Working Papers

Institute of Mathematical Economics
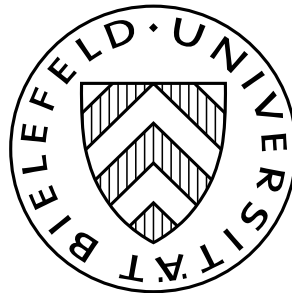
433

Mai 2010

Merging of Opinions under Uncertainty

Monika Bier and Daniel Engelage

Merging of Opinions under Uncertainty

Monika Bier*, Daniel Engelage[†‡]

May 27, 2010

Abstract

We consider long-run behavior of agents assessing risk in terms
of dynamic convex risk measures or, equivalently, utility in terms of
dynamic variational preferences in an uncertain setting. By virtue of
a robust representation, we show that all uncertainty is revealed in the
limit and agents behave as expected utility maximizer under the true
underlying distribution regardless of their initial risk anticipation. In
particular, risk assessments of distinct agents converge. This result
is a generalization of the fundamental Blackwell-Dubins Theorem, cp.
[Blackwell & Dubins, 62], to convex risk. We furthermore show the
result to hold in a *non*-time-consistent environment.

Keywords: Dynamic Convex Risk Measures, Multiple Priors, Uncer-
tainty, Robust Representation, Time-Consistency, Blackwell-Dubins

JEL-Classification: C61, C65, D81

1 Introduction

In this article, we take a look at two distinct agents deciding on investment in a risky project contingent on individual assessments of risk in a dynamic uncertain setting. We assume that they are not certain about the true underlying distribution but about sure and impossible events and agree upon those. Each decision maker, however, assesses risk in a manner she thinks appropriate, resulting in two possibly distinct measures of risk. We restrict attention to a specific class of risk measures based on plausible axioms as elaborated in [Föllmer & Schied, 04]: convex risk measures. The question we tackle in our framework is for the the long term behavior of agents. More precisely, what can be said about the evolution of the underlying risk measures inducing individual behavior. We see that, with increasing information in course of time, the two distinct risk measures will converge to each other meaning that our decision makers agree on a common appropriate measure of risk in the long run. More precisely, both will act as expected utility maximizers with respect to the true underlying distribution. In this sense all uncertainty vanishes in the limit and only risk remains. Equivalently, utility functionals as introduced in [Maccheroni et al., 06b] induced by our class of risk measures converge. Interpreted in terms of financial markets, our results show that herding is eventually inevitable. In H.P. Minsky's theory of financial instability (cp. [Schnyder, 02]), herding plays a major role for causes of financial bubbles.

Before we rigorously model the problem, we come up with an appropriate class of risk measures or, equivalently, utility functionals: In the financial industry, value at risk (VaR) still is used as a standard approach to assess and manage risk despite its well-known shortcomings. The ongoing prominence of VaR is owed to its apparent simplicity and intuitiveness. Hence, an alternative way to assess risk has to particularly compete in these respects with VaR. In our opinion, the axiomatic approach that we briefly describe now satisfies this prerequisite by virtue of a simple robust representation: Coherent risk measures were introduced in [Artzner et al., 99] in a static setting and have been generalized to a dynamic framework in [Riedel, 04]. Tangible problems in this setup are inter alia discussed in [Riedel, 10]. The equivalent theory of multiple prior preferences in a static setup is introduced in [Gilboa & Schmeidler, 89]; a dynamic generalization is given in [Epstein & Schneider, 03]. Applying coherent risk measures substantially decreases model risk as they do not assume a specific probability distribution to hold but assume a whole set of equally likely probability models. Moreover, they possess a simple robust

representation in terms of maximal expected loss. However, as they assume homogeneity, coherent risk measures do not account for liquidity risk. Though in financial applications, the Basel II accord requires a "margin of conservatism", coherent risk measures are far too conservative when estimating risk of a project as they result in a worst case approach. Furthermore, popular examples of risk measures, as e.g. entropic risk, are not coherent.

Hence, it seems worthwhile to consider a more sophisticated axiomatic approach: [Föllmer & Schied, 04] introduce convex risk measures as a generalization of coherent ones relaxing the homogeneity assumption. Equivalently, [Maccheroni et al., 06a] generalize multiple prior preferences to variational preferences. Convex risk measures are applied to a dynamic setup in [Föllmer & Penner, 06] for a stochastic payoff in the last period or, equivalently, in [Maccheroni et al., 06b] in terms of dynamic variational preferences. [Cheridito et al., 06] apply dynamic convex risk measures to stochastic payoff processes. Given a set of possible probabilistic models, convex risk measures are less conservative than coherent ones. Dynamic convex risk measures as well as dynamic variational preferences possess a robust representation in terms of minimal penalized expectation. Both approaches are equivalent as their robust representations coincide up to a factor of -1. It his, hence, more a question of taste which approach to use; the mathematical theory is identical. The minimal penalty, serving as a measure for uncertainty aversion, uniquely characterizes the risk measure or, respectively, the preference. Conditions on the minimal dynamic penalty characterize time-consistency of the dynamic convex risk measure.

We take the robust representation of a dynamic convex risk measure in terms of minimal penalty for granted. As a main result of this article we achieve a generalization of the famous Blackwell-Dubins theorem in [Blackwell & Dubins, 62] from conditional probabilities to time-consistent dynamic convex risk measures: We pose a condition on the minimal penalty in the robust representation, always satisfied by coherent risk measures, forcing the convex risk measure to converge to the conditional expected value under the true underlying distribution. Intuitively, this result states that, eventually, the uncertain distribution is revealed or, in other words, uncertainty diminishes as information is gathered but risk remains. The agent, as she has learned about the underlying distribution, is again in the framework of being an expected utility maximizer with respect to the true underlying distribution. In this sense, distinct agents assess risk in an identical way in the limit if they agree upon impossible events and apply time-consistent dynamic convex risk measures. Hence long-run

behavior of agents converges.

Our generalization of the Blackwell-Dubins theorem serves as an alternative approach to limit behavior of time-consistent dynamic convex risk measures as the one in [Föllmer & Penner, 06]. The result particularly states the existence of a limiting risk measure. As an example we consider dynamic entropic risk measures or, equivalently, dynamic multiplier preferences. We, however, show a Blackwell-Dubins type result to hold, even if we relax the time-consistency assumption. Again, we obtain existence of a limiting risk measure but in a more general manner than [Föllmer & Penner, 06] for not necessarily time-consistent convex and coherent risk measures.

Furthermore, we elaborate an example for non-time-consistent risk that satisfies the properties of our main theorem: We make explicit a learning mechanism for a penalty in terms of conditional relative entropy.

The article is structured as follows: The next section introduces the underlying probabilistic model. Section 3 elaborately discusses robust representation of dynamic convex risk measures and introduces dynamic entropic risk measures. Section 4 generalizes the Blackwell-Dubins theorem to conditional expectations. The following two sections then apply this result to coherent and convex risk measures: First, Section 5 in the time-consistent case and, then, Section 6 without assuming time-consistency. Section 7 states examples. Then we conclude.

2   Model

For our model we start with a discrete time set $t \in \{0, ..., T\}$ where T is an infinite time horizon.

Let $\mathbb{P}_0$ be the reference distribution on the underlying measurable space $(\Omega, \mathcal{F})$ with filtration $(\mathcal{F}_t)_t$. $\mathbb{P}_0$ can be seen as the true distribution of the states. Let $\mathcal{M}^e(\mathbb{P}_0)$ denote the set of all distributions on $(\Omega, \mathcal{F})$ equivalent to $\mathbb{P}_0$. Due to our assumption to only consider distributions equivalent to $\mathbb{P}_0$, the reference distribution merely fixes the null-sets of the model. This assumption has no influence on the stochastic structure of the distributions it just tells the decision makers what sure or impossible events are. An economic interpretation of this assumption was given in [Epstein & Marinacci, 06]. They related it to an axiom on preferences first postulated in [Kreps, 79]. He claimed that if a DM is ambivalent between an act x and $x \cup x'$ then he should also be ambivalent between $x \cup x''$ and $x \cup x' \cup x''$. Meaning if the possiblity of choosing x' in addition to x brings no extra utility compared to just being able to choose x, then also no additional utility should arise

from being able to choose x' supplementary to $x \cup x''$. Furthermore we define $X : \Omega \to \mathbb{R}$ to be an $\mathcal{F}$-measurable random variable which can be interpreted as a payoff at final time T. Assume X being essentially bounded with ess $\sup |X| = \kappa > 0$. Having constructed the filtered reference space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}_0)$ as above, the sets of almost surely bounded $\mathcal{F}$-measurable and $\mathcal{F}_t$-measurable random variables are denoted by $L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P}_0)$ and $L_t^\infty := L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}_0)$, respectively. All equations have to be understood $\mathbb{P}_0$-almost surely.

3 Dynamic Convex Risk Measures

Now we introduce a notion of risk measures that we consider appropriate for our framework. In this article, we apply the theory of convex risk measures as set out in [Föllmer & Penner, 06] for end-period payoffs. For payoff processes, convex risk measures are elaborated in [Cheridito et al., 06]. We do not consider the axiomatic approach to convex risk but take the robust representation of dynamic convex risk measures or, equivalently, of dynamic variational preferences as given.

Definition 3.1 (Dynamic Convex Risk & Penalty Functions). (a) *A family* $(\rho_t)_t$ *of mappings* $\rho_t : L^\infty \to L_t^\infty$ *is called a* dynamic convex risk measure *if each component* ρ_t *is a conditional convex risk measure, i.e. for all* $X \in L^\infty$, ρ_t *can be represented in terms of*

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)} \left(\mathbb{E}^\mathbb{Q} \left[-X \,|\, \mathcal{F}_t \right] - \alpha_t(\mathbb{Q}) \right),$$

where $(\alpha_t)_t$ *denotes the dynamic penalty function, i.e. a family of mappings* $\alpha_t : \mathcal{M}^e(\mathbb{P}_0) \to L_t^\infty$, $\alpha_t(\mathbb{Q}) \in \mathbb{R}_+ \cup \infty$, *closed and grounded. For technical details on the penalty see [Föllmer & Schied, 04].*
(b) *Equivalently, we define the* dynamic concave monetary utility function $(u_t)_t$ *by virtue of* $u_t = -\rho_t$, *i.e.*

$$u_t(X) = \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)} \left(\mathbb{E}^\mathbb{Q} \left[X \,|\, \mathcal{F}_t \right] + \alpha_t(\mathbb{Q}) \right).$$

Remark 3.2. (a) *By Theorem* 4.5 *in [Föllmer & Penner, 06], the above robust representation in terms of* $\mathcal{M}^e(\mathbb{P}_0)$ *is sufficient to capture all time-consistent dynamic convex risk measures.*
(b) *Assuming risk neutrality but uncertainty aversion, no discounting, and no intermediate payoff,* $(u_t)_t$ *is the robust representation of dynamic variational preferences as introduced in [Maccheroni et al., 06b]. In this sense, all our results also hold for dynamic variational preferences.*

Assumption 3.3. *In the robust representation, we assume the penalty α_t to be given by the minimal penalty $\alpha_t^{\min}$. The minimal penalty is introduced in terms of acceptance sets in [Föllmer & Penner, 06], p.64: For every $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$*

$$\alpha_t^{\min}(\mathbb{Q}) := \operatorname*{ess\,sup}_{X \in L^\infty : \rho_t(X) \leq 0} \mathbb{E}^{\mathbb{Q}}\left[-X \mid \mathcal{F}_t\right].$$

As stated in the respective references, every dynamic convex risk measure $(\rho_t)_t$ can be expressed in terms of the above robust representation uniquely by virtue of the minimal penalty and vice versa. The notion of minimal penalty is justified by the fact that every other penalty representing the same convex risk measure a.s. dominates the minimal one, cp. [Föllmer & Penner, 06]'s Remark 2.7. As the minimal penalty uniquely characterizes the convex risk measure, distinct agents assessing risk in distinct ways only differ by distinct minimal penalty functions.

Remark 3.4. *Now we elaborate more thoroughly on decision makers: Each decision maker i is endowed with an individual dynamic convex risk measure $(\rho_t^i)_t$. Hence in terms of robust representation, decision makers differ by virtue of penalty functions. The only property they share is the knowledge of sure and impossible events which is represented here by the assumption to only consider equivalent distributions in $\mathcal{M}^e(\mathbb{P}_0)$. This assumption is justified in [Föllmer & Penner, 06] from a mathematical point of view and in economic terms in our model section. In this sense, $\mathbb{P}_0$ fixes the null sets that decision makers agree upon.*

In the literature, there are three equivalent ways to introduce convex risk measures: in terms of an axiomatic system, by robust representation, and by acceptance sets. Whereas the second one is equivalent to dynamic variational preferences by robust representation, the latter one makes explicit that risk measures provide guidance for decision making: agent i accepts a risky project X as long as $\rho_t^i(X) \leq 0$.

Further assumptions on the risk measure under consideration will be posed when necessary.

Remark 3.5 (On Coherent Risk)**.** *As set out in the references, the robust representation of coherent risk is a special case of the robust representation of convex risk when the penalty is trivial, i.e. for all t it holds*

$$\alpha_t(\mathbb{P}) = \begin{cases} 0 & \text{if } \mathbb{P} \in \tilde{\mathcal{Q}}, \\ \infty & \text{else} \end{cases}$$

for some set $\tilde{\mathcal{Q}} \subset \mathcal{M}^e(\mathbb{P}_0)$ of priors. Throughout, we assume $\tilde{\mathcal{Q}}$ to be convex and weakly compact.

The following definition is a major assumption needed in order to solve tangible economic problems under convex risk.

Definition 3.6 (Time-Consistency)**.** *A dynamic convex risk measure* $(\rho_t)_t$ *is called* time-consistent *if, for all* $t, s \in \mathbb{N}$*, it holds*

$$\rho_t = \rho_t(-\rho_{t+s}).$$

Remark 3.7. *For the special approach here, [Cheridito et al., 06] show that it suffices to consider $s = 1$ in the above definition.*

Remark 3.8. *As inter alia shown in [Föllmer & Penner, 06], Theorem 4.5, time-consistency of $(\rho_t)_t$ is equivalent to a condition on the minimal penalty called* no-gain condition *in [Maccheroni et al., 06b].*

We now introduce a special class of dynamic convex risk measures that will be used in several examples later on: Dynamic entropic risk measures. Therefore, we first have to introduce:

Definition 3.9 (Relative Conditional Entropy)**.** *For $\mathbb{P} \ll \mathbb{Q}$, we define the* relative conditional entropy *of $\mathbb{P}$ with respect to $\mathbb{Q}$ at time $t \geq 0$ as*

$$\hat{H}_t(\mathbb{P}|\mathbb{Q}) := \mathbb{E}^{\mathbb{P}}\left[\log \frac{Z_T}{Z_t}\,\Big|\,\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{Z_T}{Z_t}\log \frac{Z_T}{Z_t}\,\Big|\,\mathcal{F}_t\right]\mathbb{I}_{\{Z_t>0\}},$$

where $(Z_t)_t$ by virtue of $Z_t := \frac{d\mathbb{P}}{d\mathbb{Q}}|_{\mathcal{F}_t}$ denotes the density process of $\mathbb{P}$ with respect to $\mathbb{Q}$.

Definition 3.10 (Entropic Risk Measures)**.** *Let $\delta > 0$ be arbitrary but fixed. We say that dynamic convex risk $\rho_t^e(X)$ of a random variable $X \in L^\infty$, is obtained by a* dynamic entropic risk measure *given reference model $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$ if it is of the form*

$$\rho_t^e(X) := \operatorname*{ess\,sup}_{\mathbb{P}\in\mathcal{M}^e(\mathbb{P}_0)} \left(\mathbb{E}^{\mathbb{P}}[-X|\mathcal{F}_t] - \delta \hat{H}_t(\mathbb{P}|\mathbb{Q})\right). \tag{1}$$

Remark 3.11. *The variational formula for relative entropy implies*

$$\rho_t^e(X) = \delta \log(\mathbb{E}^{\mathbb{Q}}[e^{-\frac{1}{\delta}X}|\mathcal{F}_t]).$$

Thinking of the penalty as an inverse likelihood for distributions to rule the world or a measure for uncertainty aversion, an entropic risk measure means that the agent in an uncertain setting beliefs the reference model $\mathbb{Q}$ as most likely and distributions "further away" as more unlikely. The equivalent dynamic variational preference with penalty given by relative conditional entropy are the well-known *miltiplier preferences*.

4 Adaption of Blackwell-Dubins Theorem

As a cornerstone for our main result on convergence of time-consistent dynamic convex risk measures, we first generalize the famous Blackwell-Dubins theorem, cp. [Blackwell & Dubins, 62], from conditional probabilities to conditional expectations of risky projects.

Proposition 4.1. *Let $\mathbb{Q}$ be absolutely continuous with respect to $\mathbb{P}_0$,[1] X as in the definition of the model, then*

$$\left| \mathbb{E}^{\mathbb{Q}}[X \mid \mathcal{F}_t] - \mathbb{E}^{\mathbb{P}_0}[X \mid \mathcal{F}_t] \right| \to 0 \quad \mathbb{P}_0\text{-almost surely for } t \to \infty.$$

Proof. Given $\mathbb{P}_0$ and $\mathbb{Q}$, $\mathbb{Q}$ being assumed absolutely continuous with respect to $\mathbb{P}_0$, i.e. $\frac{d\mathbb{Q}}{d\mathbb{P}_0} = q$ is well defined, and for every t, $\frac{d\mathbb{Q}(\cdot|\mathcal{F}_t)}{d\mathbb{P}_0(\cdot|\mathcal{F}_t)} = q(\cdot|\mathcal{F}_t)$. Then, the following line of equations holds $\mathbb{P}_0$-a.s.:

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[X|\mathcal{F}_t] &= \mathbb{E}^{\mathbb{Q}(\cdot|\mathcal{F}_t)}[X] \\
&= \mathbb{E}^{\mathbb{P}_0(\cdot|\mathcal{F}_t)}[q(\cdot|\mathcal{F}_t)X]
\end{aligned}
$$

and hence

$$
\begin{aligned}
\left| \mathbb{E}^{\mathbb{Q}}[X|\mathcal{F}_t] - \mathbb{E}^{\mathbb{P}_0}[X|\mathcal{F}_t] \right| &= \left| \mathbb{E}^{\mathbb{P}_0(\cdot|\mathcal{F}_t)}\left[(q(\cdot|\mathcal{F}_t) - 1)X \right] \right| \\
&\leq \kappa \left| \mathbb{E}^{\mathbb{P}_0(\cdot|\mathcal{F}_t)}\left[(q(\cdot|\mathcal{F}_t) - 1) \right] \right| \\
&= \kappa \left| \int (q(\cdot|\mathcal{F}_t) - 1)\, \mathbb{P}_0(d\cdot|\mathcal{F}_t) \right|,
\end{aligned}
$$

which converges to zero $\mathbb{P}_0$-a.s. by Blackwell-Dubins theorem as $(\mathcal{F}_t)_t$ is assumed to be a filtration and, hence, an increasing family of σ-fields. $\square$

5 Time-Consistent Risk Measures

We will now show a Blackwell-Dubins type result for coherent as well as convex risk measures in case time-consistency is assumed. We see that the risk measure eventually equals the expected value under the true parameter; in this sense, uncertainty vanishes but risk remains. Thus, the basis for learning the underlying distribution is already incorporated in convex risk measures intuitively as the domain of penalty

---

[1]Note that we have assumed all distributions to be equivalent. In particular, all those are absolutely continuous with respect to each other and this assumption is no restriction within our setup. Moreover, we can take another arbitrary but fixed distribution instead of $\mathbb{P}_0$.

consists of bayesian updated distributions. Interpreted in terms of decision makers, long-run behavior of distinct agents converges as they all behave as risk neutral expected utility maximizers with respect to the true underlying distribution in the limit.

5.1   Time-Consistent Coherent Risk

Let $(\rho_t)_t$ be a time-consistent dynamic coherent risk measure possessing robust representation

$$\rho_t(X) = \sup_{\mathbb{P} \in \tilde{\mathcal{Q}}} \mathbb{E}^{\mathbb{P}}[-X \mid \mathcal{F}_t],$$

with weakly compact and convex set of priors $\tilde{\mathcal{Q}} \subset \mathcal{M}^e(\mathbb{P}_0)$.

Proposition 5.1. *For every essentially bounded $\mathcal{F}$-measurable random variable X as in the model and time-consistent dynamic coherent risk measure $(\rho_t)_t$ we have*

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}_0}[-X \mid \mathcal{F}_t]\right| \to 0 \quad \mathbb{P}_0\text{-almost surely for } t \to \infty.$$

Proof. Thanks to the assumption of time-consistency and compactness there exists a distribution $\mathbb{P}^* \in \tilde{\mathcal{Q}}$ such that $\rho_t(X) = \mathbb{E}^{\mathbb{P}^*}[-X \mid \mathcal{F}_t]$ for all $t \in \{0, ..., T\}$ resulting in the following equation

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}_0}[-X \mid \mathcal{F}_t]\right| = \left|\mathbb{E}^{\mathbb{P}^*}[-X \mid \mathcal{F}_t] - \mathbb{E}^{\mathbb{P}_0}[-X \mid \mathcal{F}_t]\right|,$$

converging to zero as t increases and $\mathbb{P}^* \sim \mathbb{P}_0$ by Proposition 4.1. $\square$

Remark 5.2. *Note that we have not assumed $\mathbb{P}_0 \in \tilde{\mathcal{Q}}$.*

Remark 5.3. *The assumption that $\tilde{\mathcal{Q}}$ is weakly compact is crucial, as it assures that the supremum is actually attained. Additionally it is a necessary property for our result to hold, which is shown in the Proposition 5.4.*

Proposition 5.4. *Weak compactness of the set $\tilde{\mathcal{Q}}$ of priors is a necessary condition for our result in Proposition 5.1 to hold.*

Proof. For the proof, see the counterexample in section 7.2. $\square$

5.2   Time-Consistent Convex Risk

Let $(\rho_t)_t$ be a time-consistent dynamic convex risk measure, hence, possessing the following robust representation:

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)} \left\{\mathbb{E}^{\mathbb{P}}[-X \mid \mathcal{F}_t] - \alpha_t^{\min}(\mathbb{P})\right\}$$

with dynamic minimal penalty $(\alpha_t^{\min})_t$.

Assumption 5.5. *We assume* $(\rho_t)_t$ *to be* continuous from below *for all t, i.e. for every sequence of random variables* $(X_j)_j$, $X_j \in L^\infty$ *for all j, with* $X_j \nearrow X \in L^\infty$ *we have* $\lim_{j\to\infty} \rho_t(X_j) = \rho_t(X)$.

Remark 5.6. *In the coherent case, continuity from below is equivalent to weak compactness of the set* $\{\mathbb{P}|(\alpha_t(\mathbb{P}))_t = 0\} = \tilde{\mathcal{Q}}$ *of priors as inter alia shown in [Riedel, 10].*

This assumption has technical advantages as it ensures the supremum to be achieved in the robust representation of ρ_t. A proof is given in Theorem 1.2 of [Föllmer et al., 09]. It is also shown that continuity from below implies continuity from above. To sum up: continuity from above is equivalent to the existence of a robust representation. Continuity from below (which generalizes the compactness assumption in the coherent case) is equivalent to the existence of a robust representation in terms of a distinct prior distribution, the so called *worst-case distribution.*

From an economic point of view, continuity from below results from a feature of preferences already claimed in [Arrow, 71] and related to this assumption by [Chateauneuf et al., 05]. The condition on preferences we need to ask for in order to obtain this feature is called Monotone Continuity: If an act f is preferred over an act g then a consequence x is never that bad that there is no small p such that x with probability p and f with probability $(1-p)$ is still preferred over g. The same is true for good consequences mixed with g.

Formally this means, for acts $f \succ g$, a consequence x and a sequence of events $\{E_n\}_{n\in\mathbb{N}}$ with $E_1 \supseteq E_2 \supseteq ...$ and $\cap_{n\in\mathbb{N}} E_n = \emptyset$ there exists an $\bar{n} \in \mathbb{N}$ such that

$$\begin{bmatrix} x \text{ if } s \in E_{\bar{n}} \\ f(s) \text{ if } s \notin E_{\bar{n}} \end{bmatrix} \succ g \quad \text{and} \quad f \succ \begin{bmatrix} x \text{ if } s \in E_{\bar{n}} \\ g(s) \text{ if } s \notin E_{\bar{n}} \end{bmatrix}.$$

Now with the help of this assumption we can show the Blackwell-Dubins result for time-consistent convex risk measures:

Proposition 5.7. *For every essentially bounded* $\mathcal{F}$-*measurable random variable* X *and time-consistent dynamic convex risk measure* $(\rho_t)_t$, *continuous from below, it holds*

$$\left| \rho_t(X) - \mathbb{E}^{\mathbb{P}_0}[-X \,|\mathcal{F}_t] \right| \to 0 \quad \mathbb{P}_0\text{-almost surely for } t \to \infty$$

if there exists $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$ *such that* $\alpha_t^{\min}(\mathbb{P}) \to 0$ $\mathbb{P}_0$-*almost surely and* $\alpha_0^{\min}(\mathbb{P}) < \infty$.

Remark 5.8 (On the Assumption)**.** *By the main assumption in Proposition 5.7 there ought to be some* $\mathbb{P}$ *such that the penalty vanishes in*

the long run. This intuitively means that, eventually, nature at least has to pretend some distribution to be the correct one. We see that this is satisfied e.g. in the coherent or in the entropic case.

The assertion then states that it does not matter which risk measure was chosen as long as the penalty is finite in the beginning. In the time-consistent case, the penalty then vanishes for all those parameters and the convex risk eventually will be coherent.

As we will see later, in the non-time-consistent case, nature has to pay a price for not choosing a distribution time-consistently as in that case penalty has to vanish for the true underlying parameter. To conclude: when nature chooses the worst case distribution time-consistently, she merely has to pretend some distribution to be the underlying one. If she does not choose the worst case measures at any stage time-consistently, she has to reveal the true underlying distribution in the long run.

Remark 5.9. *By Theorem* 5.4 *in [Föllmer & Penner, 06] due to time-consistency the assumption* $\alpha_t^{\min}(\mathbb{P}) \to 0$ $\mathbb{P}_0$*-almost surely for some* $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$ *is equivalent to* $\alpha_t^{\min}(\mathbb{Q}) \to 0$ $\mathbb{P}_0$*-almost surely for all* $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$ *with* $\alpha_0(\mathbb{Q}) < \infty$.

Proof of the proposition. By our assumptions on $(\rho_t)_t$ there exists $\mathbb{P}^* \in \mathcal{M}^e(\mathbb{P}_0)$ such that the assertion becomes

$$\left| \mathbb{E}^{\mathbb{P}^*}[-X|\mathcal{F}_t] - \alpha_t^{\min}(\mathbb{P}^*) - \mathbb{E}^{\mathbb{P}_0}[-X|\mathcal{F}_t] \right| \to 0 \quad \mathbb{P}_0\text{-a.s.}$$

By the foregoing proposition on coherent risk, we know that this assertion holds if and only if

$$\left| \alpha_t^{\min}(\mathbb{P}^*) \right| \to 0 \quad \mathbb{P}_0\text{-a.s.}$$

As stated in Remark 5.9, Theorem 5.4 in [Föllmer & Penner, 06] implies this convergence being equivalent to

$$\left| \alpha_t^{\min}(\mathbb{P}) \right| \to 0 \quad \mathbb{P}_0\text{-a.s.}$$

for some $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$ such that $\alpha_0(\mathbb{P}) < \infty$ as assumed to hold in the assertion. □

Again, note that we have not assumed $\mathbb{P}_0$ such that $\alpha_0(\mathbb{P}_0) < \infty$.

Corollary 5.10. *Every dynamic time-consistent convex risk measure* $(\rho_t)_t$ *satisfying the assumptions of the Proposition 5.7 is asymptotically precise as in the sense of [Föllmer & Penner, 06], i.e.* $\rho_t(X) \to \rho_\infty(X) = -X$, *and vice versa.*

Proof. By the assumption of continuity from below, we know that a worst case measure in the robust representation of $(\rho_t)_t$ is actually achieved. By Theorem 5.4 (5) in [Föllmer & Penner, 06] we have that $\rho_t(X) \to \rho_\infty(X) \geq -X$ as we have assumed $\alpha_t^{\min}(\mathbb{P}_0) \to 0$. Proposition 5.11 in [Föllmer & Penner, 06] then shows the assertion.   □

Remark 5.11. *In [Föllmer & Penner, 06] time-consistency is directly used to show the existence of the limit $\rho_\infty := \lim_{t \to \infty} \rho_t$. As, by assumptions on X in the model, $\lim_{t \to \infty} (E^{\mathbb{P}_0}[-X \mid \mathcal{F}_t])$ exists we achieve existence of ρ_∞ from our result not directly from time-consistency. In our propostion the convergence of the α corresponds to asymptotic precision, however starting at a different point of view. The question now is if time-consistency is a necessary condition for our result to hold. If so, we have gained nothing, if not, we have a more general existence result for ρ_∞ than [Föllmer & Penner, 06]. We will tackle the problem of necessity of time-consistency for our result within the next section.*

Proposition 5.12. *$(\rho_t)_t$ being continuous from below is a necessary condition for the result in Theorem 5.7 to hold.*

Proof. In Proposition 5.4 we show necessity of weak compactness of the set of priors for coherent risk measures. However, weak compactness is equivalent to continuity from below and coherent risk measures are particular examples for convex ones. This proofs the assertion.   □

Remark 5.13 (On long run behavior of agents)**.** *Having considered two agents possessing risk measures $(\rho_t^1)_t$ and $(\rho_t^2)_t$, respectively, our result can be interpreted as follows: Both, $(\rho_t^1)_t$ and $(\rho_t^2)_t$ converge to the conditional expectation with respect to the true underlying distribution. In this sense, $|\rho_t^1 - \rho_t^2| \to_{t \to \infty} 0$ and both agents behave identically in the limit.*

6   Not Necessarily Time-Consistent Risk Measures

We will now achieve a Blackwell-Dubins type result for dynamic coherent and convex risk measures for which we do not pose the time-consistency assumption. However, we still assume the dynamic risk measure to be continuous from below (i.e. in the coherent case the set of priors to be weakly compact). We can still show that anticipation of risk converges to the expected value of a risky project X as defined in the model with respect to the underlying distribution $\mathbb{P}_0$.

6.1   Non-Time-Consistent Coherent Risk

We will now restate the result in a manner that time-consistency is not needed. We however need to assume that learning takes place; which is a more liberal assumption than time-consistency as seen in Section 7.3.

Definition 6.1. (a) *Given a dynamic convex risk measure* $(\rho_t)_t$, *continuous from below but not necessarily time-consistent, we call a distribution* $\mathbb{P}_t^* \in \mathcal{M}^e(\mathbb{P}_0)$ *instantaneous worst case distribution at* t *if it satisfies*[2]

$$\rho_t(X) = \mathbb{E}^{\mathbb{P}_t^*}\left[-X\,|\,\mathcal{F}_t\right] - \alpha_t^{\min}(\mathbb{P}_t^*).$$

(b) *We say* learning takes place *if there exists a* $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$ *such that the instantaneous worst case measures* $\mathbb{P}_t^* \to \mathbb{P}$ *weakly for* $t \to \infty$. *In the coherent case we need* $\mathbb{P} \in \tilde{\mathcal{Q}}$ *as the penalty is infinite otherwise.*

In this very definition, we see however, that the agent does not have to learn the true underlying $\mathbb{P}_0$. In this sense, nature might mislead her to a wrong distribution.

We can now relax the time-consistency assumption in the main result of this article. Note that time-consistency is a special case of the Definition 6.1 given continuity from below as in that case the sequence of instantaneous worst case measures is constant. Hence, we achieve the more general result:

Proposition 6.2. *Let* $(\rho_t)_t$ *be a not necessarily time-consistent dynamic coherent risk measure for which learning takes place. Then*

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}_0}[-X\,|\,\mathcal{F}_t]\right| \to 0 \quad \mathbb{P}_0\text{-almost surely for } t \to \infty$$

Proof. To make things clearer we will write the proof in terms of penalty functions and not in terms of priors. We know that a coherent risk measure has a robust representation of a convex risk measure with a penalty

$$\alpha_t^{\min}(\mathbb{P}) = \left\{ \begin{array}{ll} 0 & \text{if } \mathbb{P} \in \tilde{\mathcal{Q}}, \\ \infty & \text{else} \end{array} \right.$$

where $\tilde{\mathcal{Q}} = \{\mathbb{P}\,|\,(\alpha_t^{\min}(\mathbb{P}))_t = 0\}$, the set of priors. As we are in the case of a coherent risk measure, we particularly have $\alpha_t^{\min}(\mathbb{P}_t^*) = 0$.

---

[2]Note, that existence is locally guaranteed by continuity from below. As we however have not assumed time-consistency, the instantaneous worst case distributions at each time period may differ, hence global existence is not necessarily fulfilled.

First, note that in case $\alpha_t^{\min}(\mathbb{P}) \to \infty$ for all $\mathbb{P} \in \tilde{\mathcal{Q}}$[3], our convergence result cannot hold, as $\lim_{t\to\infty} \mathbb{E}^{\mathbb{P}_0}[-X|\mathcal{F}_t]$ exists and is finite by assumption.

Secondly, in the time-consistent (coherent as well as convex) case, it suffices to assume $\alpha_t^{\min}(\bar{\mathbb{P}}) \to 0$ for some $\bar{\mathbb{P}} \in \mathcal{M}^e(\mathbb{P}_0)$. This assumption in the time-consistent case is equivalent to $\alpha_t^{\min}(\mathbb{P}) \to 0$ for all $\mathbb{P}$ for which $\alpha_0^{\min}(\mathbb{P}) < \infty$ by Theorem 5.4 in [Föllmer & Penner, 06].

Let us now turn to the proof itself: As $(\rho_t)_t$ is assumed continuous from below, i.e. $\tilde{\mathcal{Q}}$ is assumed to be weakly compact and non-empty, we achieve an instantaneous worst case distribution at each time step, i.e. at any t, there exists $\mathbb{P}_t^* \in \mathcal{M}^e(\mathbb{P}_0)$ s.t.

$$\rho_t(X) = \mathbb{E}^{\mathbb{P}_t^*}\left[-X\,|\,\mathcal{F}_t\right] - \alpha_t^{\min}(\mathbb{P}_t^*) = \mathbb{E}^{\mathbb{P}_t^*}[-X|\mathcal{F}_t].$$

The proof is completed by showing the following convergence[4]

$$\mathbb{E}^{\mathbb{P}_n^*}[-X|\mathcal{F}_t] \to \mathbb{E}^{\mathbb{P}_0}[-X|\mathcal{F}_\infty] \qquad \text{for } n, t \to \infty.$$

In order to do this we look at the following equation for $n \geq t$ which uses the projectivity of the density, i.e. of the Radon-Nikodym derivative:

$$\mathbb{E}^{\mathbb{P}_n^*}[-X|\mathcal{F}_t] = \mathbb{E}^{\mathbb{P}_0}[-X\frac{d\mathbb{P}_n^*}{d\mathbb{P}_0}\Big|_{\mathcal{F}_n}\,|\mathcal{F}_t].$$

Define the following sequence of random variables $Y_n := -X\frac{d\mathbb{P}_n^*}{d\mathbb{P}_0}\Big|_{\mathcal{F}_n}$. These have finite expectation and thanks to our assumption that learning takes place and the original Blackwell-Dubins result we have

$$\mathbb{P}_0[\lim_{n\to\infty} Y_n = -X] = \mathbb{P}_0\left[-X\frac{d\mathbb{P}_\infty^*}{d\mathbb{P}_0}\Big|_{\mathcal{F}_\infty} = -X\right] = 1.$$

Then, by Lemma 6.4, the assertion follows. $\qquad\qquad\square$

Remark 6.3. *Again, note that we have* not *assumed* $\mathbb{P}_0 \in \tilde{\mathcal{Q}}$.

In the foregoing proof, we need a general martingale convergence result as stated in [Blackwell & Dubins, 62], Theorem 2. We know from Doob's famous martingale convergence result that

$$\mathbb{E}^{\mathbb{P}}[X|\mathcal{F}_t] = \lim_{t\to\infty} \mathbb{E}^{\mathbb{P}}[X|\mathcal{F}_\infty] \quad \mathbb{P}_0 - a.s.$$

---

[3]Of course, convergence is trivial in this case due to triviality of the penalty function.

[4]By our assumptions we know:

- $\mathbb{E}^{\mathbb{P}_n^*}[-X|\mathcal{F}_t] \to \mathbb{E}^{\mathbb{P}}[-X|\mathcal{F}_t]$ for $n \to \infty$ as $\mathbb{P}_n^* \to \mathbb{P}$ by Portemonteau's Theorem.

- $\mathbb{E}^{\mathbb{P}_n^*}[-X|\mathcal{F}_t] \to \mathbb{E}^{\mathbb{P}_n^*}[-X|\mathcal{F}_\infty]$ for $t \to \infty$ by Proposition 4.1.

The question now is, whether the result also holds when letting $n, t \to \infty$ at once.

In the time-consistent case, where $\mathbb{P}_i^* = \mathbb{P}_j^*$ for all i, j, this is immediate by Proposition 4.1.

under suitable assumptions. The question is: If $X_n \nearrow_n X$ in some sense, is it true that

$$\mathbb{E}^{\mathbb{P}}[X_n|\mathcal{F}_t] = \lim_{n,t\to\infty} \mathbb{E}^{\mathbb{P}}[X|\mathcal{F}] \quad \mathbb{P}_0 - a.s.?$$

A positive answer is given in the following lemma.

Lemma 6.4. *Fix $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$. Let $(Y_n)_n$ be a sequence of $\mathcal{F}$-measurable random variables such that $\mathbb{E}^{\mathbb{P}}[\sup_n |Y_n|] < \infty$. Assume $Y_n \to_{n\to\infty} Y$ $\mathbb{P}_0$-almost surely for some $\mathcal{F}$-measurable random variable Y. Then, it holds[5]*

$$\lim_{n,t\to\infty} \mathbb{E}^{\mathbb{P}}\left[Y_n|\,\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}}\left[Y|\,\mathcal{F}\right].$$

Proof. We re-sample the proof in [Blackwell & Dubins, 62]: For $k \in \mathbb{N}$, set $G_k := \sup\{Y_n|n \geq k\}$. If $n \geq k$, we hence have $Y_n \leq G_k$ and thus

$$\mathbb{E}^{\mathbb{P}}\left[Y_n|\,\mathcal{F}_t\right] \leq \mathbb{E}^{\mathbb{P}}\left[G_k|\,\mathcal{F}_t\right] \qquad (2)$$

for all t. Together with Doob's martingale convergence result and Lebesgue's theorem, we achieve

$$
\begin{aligned}
z \quad &:= \quad \limsup_{j\to\infty} \sup_{n,t\geq j} \mathbb{E}^{\mathbb{P}}\left[Y_n|\,\mathcal{F}_t\right] \\
&\overset{(2)}{\leq} \quad \limsup_{j\to\infty} \sup_{t\geq j} \mathbb{E}^{\mathbb{P}}\left[G_k|\,\mathcal{F}_t\right] \\
&= \quad \lim_{t\to\infty} \mathbb{E}^{\mathbb{P}}\left[G_k|\,\mathcal{F}_t\right] \\
&\overset{\text{Doob}}{=} \quad \mathbb{E}^{\mathbb{P}}\left[G_k|\,\mathcal{F}\right]
\end{aligned}
$$

and

$$z \leq \lim_{k\to\infty} \mathbb{E}^{\mathbb{P}}\left[G_k \mid \mathcal{F}\right] \overset{\text{Lebesgue}}{=} \mathbb{E}^{\mathbb{P}}\left[Y \mid \mathcal{F}\right].$$

In the same token,

$$x := \liminf_{j\to\infty} \inf_{t,n\geq j} \mathbb{E}^{\mathbb{P}}\left[Y_n|\,\mathcal{F}_t\right] \geq \mathbb{E}^{\mathbb{P}}\left[Y|\mathcal{F}\right],$$

which completes the proof since

$$x = \liminf_{j\to\infty} \inf_{t,n\geq j} \mathbb{E}^{\mathbb{P}}\left[Y_n|\,\mathcal{F}_t\right] \leq \limsup_{j\to\infty} \sup_{n,t\geq j} \mathbb{E}^{\mathbb{P}}\left[Y_n|\,\mathcal{F}_t\right] = z.$$

$\square$

---

[5]The convergence in the assertion of the lemma can also be shown in L^1.

Remark 6.5. *Note, that the above new version of the fundamental result particularly holds for time-consistent dynamic coherent risk measures as then such a limiting $\mathbb{P}$ as in the Definition 6.1(b) always exists, the worst case one. However, we particularly have an existence result for the limit $\rho_\infty := \lim_{t\to\infty} \rho_t$ in the non time-consistent case and thus a more general existence result than in [Föllmer & Penner, 06].*

6.2   Non-Time-Consistent Convex Risk

As in the case of coherent risk measures, we now state our generalization of the Blackwell-Dubins theorem when the dynamic convex risk measure is *not* assumed to be time-consistent. As in the coherent case, we assume that learning takes place, i.e. there exists $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$ such that the instantaneous worst case $\mathbb{P}_t^* \to \mathbb{P}$ as $t \to \infty$. Furthermore, we have to assume $\alpha_t^{\min}(\mathbb{P}_t^*) \to 0$ as $t \to \infty$:[6] As in the foregoing proof, we achieve convergence of the conditional expectations under the family of instantaneous worst case distributions to the conditional expectation under $\mathbb{P}_0$.

Proposition 6.6. *For every risky project X as set out in the model and dynamic convex risk measure $(\rho_t)_t$, continuous from below but not necessarily time-consistent, we have*

$$\left| \rho_t(X) - \mathbb{E}^{\mathbb{P}_0}[-X \mid \mathcal{F}_t] \right| \to 0 \quad \mathbb{P}_0\text{-almost surely for } t \to \infty$$

if learning takes place for an instantaneous worst case sequence $(\mathbb{P}_t^*)_t$ toward some $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$ and we have

$$\alpha_t^{\min}(\mathbb{P}_t^*) \to 0 \quad \mathbb{P}_0\text{-almost surely for } t \to \infty.$$

Proof. Applying the procedure used in the proof of Proposition 6.2 to the proof of Proposition 5.7 shows the assertion. $\square$

7   Examples

In this section, we first consider dynamic entropic risk measures as a prominent economic example of time-consistent dynamic convex risk measures. From a preference based perspective, this example can equivalently be stated in terms of multiplier preferences. In the second part we state a counterexample serving as proof for Proposition 5.4 and 5.12. Lastly, we consider a dynamic risk measure that is not time-consistent but satisfies the properties of Proposition 6.6.

---

[6]Note, again, we do not have to assume $\alpha_t^{\min}(\mathbb{P}_0) \to 0$.

7.1   Entropic Risk

Here, we will have a look at time-consistent dynamic entropic risk measure $(\rho_t^e)_t$. Recall its Definition 3.10 in terms of

$$\rho_t^{\mathrm{e}}(X) := \delta \log \mathbb{E}\left[e^{-\gamma X}\middle| \mathcal{F}_t\right]$$

for some model parameter $\delta > 0$. A fundamental result shows that the robust representation of dynamic entropic risk is given in terms of conditional relative entropy as penalty function, i.e. for all n, we have for $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$ and penalty's reference distribution $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$

$$\alpha_t^{\min}(\mathbb{P}) = \frac{1}{\gamma}\hat{H}_t(\mathbb{P}|\mathbb{Q}) := \frac{1}{\gamma}\mathbb{E}^{\mathbb{P}}\left[\ln\frac{Z_T}{Z_t}\middle| \mathcal{F}_t\right],$$

where $Z_t := \frac{d\mathbb{P}}{d\mathbb{Q}}\Big|_{\mathcal{F}_t}$, the Radon-Nikodym derivative of $\mathbb{P}$ with respect to $\mathbb{Q}$ conditional on $\mathcal{F}_t$.

The fundamental Blackwell-Dubins Theorem immediately shows that $|\mathbb{P}(\cdot|\mathcal{F}_t) - \mathbb{Q}(\cdot|\mathcal{F}_t)| \to 0$ for every $\mathbb{P}, \mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$. Hence, we have that $\frac{Z_T}{Z_t} \to 1$ $\mathbb{P}_0$-a.s. for $t \to \infty$ and hence

$$\alpha_t^{\min}(\mathbb{P}) \to 0$$

showing Proposition 5.7 to hold. This is an alternative way to show the last assertion in Theorem 6.3 in [Föllmer & Penner, 06] directly.

7.2   Counterexample

To show necessity of continuity from below in Proposition 5.7 we consider the following example introduced in [Föllmer & Penner, 06]:

The underlying probability space consists of the state space $\Omega = (0, 1]$ endowed with the Lebesgue measure $\mathbb{P}_0$ and a filtration $(\mathcal{F}_t)_t$ generated by the dyadic partitions of Ω. This means $\mathcal{F}_t$ is generated by the sets $J_{t,k} := (k2^{-t}, (k+1)2^{-t})$ for $k = 0, ..., 2^{t-1}$. In this setting [Föllmer & Penner, 06] construct a time-consistent coherent and therefore convex risk measures with $\alpha_t^{\min}(\mathbb{P}_0) \to 0$ $\mathbb{P}_0$-a.s. of the following form:

$$\rho_t(X) = -\operatorname{ess\,sup}\{m \in L_t^\infty \,|m \le X\}.$$

That this sequence from all properties assumed in Proposition 5.7 is only missing continuity from below (equivalent to weak compactness of the set of priors) can be seen in the following way: Let t be arbitrary but fixed and X defined by virtue of

$$X(\omega) = \begin{cases} 0 & \text{for } \omega \in (0, (2^t - 1)2^{-t}], \\ 1 & \text{else.} \end{cases}$$

Then we can construct a sequence $(X_n)_n$, $X_n \nearrow X$, such that $\rho_t(X_n) = 0$ for all n but $\rho_t(X) = -X \neq 0$. This shows $(\rho_t)_t$ not being continuous from below.

Now we still have to show that for this construction the statement of our proposition is not fulfilled. To verify this look at a set A assumed to be $\mathcal{F} := \sigma(\bigcup_{t \geq 0} \mathcal{F}_t)$-measurable such that $\mathbb{P}_0[A] > 0$ and $\mathbb{P}_0[A^c \cap J_{t,k}] \neq 0$ for all t and k. For this set, it holds

$$\lim_{t \to \infty} \left| \rho_t(\mathbb{1}_A) - \mathbb{E}^{\mathbb{P}_0}[-\mathbb{1}_A \,|\mathcal{F}_t] \right| = \lim_{t \to \infty} |0 + \mathbb{P}_0[A \,|\mathcal{F}_t]| = \mathbb{P}_0[A] > 0$$

and hence necessity of the continuity assumption is shown.

The skeptical reader might now object that such a set A might not exist. For sake of completeness we briefly quote a set A from [Föllmer & Penner, 06] that satisfies our assumptions: Let A be defined by virtue of its complement

$$A := \left(\bigcup_{t=1}^{\infty} \bigcup_{k=1}^{2^t - 1} U_{\epsilon_t}(k 2^{-t}) \right)^c,$$

where U_{ϵ_t} denotes the ϵ_t-neighborhood and $\epsilon_t \in]0, 2^{-2t}]$.

7.3   A Non Time-Consistent Example

Our last example is not only worth considering as it constitutes a non time-consistent convex risk measure satisfying the properties of our main result but also as it explicitly states a learning mechanism by virtue of the minimal penalty. To us it seems that this example is more conveniently posed in a parametric setting. Hence, let a distribution $\mathbb{P}^\theta \in \mathcal{M}^e(\mathbb{P}^{\theta_0})$ on the measurable space $(\Omega, \mathcal{F})$ with filtration $(\mathcal{F}_t)_t$ be uniquely given by a parameter $\theta \in \Theta$. Assume the parameter space Θ such that all induced distributions $\mathbb{P}^\theta$, $\theta \in \Theta$, are equivalent to $\mathbb{P}^{\theta_0}$ for some fixed reference parameter $\theta_0 \in \Theta$. We have to add more structure to the underlying reference space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P}^{\theta_0})$: We fix $(S, \mathcal{A})$ as a measure space where S describes the possible states of the world at a fixed point in time t and set $\Omega := \bigotimes_{t=0}^{T} S_t$, $S_t = S$. On this space let $\mathcal{F}$ be the product σ-field generated by all projections $\pi_t : \Omega \to S_t$ and let the elements of the filtration $\mathcal{F}_t$ be generated by the sequence $\pi_1, ..., \pi_t$. We assume $\theta = (\theta_t)_t \in \Theta$; every entity θ_t characterizes a distribution in $\mathcal{M}(S_t)$ possibly dependent on $(\theta_i)_{i<t}$. The family $\theta = (\theta_t)_t$ then defines a prior $\mathbb{P}^\theta \in \mathcal{M}^e(\mathbb{P}^{\theta_0})$. Set $\theta^t := (\theta_1, \ldots, \theta_t)$ and $\mathbb{P}^{\theta_t}$ denote the marginal distribution at t induced by θ_t.

We now introduce a model for wich dynamic entropic risk measures in Definition 3.10 serve as a vehicle: We choose the best fitting distribution as reference distribution in the conditional relative entropy.

Definition 7.1 (Experience Based Entropic Risk). *A penalty $(\hat{\alpha}_t)_t$ is said to be achieved by* experience based entropic learning *if given as*

$$\hat{\alpha}_t(\eta) := \delta \hat{H}_t(\mathbb{P}^\eta | \mathbb{P}^{\hat{\theta}})$$

for $\delta > 0$, $\eta = (\eta_t)_t \in \Theta$ and $\hat{\theta} = (\hat{\theta}_t)_t$ achieved in the following manner: Being at time t, the reference family $\hat{\theta}$ of parameters is achieved by

$$\hat{\theta}_i = \begin{cases} \hat{\theta}_i & i \leq t, \\ \hat{\theta}_t & i > t, \end{cases}$$

where $\hat{\theta}_i$ is the maximum likelihood estimator given past observations. The resulting convex risk measure $(\hat{\rho}_t)_t$ incorporating this very penalty function is then called experience based entropic risk measure*.*

Remark 7.2. *$(\hat{\alpha}_t)_t$ is well defined as penalty as inter alia shown in [Föllmer & Schied, 04]. Hence, the model is well defined, i.e. $(\hat{\rho}_t)_t$ is a dynamic convex risk measure, which also directly follows from the axioms.*

Now, as the reference distribution is stochastic, we achieve:

Proposition 7.3. *Experience based entropic risk is in general not time-consistent.*

Proof. As proof we construct the following counterexample. $\qquad \square$

Example 7.4 (Entropic Risk in a Tree). *Since our example is mainly for demonstration purposes we restrict ourselves to a simple Cox-Ross-Rubinstein model with 3 time periods. Each time period is independent of those before. One could imagine that in every time period a different coin is thrown and the result of the coin toss determines the realization in the tree, e.g. from heads results up and from tails down. The payoffs of our random variable X are limited to the last time-period and are as shown in the figure below. For tractability reasons we also confine ourselves to a single likelihood function $l(\cdot \mid \theta)$. The probability for going up in this tree will always be assumed to lie in the interval $[a, b]$ where $0 < a \leq b < 1$.*

Time-period 2: *Since we want to show a contradiction to time-consistency we will show that the recursive formula*

$$\hat{\rho}_t(X) = \hat{\rho}_t(-\hat{\rho}_{t+s}(X)) \qquad \text{for all } t \in [0, T] \text{ and } s \in \mathbb{N}$$
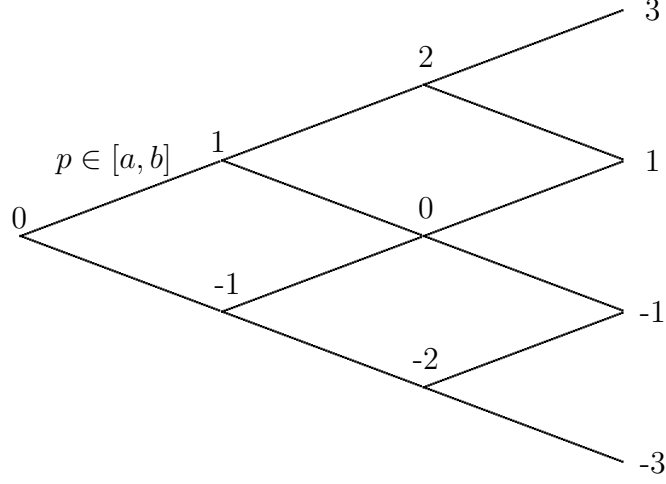
Figure 1: Cox-Ross-Rubinstein Model

is violated. So we start with the calculation of $\rho_2(X)$ for the different sets in $\mathcal{F}_2$

$$\hat{\rho}_2(X)(\text{up}, \text{up})$$

$$= \underset{p \in [a,b]}{\text{ess sup}} \, \mathbb{E}\left[-X \mid \mathcal{F}_2\right](\text{up}, \text{up}) - \mathbb{E}\left[\ln\left(\frac{\theta_2}{\theta_2^*}\right) \mid \mathcal{F}_2\right](\text{up}, \text{up})$$

$$= \underset{p \in [a,b]}{\sup} \left(-3p - 1 + p - p \ln \frac{p}{b} - (1-p)\ln\left(\frac{1-p}{1-b}\right)\right)$$

$$= \ln\left(be^{-3} + (1-b)e^{-1}\right),$$

where the reference distribution $\mathbb{P}^{\theta^*}$ induced by θ^* is determined by the following maximization:

$$\theta^* = (\theta_0^*, \theta_1^*, \theta_2^*), \quad \theta_2^* \in \underset{\theta_2 \in [a,b]}{\arg \max} \, l(\text{up} \mid \theta_2)$$

giving us the maximum-likelihood estimator for what happened in the last time-period which we also think is the right distribution for the next time-period.

The result of this computation can also be obtained by using a variational form which can for example be found in [Föllmer & Penner, 06] and takes the following form

$$\hat{\rho}_t(X) = \ln \mathbb{E}^{\mathbb{P}^{\theta^*}}\left[\exp(-X) \mid \mathcal{F}_t\right],$$

where $\mathbb{P}^{\theta^*}$ is again the reference distribution the decision maker establishes by looking at the past, which, as we look at naive learning, will

again only be what happened in the last period. Since this gives way for an easier and quicker computation we will use this form for the following calculations:

$$\hat{\rho}_2(X)(\text{down}, \text{up}) = \hat{\rho}_2(X)(\text{up}, \text{down})$$

$$= \ln \mathbb{E}^{\mathbb{P}^{\theta^*}}\left[\exp(-X) \mid \mathcal{F}_2\right](\text{down}, \text{up})$$

$$= \ln \left(\frac{1}{2}e^{-1} + \frac{1}{2}e^1\right),$$

if $\frac{1}{2} \in [a, b]$

For the last possible event in time 2 our risk-measure takes the following value:

$$\begin{aligned}
\hat{\rho}_2(X)(\text{down}, \text{down}) &= \ln \mathbb{E}^{\mathbb{P}^{\theta^*}}\left[\exp(-X) \mid \mathcal{F}_2\right](\text{down}, \text{down}) \\
&= \ln \left(ae^1 + (1-a)e^3\right).
\end{aligned}$$

Time-period 1: *If for the next time-period we maintain the assumption of time-consistency and make use of the recursive formula, using the variational form as we did above will yield*

$$\hat{\rho}_1(X)(\text{up}) = \hat{\rho}_1(-\hat{\rho}_2(X))(\text{up}) = \ln \mathbb{E}^{\mathbb{P}^{\theta^*}}\left[\exp(\hat{\rho}_2(X)) \mid \mathcal{F}_1\right](\text{up})$$

$$= \ln \left(b\left(be^{-3} + (1-b)e^{-1}\right) + (1-b)\frac{1}{2}\left(e^{-1} + e^1\right)\right)$$

$$= \ln \left(b^2 e^{-3} + \left(\frac{1}{2} + \frac{1}{2}b - b^2\right)e^{-1} + \frac{1}{2}e^1\right).$$

Now if we calculate $\hat{\rho}_1(X)(up)$ without the time-consistency assumption meaning we cannot use the recursive formula we obtain the following equation:

$$\hat{\rho}_1(X)(up) = \operatorname*{ess\,sup}_{p,q \in [a,b]} \mathbb{E}^{p,q}\left[-X \mid \mathcal{F}_1\right](up) - \mathbb{E}^{p,q}\left[\ln\left(\frac{\theta_1 \theta_2}{\theta_1^* \theta_2^*}\right) \mid \mathcal{F}_1\right](up)$$

$$= \ln \left(b2e^{-3} + 2b(1-b)e^{-1} + (1-b)2e^1\right).$$

This clearly is in general not the same as we obtained under the assumption of time-consistency. However if our dynamic experience based entropic risk measure were time-consistent these calculations should give us the same results. Hence this example clearly shows us that the assumption of our risk measure being time-consistent only leads up to contradictions and can therefore not be true.

Having observed experience based entropic risk $(\hat{\rho}_t)_t$ not being

time-consistent,[7] we show that it nevertheless satisfies the conditions of Proposition 6.6. By standard results on conditional entropic risk measures, $(\hat{\rho}_t)_t$ is continuous from below. Let us restrict ourselves to the iid case: We know that we achieve $\hat{\theta}_t \to \bar{\theta}_0$, $\mathbb{P}^{\theta_0}$-a.s., where $\theta_0 = (\bar{\theta}_0)_t$ for some $\bar{\theta}_0$ inducing a marginal distribution in $\mathcal{M}(S_t)$.

Furthermore, Proposition 6.6 is applicable and hence, our generalization of the Blackwell-Dubins theorem holds for experience based entropic risk. Indeed: By definition of the penalty and our considerations in Section 7.1, $\hat{\alpha}_t^{\min}(\theta) := \delta \hat{H}_t(\mathbb{P}^\theta | \mathbb{P}^{\hat{\theta}}) \to 0$ as $t \to \infty$ for all $\theta \in \Theta$ by the fundamental Blackwell-Dubins theorem. Secondly, as the maximum likelihood estimator is asymptotically stable, i.e. $\hat{\theta}_t \to \bar{\theta}_0$, the conditional reference distributions $\mathbb{P}^{\hat{\theta}}(\cdot | \mathcal{F}_t)$ of the relative conditional entropy converge. Hence the worst-case instantaneous distributions $\mathbb{P}_t^*$ converge as in Definition 6.1 due to continuity of the entropy and as the effective domain of the penalty is given by conditional distributions, a fact that is made particularly precise in [Maccheroni et al., 06b].

---

[7]More generally, it can be shown that learning leads to time-inconsistency in the entropic case no matter what mechanism is used to achieve the reference distribution: A reference distribution $\mathbb{P}^{\tilde{\theta}}$ for experience based entropic risk is said to be obtained by *general learning* if the family $(\tilde{\theta}_t)_t$ is a family of random variables. We call the resulting dynamic convex risk measure $(\tilde{\rho}_t^g)_t$ defined by virtue of $\tilde{\alpha}_t^g := \hat{H}_t(\cdot | (\tilde{\theta}_t)_t)$ in the robust representation *general experience based entropic risk*.

However, general experience based entropic risk $(\tilde{\rho}_t^g)_t$ is in general not time-consistent. Intuitively, the minimal penalty function uniquely defines a risk measure. Changing the reference distribution due to learning results in a different minimal penalty and hence, a distinct risk measure. More formally, this can be seen as follows: Let $\tilde{\theta} = (\tilde{\theta}_1, \dots)$ be obtained by general learning and ${}^t\tilde{\theta}$ such that $\mathbb{P}^{{}^t\tilde{\theta}} = \mathbb{P}^{\tilde{\theta}}(\cdot | \mathcal{F}_t)$. Let $Z_{t+1} := \left. \frac{d\mathbb{Q}^{{}^t\tilde{\theta}}}{d\mathbb{Q}^{{}^{t+1}\tilde{\theta}}} \right|_{\mathcal{F}_{t+1}}$. Then, we have

$$
\begin{aligned}
\tilde{\rho}_t^g(X) &= \ln \mathbb{E}^{\mathbb{Q}^{{}^t\tilde{\theta}}} \left[e^{\ln \mathbb{E}^{\mathbb{Q}^{{}^{t+1}\tilde{\theta}}} \left[\frac{Z_T}{Z_{t+1}} e^{-X} \middle| \mathcal{F}_{t+1} \right]} \middle| \mathcal{F}_t \right] \\
&= \ln \mathbb{E}^{\mathbb{Q}^{{}^t\tilde{\theta}}} \left[e^{-(-\rho_{t+1}(X - \ln(\frac{Z_T}{Z_{t+1}})))} \middle| \mathcal{F}_t \right] \\
&= \tilde{\rho}_t^g(-\tilde{\rho}_{t+1}^g(X - \ln(\frac{Z_T}{Z_{t+1}}))) \\
&\neq \tilde{\rho}_t^g(-\tilde{\rho}_{t+1}^g(X)),
\end{aligned}
$$

if $\frac{Z_T}{Z_{t+1}} \neq 1$ a.s., i.e. if, intuitively speaking, learning actually takes place and, hence, the reference distributions at distinct time periods differ.

8   Conclusions

The major contribution of our results is to carry over the famous Blackwell-Dubins theorem from probability distributions to convex risk measures. We have shown that two agents with possibly distinct risk assessments merely have to agree on sure and impossibly events for their attitude towards a project to merge in the long run. Starting with individual dynamic convex risk measures both act as expected utility maximizers with respect to the true underlying distribution in the limit. It is particularly striking that the results still hold when time-consistency is not posed as an assumption.

We therefore introduced a generalization of the famous Blackwell-Dubins theorem on "Merging of Opinions" to conditional expected values. Existence of a worst case distribution due to continuity from below and time-consistency then allowed for a further generalization to coherent and convex risk measures. In particular, we have obtained the existence of the limiting risk measure ρ_∞ in that case.

By virtue of a counterexample, we have shown necessity of continuity from below for our result. However, we have shown that time-consistency is not necessary for the Blackwell-Dubins type result to hold. In particular, we have obtained a more general existence result for the limiting risk measure ρ_∞ than in [Föllmer & Penner, 06].

Further research should be conducted in the direction of our results. First, of course, the riddle of explicitly constructing convex risk measures by virtue of the penalty function is still to solve; in particular, how a learning mechanism might be introduced without destroying the assumption of time-consistency. Weaker notions of time-consistency that are satisfied in a "learning" environment should be introduced along with a comprehensive theory allowing for solutions of tangible economic and social problems.

In the article at hand, we have considered risky projects with final payoffs, i.e. random variables of the form $X \in \mathcal{F}$. We have shown convergence of convex risk measures to the conditional expected value with respect to the true underlying distribution: a generalization of the Blackwell-Dubins theorem to (not necessarily time-consistent) convex risk measures for final payoffs. To us it seems being an interesting, yet challenging, task to generalize our result to the case of convex risk measures for stochastic payoff processes $(X_t)_t$ with respect to some filtration $(\mathcal{F}_t)_t$, where each X_t denotes the stochastic payoff in period t. [Cheridito et al., 06] introduce dynamic convex risk measures for these stochastic processes and elaborately discuss time-consistency issues but do not inspect limiting behavior. A major difficulty in the case of stochastic processes is that the assumption of equivalent distributions

should be replaced by local equivalence, cp. [Riedel, 10]. Hence, the main question turns out to be if the result still holds assuming local instead of global equivalence as done here.

References

[Artzner et al., 99] Artzner, Philippe; Delbaen, Freddy; Eber, Jean-Marc & Heath, David: *Coherent Measures of Risk*, Mathematical Finance, Vol.9, No.3, pp.203-228, 1999.

[Arrow, 71] Arrow, Kenneth: *Essays in the Theory of Risk Bearing*, Markham Publishing Co., Chicago, 1971.

[Blackwell & Dubins, 62] Blackwell, David & Dubins, Lester: *Merging of Opinions with Increasing Information*, Ann. Math. Statistics, Vol. 33, No. 3, pp. 882-886, 1962.

[Chateauneuf et al., 05] Chateauneuf, Alain; Maccheroni, Fabion; Marinacci, Massimo & Tallon, Jean-Marc: *Monotone Continuous Multiple Priors* , Economic Theory Vol. 26 No.4, pp. 973-982, 2005.

[Cheridito et al., 06] Cheridito, Patrick; Delbaen, Freddy & Kupper, Michael: *Dynamic Monetary Risk Measures for Bounded Discrete-Time Processes*, Electronic Journal of Probability, Vol.11, No.3, pp.57-106, 2006.

[Epstein & Marinacci, 06] Epstein, Larry & Marinacci, Massimo: *Mutual Absolute Continuity of Multiple Priors* , Journal of Economic Theory, Vol.137, No.1, pp.716-720, 2007.

[Epstein & Schneider, 03] Epstein, Larry & Schneider, Martin: *Recursive Multiple Priors*, Journal of Economic Theory, Vol.113, pp.1-13, 2003.

[Epstein & Schneider, 07] Epstein, Larry & Schneider, Martin: *Learning under Ambiguity*, Review of Economic Studies, Vol.74, pp.1275-1303, 2007.

[Föllmer & Schied, 04] Föllmer, Hans & Schied, Alexander: *Stochastic Finance, An Introduction in Discrete Time*, 2nd edition, Walter DeGruyter, 2004.

[Föllmer & Penner, 06] Föllmer, Hans & Penner, Irene: *Convex Risk Measures and the Dynamics of their Penalty Functions*, Statistics and Decisions, Vol.24, pp.61-96, 2006.

[Föllmer et al., 09] Föller, Hans; Schied, Alexander & Weber, Stefan *Robust Preferences and Robust Portfolio Choice*, Handbook of Numerical Analysis, Vol.15, pp.29-89, 2009.

[Gilboa & Schmeidler, 89] Gilboa, Itzhak & Schmeidler, David: *Maximin Expected Utility with non-unique Prior*, Journal of Mathematical Economics, Vol.18, pp.141-153, 1989.

[Kreps, 79] Kreps, David: *A Representation Theorem for 'Preference for Flexibility'*, Econometrica, Vol. 47, No.3, pp.565-577, 1979.

[Maccheroni et al., 06a] Maccheroni, Fabio; Marinacci, Massimo & Rustichini, Aldo: *Ambiguity Aversion, Robustness, and the Variational Representation of Preferences*, Econometrica, Vol.74, No.6, pp.1447-1498, 2006.

[Maccheroni et al., 06b] Maccheroni, Fabio; Marinacci, Massimo & Rustichini, Aldo: *Dynamic Variational Preferences*, Journal of Economic Theory, Vol.128, pp. 4-44, 2006.

[Riedel, 04] Riedel, Frank: *Dynamic Coherent Risk Measures*, Stochastic Processes and their Applications, Vol.112, pp.185-200, 2004.

[Riedel, 10] Riedel, Frank: *Optimal Stopping with Multiple Priors*, Econometrica, Vol.77, No.3, pp.857-908, 2010.

[Schied, 07] Schied, Alexander: *Optimal Investment for Risk- and Ambiguity-Averse Preferences: A Duality Approach*, Finance and Stochastics, Vol.11, pp.107-129, 2007.

[Schnyder, 02] Schnyder, Mark: *Die Hypothese finanzieller Instabilität von Hyman P. Minsky – Ein Versuch der theoretischen Abgrenzung und Erweiterung*, Doctoral Thesis, Freiburg University, 2002.