_\Delta \phi(\int_S u(f(s))) dP(s)) d\mu(P))$,

[11] As a simple example consider the Elsberg-Schmeidler thought experiment in which "nature" chooses a coin $s \in \mathcal{S}$ which has probability $(a, 1-a)$ on the heads and tails respectively. The action is a lottery on the heads and the tails, i.e., a random variable with a probability distribution depending on s. For example, betting 1 euro on heads corresponds to the lottery $f = \mathbf{1}_H$ and $f(s)$ will then yield 1 euro with probability a and 0 euros with probability $1-a$. More complicated lotteries $f(s)$ can also be selected.

where ϕ is a continuous and strictly increasing function and μ is a probability measure on Δ. Smooth preferences are not necessarily distinct from variational preferences; as shown in Cerreia-Vioglio et al. (2011), if ϕ is a function corresponding to constant absolute risk aversion[12] then the smooth preferences reduce to a variational preference.

In closing this short review, we wish to mention that uncertainty averse preferences satisfy a well-defined set of axioms. The first set of axioms which allows us to characterize a preference relation as uncertainty averse are standard rationality axioms (weak order and monotonicity) along with a convexity axiom according to which $f_1 \sim f_2$ implies that $\alpha f_1 + (1-\alpha) f_2 \geq f_1$ for any $\alpha \in (0,1)$. The remaining axioms on the preference relation are somewhat technical, such as for instance continuity assumptions or a weak version of the risk independence axiom which holds for the constant acts only, i.e., for any triple of constant acts[13] $x_1, x_2, x_3 \in \mathbb{X}$, if $x_1 \sim x_2$ then $\alpha x_1 + (1-\alpha) x_3 \sim \alpha x_1 + (1-\alpha) x_3$ for any $\alpha \in (0,1)$. Further refinement of the technical axioms yield more refined results on the mapping G. For details, we refer to Cerreia-Vioglio et al. (2011).

Let $\mathcal{S}$ be a state space modelling the various states of nature, and consider mappings $f : \mathcal{S} \to \mathbb{X}$ which model the outcome of an action effected by the agent which is dependent on the state of the world; $f(s)$ is the effect of action f if the state of the world $s \in \mathcal{S}$ materializes. The constant act $f(s) = x$ for all $s \in \mathcal{S}$ will be simply denoted by x. The agent is equipped with a preference relation on the set of acts, $\succeq$. Following Maccheroni et al. (2006) we assume that the preference relation satisfies the following properties.

Assumption EC.1 **(Properties of $\succeq$)**. *The preference relation $\succeq$ satisfies the following properties:*

A.1 Weak order, i.e., $\succeq$ is complete and transitive.

A.2 Weak certainty independence, i.e., $\lambda f_1 + (1-\lambda) x \succeq \lambda f_2 + (1-\lambda) x$ implies that $\lambda f_1 + (1-\lambda) y \succeq \lambda f_2 + (1-\lambda) y$, for any acts f_1, f_2, any constant acts x, y and $\lambda \in (0,1)$.

A.3 Continuity, i.e., the sets $\{\lambda \in [0,1] : \lambda f_1 + (1-\lambda) f_2 \succeq f_3\}$ and $\{\lambda \in [0,1] : f_3 \succeq \lambda f_1 + (1-\lambda) f_2\}$ are closed sets for any acts f_1, f_2, f_3.

A.4 Monotonicity, i.e., if $f_1(s) \succeq f_2(s)$ for any $s \in \mathcal{S}$ then $f_1 \succeq f_2$.

A.5 Uncertainty aversion, i.e., $f_1 \sim f_2$ implies that $\lambda f_1 + (1-\lambda) f_2 \succeq f_1$ for any acts f_1, f_2 and any $\lambda \in (0,1)$.

A.6 Nondegeneracy, i.e., $f_1 \succ f_2$ for some acts f_1, f_2.

Definition EC.1 (Variational preferences, Maccheroni et al. (2006)). A preference relation $\succeq$ on the space of acts is called variational if it satisfies the axioms of Assumption EC.1.

[12] That is, $\phi(t) = -a \exp(-\theta t) + b$ or $\phi(t) = at = b$, with $a, \theta > 0$ and $b \in \mathbb{R}$.

[13] That is, acts such that $x(s) = x$ for every $s \in \mathcal{S}$.