On Dynamic Coherent and Convex Risk Measures: Risk Optimal Behavior and Information Gains

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors

der Wirtschafts- und Gesellschaftswissenschaften

durch die

Rechts- und Staatswissenschaftliche Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

vorgelegt von

Daniel Engelage

aus Minden/Westfalen

Bonn 2009

Dekan:	Prof. Dr. Christian Hillgruber
Erstreferent:	Prof. Dr. Frank Riedel
Zweitreferent:	JProf. Dr. Eva Lütkebohmert-Holtz

Tag der mündlichen Prüfung: 21.08.2009

Abstract

We consider tangible economic problems for agents assessing risk by virtue of dynamic coherent and convex risk measures or, equivalently, utility in terms of dynamic multiple priors and variational preferences in an uncertain environment.

Solutions to the Best-Choice problem for a risky number of applicants are well-known. In Chapter 2, we set up a model with an ambiguous number of applicants when the agent assess utility with multiple prior preferences. We achieve a solution by virtue of multiple prior Snell envelopes for a model based on so called assessments. The main result enhances us with conditions for the ambiguous problem to possess finitely many stopping islands.

In Chapter 3 we consider general optimal stopping problems for an agent assessing utility by virtue of dynamic variational preferences. Introducing variational supermartingales and an accompanying theory, we obtain optimal solutions for the stopping problem and a minimax result. To illustrate, we consider prominent examples: dynamic entropic risk measures and a dynamic version of generalized average value at risk.

In Chapter 4, we tackle the problem how anticipation of risk in an uncertain environment changes when information is gathered in course of time. A constructive approach by virtue of the minimal penalty function for dynamic convex risk measures reveals time-consistency problems. Taking the robust representation of dynamic convex risk measures as given, we show that all uncertainty is revealed in the limit, i.e. agents behave as expected utility maximizers given the true underlying distribution. This result is a generalization of the fundamental Blackwell-Dubins theorem showing coherent as well as convex risk measures to merge in the long run.

Keywords: Uncertainty, Dynamic Variational Preferences, Dynamic Multiple Prior Preferences, Dynamic Convex Risk Measures, Dynamic Coherent Risk Measures, Dynamic Penalty, Robust Representation, Time-Consistency, Best-Choice Problem, Optimal Stopping, Blackwell-Dubins Theorem

If we begin with certainties,
we shall end in doubts;
if we begin with doubts,
and are patient,
we shall end in certainties.

Marcus Aurelius

Acknowledgments

I am deeply grateful to my advisor Professor Dr. Frank Riedel, Institute of Mathematical Economics (IMW), Bielefeld University, and Professor Dr. Patrick Cheridito, Dept. for Operations Research and Financial Engineering (ORFE), Princeton University. This doctoral thesis would not have been possible without their support, advice and ideas for further investigation. Both were very amenable to this project and helped me solving severe problems that I have encountered throughout my work. Fruitful discussions with Professor Riedel as well as his vital comments and suggestions have profoundly improved this thesis. Besides mathematical strictness, he taught me the importance of intuitive applicability for theoretical results. Throughout my doctoral studies, I have thus not only gained methodological and technical knowledge in economic theory but also implementation skills in industry. In addition, I am indebted to JProfessor Dr. Eva Lütkebohmert-Holtz for her efforts in surveying this thesis.

The underlying articles would have barely been conceivable without the help of Monika Bier, coauthoring the fourth chapter, who was always accessible. We have enjoyed many inspiring discussions over the past four years.

I acknowledge financial grants from the German Academic Exchange Service (DAAD) and the German Research Foundation (DFG) through the Bonn Graduate School of Economics (BGSE). I am deeply sincere for institutional support by the BGSE, in particular the efforts by Professor Dr. Urs Schweizer, Dr. Silke Kinzig and Corinna Lehmann as well as the entire faculty.

Sincere thanks are given to my fellow students at Bonn Graduate School of Economics and Princeton University, especially Almira Buzaushina and Natalia Besedovsky for intellectual as well as moral support. Many fellow students have grown on me and became sincere friends over the past years. Finally, special thanks go to my family for supporting me throughout my studies.

Bonn, June 2009 *Daniel Engelage*

Contents

Chapter 1

General Introduction

In light of the current financial crisis accompanied by an unprecedented amount of uncertainty in markets, financial industry as well as market supervisors are in need of sophisticated yet applicable instruments to quantify and manage risk. Therefore, the general question how agents anticipate risk in uncertain environments is not just one of theoretical interest in economists but necessitates a wholehearted and understandably framed answer procurable to be adopted by professionals in real world practice.

1.1 An Axiomatic Approach to Risk Measurement

For the financial industry, value at risk (VaR) still seems to be the standard approach in quantification of risk despite its several well known shortcomings elaborately discussed e.g. in [McNeil et al., 05]: A danger in applying VaR is the possibility of accumulating a highly risky portfolio and the fact that diversification effects might not be accounted for. The prominence of VaR as industry standard is mainly owed to its simplicity and intuitiveness. In overcoming these shortcomings, alternative approaches to risk assessment have to be introduced which result in risk measures that are easily communicated, intuitive and straightforward to implement for solving tangible problems. As an example, alternative risk measures have to be readably applicable to minimal capital requirement models in line with the Basel II accord to ensure financial stability for banking institutions while being easily manageable.

A sensible axiomatic approach to quantify risk was first mentioned in [Artzner et al., 99] for a static setting: The authors introduced the notion of *coherent risk measures* assessing risk of projects considered as real valued random variables. Several other references as [Delbaen, 02] advanced upon this approach for more general probability spaces. The approach to coherent risk measures is based on four quite intuitive axioms and leads to a simple and hence applicable *robust representation* that we encounter later. We will rigorously introduce the underlying notion of risk measures in the respective chapters of this thesis. However, for the sake of completeness and an intuitive understanding at this stage, the four axioms for a risk measure to be coherent are given by *monotonicity*, *cash invariance*, *sub-additivity* and *positive homogeneity of degree one*. The major advantage of coherent risk measures is their simple and intuitive robust representation in terms of maximized expected loss as elaborated below. Furthermore, coherent risk measures do not necessitate a specific probabilistic model and hence help to significantly reduce *model risk* in applications. However, coherent risk measures have two major shortcomings: First, they overestimate risk as they lead to a worst-case approach by virtue of robust representation: An issue that has to be scoped with from point of view of financial institutions having an intrinsic interest in assessing risk not too conservatively when calculating minimal capital requirements. Secondly, due to the assumption of homogeneity, coherent risk measures do not take into account *liquidity risk* as one of the major problems in the current financial crisis.

As an advancement, *convex risk measures* are introduced inter alia in [Föllmer & Schied, 04] for a static setting: The assumptions of sub-additivity and homogeneity are replaced by *convexity*, intuitively stating that diversification reduces risk. It is immediately seen that coherent risk measures are a special class of convex ones.

The prominent VaR is neither coherent nor convex. However, average value at risk (AVaR), also called expected shortfall or conditional value at risk in respective literature, is coherent as it, intuitively speaking, considers not just quantiles but has a closer look in the respective tails of a distribution. The most prominent example for a convex risk measure is *entropic risk* conveying an elegant intuition discussed below.

Of course, financial markets are intrinsically dynamic and agents are supposed to use information they gain in course of time. Hence, *dynamic convex risk measures* are considered in many of the cited references. Dynamic coherent risk measures can inter alia be found in [Riedel, 04] or [Artzner et al., 07]. Wholehearted elaborations of dynamic convex risk measures are given in [Föllmer & Penner, 06] or [Föllmer et al., 09] for risky projects seen as payoffs in the last period and in [Cheridito et al, 06] for risky projects seen as stochastic processes.

To give some flesh to the bone without being mathematically precise at this stage, consider a risky project X and an information process given by filtration $(\mathcal{F}_t)_t$. We then call a family $(\rho_t)_t$ of $(\mathcal{F}_t)_t$-adapted random variables a dynamic convex risk measure if each ρ_t is a *conditional convex risk measure* and hence possesses the robust representation

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

for some *dynamic (minimal) penalty function* $(\alpha_t)_t$. Intuitively, at time t the agent evaluates risk of a position X as the maximal conditional expected loss with respect to all possible distributions but has to be compensated by nature for choosing a specific distribution in terms of the non-negative penalty. In this sense, robust representation of convex risk is a *maximized penalized expected loss.* Intuitively, the smaller the penalty the more likely the agent considers the respective distribution to be the correct probabilistic model ruling the world. As coherent risk measures are just a special case of convex ones, they also satisfy this robust representation but in terms of a much simpler penalty that can only take the values zero or infinity and is, hence, called *trivial penalty* further on. Throughout we consider robust representation in terms of minimal penalty.

By virtue of the robust representation above, a convex risk measure is uniquely characterized by its minimal penalty function. Given a coherent and a convex risk measure for which the sets of distributions with infinite penalty coincide, we see that the convex risk measure assesses risk more liberally than the coherent one: A conservative over estimation of risk when using coherent risk measures is the price we have to pay for substantially reducing model risk. In other terms, if two agents assess risk in a convex manner, the first one with penalty $(\alpha_t^1)_t$, the second one with penalty $(\alpha_t^2)_t$, then, given $(\alpha_t^1)_t \geq (\alpha_t^2)_t$, the first agent is less uncertainty averse. In this sense, the penalty is a measure for *uncertainty aversion*. In other terms, $(\rho_t^1)_t$ assesses risk more liberal than $(\rho_t^2)_t$.

When considering dynamic problems under convex risk, the integral question is how conditional convex risk measures at distinct time-periods are connected. To scope with this issue, the notion of *time-consistency* was introduced. It is inter alia elaborately discussed in [Föllmer & Penner, 06] and [Cheridito et al, 06]. Formally, time-consistency is defined as $\rho_t = \rho_t(-\rho_{t+1})$, a Bellman equation for nature, the intuition of which is given in the respective chapters of this thesis. By virtue of the robust representation, time-consistency of a dynamic convex risk measure can equivalently be stated as a property of the minimal penalty function, called the *no-gain condition*. In the coherent case, time-consistency reduces to a *stability condition* on the set of distributions for which penalty vanishes, the set of *multiple priors*. This stability condition and equivalent notions are inter alia discussed in [Riedel, 09].

1.2 A Preference Based Alternative

So far, we have focused on risk measures as underlying objects. Equivalently, we can build our results on a preference based point of view of the problem. *Multiple prior preferences* were introduced in [Gilboa & Schmeidler, 89] and applied to a dynamic framework in [Epstein & Schneider, 03]. These types of preferences are, assuming ambiguity aversion but risk neutrality as well as a discount factor of unity and no intermediate payoffs, equivalent to coherent risk measures: Robust representation of multiple prior preferences is the same as the one for coherent risk measures up to a minus sign. In that sense, an agent evaluating utility of a risky project in an uncertain environment in terms of multiple priors, considers the minimal expected payoff with respect to all distributions she deems likely to rule the world, i.e. have a vanishing penalty. [Riedel, 09] approaches optimal stopping problems with respect to multiple priors and thereto generalizes the *Snell envelope* approach appropriately.

The preference based equivalent to convex risk measures is given by *variational preferences*, introduced in a static set up in [Maccheroni et al., 06a] and generalized to a dynamic framework in [Maccheroni et al., 06b]. For the sake of intuitive convenience, [Cheridito et al, 06] actually state their theory of time-consistent dynamic convex risk measures in terms of utility functionals instead of risk measures. As for the equivalence of coherent risk measures and multiple priors, the robust representation of variational preferences coincides with that of convex risk measures up to a minus sign: Robust representation of variational preferences might hence be seen as a minimal penalized expectation. In the dynamic setting, time-consistency considerations are the same for the preference based approach as for the one in terms of risk measures and result in the no-gain condition on the minimal penalty function.

Given these considerations, we note that it does not matter for our insights whether we apply the preference based approach or the ansatz by virtue of risk measures: Each chapter may be reformulated in terms of the other approach. However, in chapters 3 and 4, we consider a theory in terms of dynamic variational preferences. Chapter 5 is based on dynamic convex risk measures.

1.3 Particular Considerations

The main chapters of this thesis, each of which self contained in notation, are based on three articles. The first two consider optimal behavior of agents assessing risk in terms of coherent and convex risk measures or, equivalently, assessing utility in terms of multiple prior preferences and variational preferences. The third one is concerned with merging of dynamic convex risk measures as information is gained in course of time. The latter chapter is coauthored by Monika Bier.

As we have already mentioned, there are basically three distinct but equivalent ways to introduce convex and hence coherent risk measures. First, by virtue of an axiomatic system. Secondly, through a robust representation as given above. Lastly, in terms of *acceptance sets*. The latter approach makes explicit that, intuitively, a risk measure might be seen as the smallest amount of numeraire that is necessary to make the agent accept a risky project. This intuitively shows the tight connection of risk measures to preferences. The starting point for our discussions in the subsequent chapters, however, will be the robust representation of convex risk measures or variational preferences, respectively. In this sense, we build our models on fundamental results concerning the representation of time-consistent dynamic convex risk measures as inter alia stated in [Föllmer & Penner, 06].

In **Chapter 2** we generalize the so called *Best-Choice problem* to multiple priors. Extensions of the "simple" Best-Choice or Secretary problem are inter alia introduced in [Gilbert & Mosteller, 66] or [Freeman, 83]. Solutions to the problem for a risky number of applicants, i.e. when the number of applicants is given by a random variable with a known distribution, can be found in [Presman & Sonin, 72], [Stewart, 81], [Petrucelli, 83] and [Irle, 80]. Here, we set up a model with an ambiguous number of applicants, i.e. a distinct distribution on the random number of applicants is not known. An impossibility result shows the natural ambiguous generalization of the risky model not to be solvable in terms of a time-consistent approach. We achieve a solution by virtue of the multiple prior Snell envelope introduced in [Riedel, 09] for the ambiguous model based on so called *assessments*. The main result enhances us with conditions for the ambiguous problem to possess finitely many stopping islands and constitutes a generalization of the main result in [Presman & Sonin, 72]. A major practical contribution of our ambiguous model is elimination of model risk that is highly apparent in the risky setup of the problem. However, before building our own model for the Best-Choice problem under ambiguity, we take some time to review extensions of the problem with a fixed number of applicants and discuss distinct approaches to the problem with a risky number of applicants.

In **Chapter 3** we consider general *optimal stopping* problems of payoff processes for an agent assessing risk in a convex manner as set out in [Cheridito et al, 06] or, equivalently, assessing utility by virtue of dynamic variational preferences as in [Maccheroni et al., 06b]. By generalizing the approach in [Riedel, 09] from the coherent to the convex case introducing *variational supermartingales* and an accompanying theory, we obtain optimal solutions for the stopping problem and a *minimax result.* As a byproduct, we generalize the model in [Maccheroni et al., 06b] to the case of infinite probability spaces. To illustrate the main results, we consider prominent examples: *dynamic entropic risk measures* and a dynamic version of *generalized average value at risk* (gAVaR); for our theory to be applicable, we have to introduce a time-consistent dynamic version of gAVaR.

Having discussed risk optimal behavior of agents in the foregoing two chapters, in **Chapter 4**, coauthored by Monika Bier, we answer the following question: How does anticipation of risk and, hence, optimal behavior in an uncertain environment change when *information* is gathered in course of time? We answer this question in terms of dynamic convex risk measures or, equivalently, dynamic variational preferences. Therefore, we first introduce a *constructive approach* by virtue of the minimal penalty function conceived as likelihood of priors showing that time-consistency turns out to be a major problem when explicitly constructing a dynamic penalty. Hence, in the second part of that chapter we take the robust representation of dynamic convex risk measures as given and show that all uncertainty is revealed in the limit, i.e. distinct agents behave as expected utility maximizers given the true underlying distribution. In other terms, distinct dynamic convex risk measures *merge* to conditional expectation with respect to the underlying distribution as information increases. Note, it is just uncertainty that is revealed: There is still risk going on by virtue of the underlying distribution. This result is a generalization of the fundamental *Blackwell-Dubins theorem*, cp. [Blackwell & Dubins, 62], to convex risk measures. A particular achievement is the extension of the Blackwell-Dubins theorem to *not necessarily time-consistent* convex risk measures. We thus obtain a more general existence result for limiting risk measures than [Föllmer & Penner, 06]. As an application we consider dynamic entropic risk measures.

So far, we have just quite briefly discussed related literature. As the subject matters of the underlying articles are quite different, a scientific placement of our results within the literature seems cumbersome in this general introduction. Hence, elaborate discussions on literature and relevance of our results are stated in the respective chapters.

As the intuition of our results can mostly be inferred from the respective mathematical proofs, we have decided to state them within the chapters and not in separate appendices.

Chapter 2

The Best-Choice Problem with an Ambiguous Number of Applicants

2.1 Introduction

The *Best-Choice* or *Secretary problem* is not just a popular anecdote you can tell at dinner parties but constitutes a whole field in *stochastic optimization theory*. The origin of this problem is not quite clear today but traces back to the 1950s. Historical abridgments may be found in [Freeman, 83] or [Ferguson, 89]. The latter article summarizes the "simple" Secretary problem as follows:

- You are ought to assign a position as a secretary to exactly one of $n \in \mathbb{N}$ applicants; n is known.

- Applicants are interviewed sequentially in random order and ranked relative to the ones already interviewed. The decision to accept an applicant, i.e. to stop the process of job interviews, is based on relative ranks only.

- Once rejected, a job candidate cannot be recalled upon.
- You want to have the best secretary: You obtain payoff 1 if choosing the best applicant among all n and zero else. Put equivalently, you want to find a stopping time for the interview process maximizing the probability of accepting the best applicant.

The solution to this formulation, i.e. the stopping rule that maximizes the probability of choosing the best applicant, is well known:

- Given $s \in \mathbb{N}$. Reject the first $s-1$ applicants and then accept the first relatively best thereafter.
- Choose s to maximize the probability of choosing the best applicant among all n within the last $n-s$. For $n \gg 0$, $s \approx \frac{1}{e}n \approx \frac{1}{3}n$.

Due to the variety of distinct formulations of the Secretary problems, it seems worthwile to consider a generic definition:

Definition 2.1.1 ([Ferguson, 89], p.284)**.** *A* Secretary Problem *is a sequential observation and selection problem in which the payoff [and the decision to stop] depends on the observations only through their relative ranks and not otherwise on their actual values.*[1]

Though we use the above definition, three types of Secretary Problems are customarily distinguished:

- The *no-information* problem: only the rank of an upcoming applicant is observable. All orderings are equally probable.[2]

[1] In this very definition, we already see the problem of talking about ambiguity: By definition, a Secretary problem is considered under ambiguity as there is no distribution of actual values known. However, when we talk of ambiguity in context of this problem, we mean an ambiguous number of applicants.

[2] [Chudjakow & Riedel, 09] introduce ambiguity about the orderings.

- The *full-information* problem: To each applicant, an actual value can be attached. These values are distributed with a known probability distribution.

- The *partial-information* problem: Actual values of applicants are observed but the distribution is only partially known, i.e. belongs to some family $(F_\theta)_{\theta\in\Theta}$ with unknown parameter θ.

As mentioned in [Ferguson, 89], a first rigorous approach to the Secretary problem is elaborated in [Lindley, 61]: A solution to the finite horizon problem as well as an approximation for the infinite horizon problem is discussed. In advance, a more general utility function than above is considered. Surprisingly, the partial-information and the no-information problem are quite similar: [Stewart, 78] shows that a non-informative prior leads to the same solution as the no-information problem, i.e. bayesian learning does not contribute to maximizing the probability of choosing the best applicant.

The concern of this article is to extend the Secretary problem to *ambiguity*. Introducing ambiguity may be done in two distinct ways: First, in the no-information case, ambiguity is introduced about the number n of applicants. In the simple problem above, n is fixed and known to the observer. Several extensions relax this assumption by introducing *risk*: [Presman & Sonin, 72] assume a random number N of applicants being distributed by a known prior distribution on $\mathbb{N}$. Another approach, makes use of applicants arriving at poisson random times with known parameter and choice to be accomplished before a fixed time horizon, e.g. [Stewart, 81]. In our approach, we assume applicants arriving at fixed times $1, 2, \dots$ and introducing ambiguity over the number of applicants N in terms of multiple priors. A second approach would be introducing ambiguity over arrival times, e.g. ambiguous poisson arrival times, up to a fixed time horizon T.

Secondly, ambiguity could be introduced over the actual qualification of applicants in the partial information setting, usually considered to be risky but not uncertain. In this sense, the no-information problem considered here is a case of maximal ambiguity on qualifications.

To be precise, here we *tackle the no-information problem with an ambiguous number of applicants*. Our approach to ambiguity is based on (recursive) multiple priors as in [Gilboa & Schmeidler, 89] and [Epstein & Schneider, 03] on the number of applicants, applied to optimal stopping as in [Riedel, 09]. An alternative approach to ambiguity makes use of non-additive measures, so called *capacities*, and corresponding *Choquet integrals* with respect to those (cp. [Föllmer & Schied, 04]). These allow for uncertainty averse as well as uncertainty loving agents (cp. [Skulj, 01]) and the degree of convexity of the capacity is a measure for uncertainty averseness. However, for optimal stopping problems the *multiple priors* framework seems more adequate. Under the assumption of uncertainty averseness both approaches are equivalent, as stated in [Chateauneuf, 1991].

[Riedel, 09] shows that an uncertainty averse but risk neutral agent in a time-consistent dynamic ambiguous setup behave as expected utility maximizer with respect to some *worst-case distribution* as she plays against a malevolent nature, underpinning [Gilbert & Mosteller, 66], where the problem is modeled as a two person game: one player chooses the applicant, the other the order in which applicants are presented in order to minimize the observers probability of choosing the best. Such a two person game in a risky and in an ambiguous context is also discussed in [Bruss, 84] and [Hill & Krengel, 91], respectively.

In course of modeling and solving the ambiguous Best-Choice problem, we have also to tackle the following problem: As usual, a stopping time only depends on the information gathered so far. Hence, in the no-information case, the decision to stop at time t only hinges on the relative rank of the t^{th} applicant. In case of a fixed or a risky number of applicants, if t is a candidate, the optimal solution is measurable with respect to the σ-algebra generated by the relative rankings up to time t, i.e. the stopping rule is a random variable but at the realization at time t it is known whether to stop or not for sure. Hence, we call these *deterministic stopping rules*. In [Presman & Sonin, 72] it is shown that randomization at time t does not increase expected payoff, a result extended in [Abdel-Hamid et al., 82]. Thus, it is enough to consider deterministic stopping rules. However, [Hill & Krengel, 91] consider *randomized stopping rules*. Such a rule τ is not measurable with respect to the sigma algebra generated by the relative ranks up to time t but satisfies $\{\tau = t\} \in \sigma(R_1, U_1, \ldots, R_t, U_t)$, where the R_i's denote relative ranks and the U_i's independent random experiments. Intuitively: At time t, the stopping rule specifies a random experiment, e.g. tossing a coin, whose outcome determines stopping or not. In other words, at time t, we stop with a probability that is fixed upon realization at t. Randomized stopping times are discussed in [Siegmund, 67]. As we will see, in our model it suffices to consider deterministic rules.

We will encounter that a straightforward ambiguous generalization of the risky setup in [Presman & Sonin, 72] is not only doubtful from an economic perspective but also does not satisfy the crucial time-consistency condition needed for solving the problem: We show an intuitive impossibility result stating that time-consistency cannot be achieved by virtue of a set of priors on $\mathbb{N}$ and come up with a distinct approach based on so called *assessments* $\mu := (\mu_i)_i$, i.e. families of distributions on the number of applicants, where μ_i may be thought of as the distribution on the number of applicants the agent considers being correct upon observing the i^{th} applicant. Multiple priors in this framework then correspond to the distributions of the *candidate process* induced by multiple assessments.

Having obtained an adequate model in terms of assessments inducing time-consistent multiple priors, we solve the problem by virtue of *minimax Snell envelopes* as introduced in [Riedel, 09] and obtain our main result: The ambiguous version of Theorem 3.1 in [Presman & Sonin, 72] giving necessary and sufficient conditions for the solution to the ambiguous Best-Choice problem to consist of finitely many stopping islands. A *stopping island* is, intuitively speaking, a set of applicants, which, if observed to be better than all applicants interviewed before, are optimal to be chosen. The theorem furthermore characterizes these stopping islands. To understand the importance of such a theorem it has to be noted, that the "simple" Best-Choice problem is monotone and hence there is just one stopping island up to infinity. This monotonicity property does not hold any longer in the risky as well as in the ambiguous case.

Before turning to a mathematical formulation and solution to the issue, we should ask if the Best-Choice problem is worthwhile for applications or if it is just for theoretical considerations. [Stewart, 78] gives two examples, a third is given in [Gilbert & Mosteller, 66]; the fourth is the usual application thought of today:

- Selling a single item: You have your old car for sale but no information on the market price. Prospective buyers arrive in random order telling the amount willing to pay. Either you stop and sell your car to some buyer or you send him away.[3]

- Exploration of resources: You are exploring oil deposits in the Middle East. When you have found a deposit you either stop and exploit it or you go on exploring. If not exploiting a deposit, someone else will do so.

- Atomic bomb inspection programs: You try to maximize the probability of finding a repository where illegal weapons-usable plutonium is stored.[4]

[3] [Stewart, 78] argues that the first example should be considered in the context of poisson arrival times.

[4] This particular example is more intuitive when modeled as a two person game as e.g. in [Gilbert & Mosteller, 66]: You want to maximize the probability of finding the repository by choosing an appropriate stopping rule whereas your opponent tries to minimize this probability by choosing the appropriate random order of examined repositories.

- Optimal exercise of an option or other financial derivatives.

As we see, the application changes over time but the integral problem, and in particular the mathematical methods, remain the same.

Having discussed the importance of the problem in economics, the last question to answer is: What value is added when considering this problem in an ambiguous set-up? First, we substantially decrease model risk apparent in the risky setup of the problem as a probabilistic model regarding the number of applicants has to be chosen and this respective model might just be wrong. Secondly, having a look at financial markets nowadays, a lot of uncertainty is "going on" there. No clear cut probability distributions can be attached to derivatives as not enough information is available or volatility is hitting in too strongly. In this case, an ambiguous or, equivalently, coherent approach seems a valuable ansatz for solving problems as e.g. pricing derivatives. In particular, with no information available, expert judgement tends to favor worst case solutions being theoretically underpinned within our framework.

The article is structured as follows: The next section discusses the "simple" Best-Choice problem and related extensions. The third section introduces the Secretary problem with a risky number of applicants, first by showing distinct approaches to model the problem and then discussing a concrete model. The fourth section is the main part of this article: We first recall the approach to optimal stopping as set out in [Riedel, 09]. Then, we introduce a direct generalization of the risky Best-Choice Problem to an ambiguous one and show why this is not feasible. Thereafter, we model the ambiguous problem in terms of so called assessments and solve it by virtue of the multiple prior Snell envelope. The main result of this article generalizes the main result in [Presman & Sonin, 72] to the ambiguous case and gives necessary and sufficient conditions for the solution to the ambiguous Best-Choice problem to consist of finitely many stopping islands. The last section concludes.

2.2 The Fundamental Problem

Before turning to the ambiguous model, we briefly set out the "simple" Best-Choice problem with a fixed number of applicants $n \in \mathbb{N}$, known to the agent. This section is divided into two parts: the no-information and the full-information problem. We achieve optimal strategies and respective choice probabilities as well as asymptotic results.

2.2.1 The No-Information Best Choice Problem

[Gilbert & Mosteller, 66] restate the problem in the following fashion: Given an urn with n balls, each with a different number but the range of numbers not known to the agent. The balls are drawn sequentially without replacement. The agent is reported the number on the ball but does not know which numbers are left in the urn, in particular has no information on the distribution of draws or even its range. Hence, the decision can only depend on *relative ranks* of draws. Equivalently, we could, as in [Gilbert & Mosteller, 66], just report the current rank to the agent. When reported the current rank, the agent must choose between keeping the current ball or continuing drawing. The problem is to maximize the probability of choosing the ball with the largest number among all n balls, or in other words to stop at the true maximum of the sequence. Equivalently, we endow the agent with a utility function only accounting for the best and the agent has to maximize expected utility. We will now make the problem rigorous.

Definition 2.2.1. (a) *Let Y_i denote the* relative rank *of the i^{th} applicant among the first i, $i \leq n$. Let $\bar{Y}_i$ denote its* absolute rank *among all n applicants.*
(b) *We call applicant i a* candidate *(or current maximum), if $Y_i = 1$. We call i the* true maximum *(or the best) if $\bar{Y}_i = 1$.*

More formally: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some arbitrary underlying probability space, $\bar{Y}_i : \Omega \to \{1, \ldots, n\}$, $i = 1, \ldots, n$, the absolute rank and $Y_i : \Omega \to \{1, \ldots, i\}$, $i = 1, \ldots, n$, the relative rank. Define the filtration $(\mathcal{F}_i)_{i \leq n}$ by $\mathcal{F}_i := \sigma(Y_1, \ldots, Y_i)$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Having in mind the intuition of a filtration as information process, $(\mathcal{F}_i)_{i \leq n}$ states: Upon arrival of applicant i, the agent can only observe her relative rank. In particular $\bar{Y}_i$ is not $\mathcal{F}_i$-measurable for $i < n$.

Define a (random) utility function $u : \{1, \ldots, n\} \times \Omega \to \mathbb{R}$ from stopping at applicant i for the agent as follows:

$$u_i = \begin{cases} 1 & \text{if } \bar{Y}_i = 1, \\ 0 & \text{else.} \end{cases}$$

As we see, $u_i \notin \mathcal{F}_i$, i.e. $(u_i)_{i \leq n}$ is not adapted and hence not an admissible payoff process for our problem.[5] The natural way to introduce an *adapted payoff process* built on this utility function is to consider its projection on $(\mathcal{F}_i)_{i \leq n}$, i.e. its *conditional expectation*. Hence, we define the adapted payoff process $(X_i)_{i \leq n}$ from stopping at applicant i upon observing by

$$X_i := \mathbb{E}^{\mathbb{P}} \left[u_i \middle| \mathcal{F}_i \right] = \mathbb{P}(\bar{Y}_i = 1 | Y_1, \ldots Y_i, i) = \mathbb{P}(\bar{Y}_i = 1 | Y_i, i),$$

where the last equation reflects the Marcovian nature of the problem. In words, X_i is the expected payoff from stopping at applicant i or, in other terms, the probability of applicant i being the best given her current rank. Let $\mathbf{T}$ denote the set of all stopping times, i.e. all mappings $\tau : \Omega \to \{1, \ldots, n\}$ such that $\{\tau \leq i\} \in \mathcal{F}_i$, then the *no-information Best Choice problem* is defined by its *value function* $(V_i)_{i \leq n}$:

Remark 2.2.2 (Agent's Problem)**.** *For $n \geq i \geq 0$ the value function $V :=$*

[5] Intuitively, to evaluate $u(i, \cdot)$ we need all information up to the last applicant n.

$(V_i)_{i\leq n}$ *of the problem is given by*

$$\begin{aligned} V_i := \max_{\tau\in\mathbf{T},\tau\geq i} \mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_i] &= \max_{\tau\in\mathbf{T},\tau\geq i} \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}[u_\tau|\mathcal{F}_\tau]\big|\,\mathcal{F}_i\right] \\ &= \max_{\tau\in\mathbf{T},\tau\geq i} \mathbb{E}^{\mathbb{P}}[u_\tau|\mathcal{F}_i] \\ &= \max_{\tau\in\mathbf{T},\tau\geq i} \mathbb{P}(\bar{Y}_\tau = 1|\mathcal{F}_i) \qquad (2.1) \\ &= \max_{\tau\in\mathbf{T},\tau\geq i} \mathbb{P}(\bar{Y}_\tau = 1|Y_i, i). \end{aligned}$$

Note that V_i is an $\mathcal{F}_i$-measurable random variable.

Proposition 2.2.3. *Equation (2.1), the Best-Choice problem, is solved by the smallest optimal stopping time:*

$$\tau^* := \min_i\{i \geq s^*|Y_i = 1\} \wedge n,$$

where s^* *solves*

$$\sum_{k=s^*}^{n-1} \frac{1}{k} \leq 1 < \sum_{k=s^*-1}^{n-1} \frac{1}{k}.$$

Proof. As stated in [Neveu, 75], Section VI.1, the value function $(V_s)_{s\leq n}$ of an optimal stopping problem satisfies the *Bellman equation* and hence

$$V_s = \max\left\{\underbrace{\mathbb{P}(\bar{Y}_s = 1|\mathcal{F}_s)}_{=X_s}; \mathbb{E}^{\mathbb{P}}[V_{s+1}|\mathcal{F}_s]\right\}.$$

for $s < n$ and $V_n := u_n = X_n$. Let's assume interviewing the s^{th} applicant, $s \leq n$. If $s = n$, we always stop as there is no better to come even if she is not ranked first. If she is not a candidate, i.e. $Y_s > 1$, and $s < n$, the value is given by

$$\begin{aligned} V_s &= \max\{\mathbb{P}(\bar{Y}_s = 1|Y_s > 1, s \leq n); \mathbb{E}^{\mathbb{P}}[V_{s+1}|Y_s > 1, s \leq n]\} \\ &= \mathbb{E}^{\mathbb{P}}[V_{s+1}|Y_s > 1, s \leq n] = \max_{\tau > s} \mathbb{P}[\bar{Y}_\tau = 1|Y_s > 1, s \leq n], \end{aligned}$$

since $Y_s > 1$ implies $\bar{Y}_s > 1$ $\mathbb{P}$-a.s. Hence, an applicant not being a candidate is never accepted and stopping does not occur at s.

Now assume s to be a candidate. In this case $\mathbb{P}(\bar{Y}_s = 1|Y_s) > 0$ and hence, if $s < n$,

$$V_s = \max\{\mathbb{P}(\bar{Y}_s = 1|Y_s = 1); \mathbb{E}^{\mathbb{P}}[V_{s+1}|Y_s = 1]\}$$

is non-trivial. By the *principle of backwards induction*, we stop at s, if the probability of applicant s being best exceeds that of choosing the best applicant from applicant $s+1$ onwards, i.e.

$$\mathbb{P}(\bar{Y}_s = 1|Y_s = 1) \geq \mathbb{E}^{\mathbb{P}}[V_{s+1}|Y_s = 1]. \tag{2.2}$$

In other words, the payoff at s exceeds the conditional expected payoff from going on with optimal stopping strategy. We have

$$\mathbb{P}(\bar{Y}_s = 1|Y_s = 1) = \frac{s}{n}.$$

This term is increasing in s, i.e. the later we observe a candidate, the higher the probability that she is best. The second part of the value function above, i.e. the probability of winning with the best strategy from $s+1$ onwards, is, by monotonicity of probability measures, decreasing in s. *Hence, the optimal strategy is of the form: pass the first s^* draws and take the first candidate thereafter.*[6]

Given this form of an optimal strategy, we now compute $\mathbb{E}^{\mathbb{P}}[V_{s^*+1}|\mathcal{F}_{s^*}]$ as the probability of winning with the optimal strategy when rejecting s applicants. By combinatoric considerations, we have for all k, and $s^* \leq k$

$$\mathbb{P}(\bar{Y}_k = 1) = \frac{1}{n}, \quad \mathbb{P}(Y_{s^*+1} > 1, \ldots, Y_{k-1} > 1) = \frac{s^*}{k-1},$$

meaning that the relatively best applicant in $1, \ldots, k-1$ is in $1, \ldots, s^* - 1$. Hence, by independence,

$$\mathbb{P}(Y_{s^*+1} > 1, \ldots, Y_{k-1} > 1 \wedge \bar{Y}_k = 1) = \frac{s^*}{n(k-1)}.$$

[6] In the process of modeling the problem in the risky and the ambiguous setup, we see that this monotonicity property does not necessarily hold when n is not deterministic.

Summing up, we achieve the probability of accepting the best applicant when accepting the first candidate after applicant s^*, i.e. the probability of winning with a strategy of the optimal type:

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}[V_{s^*+1}|\mathcal{F}_{s^*}]\\ =\ & \mathbb{P}\left(\bigcup_{k=s^*+1}^{n}\{Y_s^*+1>1,\ldots,Y_{k-1}>1\wedge\bar{Y}_k=1\}\right)=\sum_{k=s^*+1}^{n}\frac{s^*}{n(k-1)}.\end{aligned}$$

By equation (2.2), we have the optimal s, say s^*, to satisfy

$$\frac{s}{n}\geq\frac{s}{n}\sum_{k=s+1}^{n}\frac{1}{k-1}\quad\wedge\quad\frac{s-1}{n}<\frac{s-1}{n}\sum_{k=s}^{n}\frac{1}{k-1},$$

or, equivalently, s^* solves

$$\sum_{k=s^*}^{n-1}\frac{1}{k}\leq 1<\sum_{k=s^*-1}^{n-1}\frac{1}{k}.$$

□

Remark 2.2.4. *The intuition of the last inequality is immediate: The expected number of candidates following* s^*-1 *has to be at least* 1*, whereas the expected number of candidates following the one at* s^* *has to be less than one, having in mind that the last candidate is the best applicant. Intuitively, a strategy that passes the first* s^* *observations may fail if the best applicant already appears among the first* s^* *ones or if between the* $(s^*+1)^{st}$ *and the best applicant there is candidate who is then mistakenly chosen.*

Approximate results for s^* are available: For large n, we can use the Euler approximation and obtain

$$\mathbb{P}(\bigcup_{k=s}^{n}\{Y_s>1,\ldots,Y_{k-1}>1\wedge\bar{Y}_k=1\})\approx\frac{s}{n}\ln\frac{n}{s}.$$

Maximizing the last term yields $s^*=\frac{n}{e}$ and a corresponding probability of choosing the best of $\frac{1}{e}$, where e denotes the Euler constant.

Remark 2.2.5 (On Snell envelopes). *The foregoing proof was explicitly achieved in terms of backward induction via the Bellman equation. This is, however, just an explicit way of solving optimal stopping problems in terms of* Snell envelopes $(U_i)_{i\leq n}$.[7] *For the sake of completeness, we briefly reconsider the foregoing proof: For an adapted processes* $(X_i)_{i\leq n}$, *the minimal optimal stopping time is given by* $\tau^* = \inf\{i \geq 0 | X_i \geq U_i\}$, *where the Snell envelope* $(U_i)_{i\leq n}$ *for the no-information Best-Choice problem is recursively defined by*

$$\begin{aligned} U_n &:= X_n = \mathbb{P}(\bar{Y}_n = 1|\mathcal{F}_n) = 1_{\{Y_n=1\}}(Y_n) \\ U_i &:= \max\left\{X_i; \mathbb{E}^{\mathbb{P}}[U_{i+1}|\mathcal{F}_i]\right\} \\ &= \max\left\{\mathbb{P}(\bar{Y}_i = 1|\mathcal{F}_i); \mathbb{E}^{\mathbb{P}}[U_{i+1}|\mathcal{F}_i]\right\} \end{aligned}$$

for $i < n$. *We see that this is just the Bellman equation.*[8] *Having in mind – as already extensively used – that the payoff only depends on the observed rank of the applicant, we evaluate the distinct parts of the Snell envelope:*

$$\begin{aligned} \mathbb{P}(\bar{Y}_i = 1|Y_i > 1) &= 0, \\ \mathbb{P}(\bar{Y}_i = 1|Y_i = 1) &= \frac{i}{n}, \\ \mathbb{E}^{\mathbb{P}}[U_{i+1}|\mathcal{F}_i] &= \frac{i}{n}\sum_{k=i}^{n-1}\frac{1}{k}, \end{aligned}$$

where the last equation is shown in the foregoing proof. Hence,

$$U_i = \max\left\{\frac{i}{n}\mathbb{I}_{\{Y_1=1\}} + 0\mathbb{I}_{\{Y_1>1\}}; \frac{i}{n}\sum_{k=i}^{n-1}\frac{1}{k}\right\}$$

and we obtain as smallest optimal stopping rule

$$\begin{aligned} \tau^* &= \min\left\{i \geq 1 \;\middle|\; \frac{i}{n} \geq \frac{i}{n}\sum_{k=i}^{n-1}\frac{1}{k} \quad \wedge \quad Y_i = 1\right\} \\ &= \min\left\{i \geq s \;\middle|\; s = \arg\min_{t\leq n}\left\{\frac{t}{n} \geq \frac{t}{n}\sum_{k=t}^{n-1}\frac{1}{k}\right\} \quad \wedge \quad Y_i = 1\right\} \end{aligned}$$

[7] The theory of Snell envelopes will be discussed in more detail in course of this article.

[8] In particular, we have $(U_i)_{i\leq n} = (V_i)_{i\leq n}$.

$$= \quad \min\left\{ i \geq s \quad \middle| \quad \sum_{k=s}^{n-1} \frac{1}{k} \leq 1 < \sum_{k=s-1}^{n-1} \frac{1}{k} \quad \wedge \quad Y_i = 1 \right\}$$
$$\overset{n\to\infty}{\approx} \quad \min\left\{ i \geq \frac{1}{e} \quad \middle| \quad Y_i = 1 \right\},$$

a solution, which of course equals our result in Proposition 2.2.3.

2.2.2 The Full-Information Best-Choice Problem

For the sake of completeness we briefly consider the Best Choice problem without any ambiguity: We have full knowledge about the distribution of applicants' qualifications as well as the number of applicants is fixed.

Let therefore $(W_i)_{i\leq n}$ be a family of random variables, iid with distribution F each. The agent wants to maximize the probability of choosing the largest draw. Since only the largest counts and nothing else, we may without loss of generality set $F = U[0,1]$, the uniform distribution on the interval $[0,1]$. Now, we call the i^{th} draw a candidate if $W_i = \max_{k\leq i}\{W_k\}$.

In the no-information problem, s^* observations were needed to gain information. This is not the case here: If we, for example, observe the first draw very close to unity, the probability of larger observations is relatively small and hence, it might even be optimal, to accept the first draw. Thus, in the current problem, the decision is not only contingent on an applicant being a candidate or not and its time of observation but also on her current value. As we will see, the general rule turns out to be: Accept the first candidate exceeding some decision number corresponding to the qualification of that applicant.

The sequence of optimal decision numbers may be obtained by backward induction and only depends on the number of remaining draws: The last draw W_n is always accepted. Hence, the decision number is $b_1 = 0$. Assume that we have not accepted an applicant up to the second to last, $W_{n-1} = w$, and that $W_{n-1} = \max_{i\leq n-1} W_i$. Recall, that we will never accept non-candidates.

Then

$$\mathbb{P}(W_n \geq w) = 1 - w.$$

Hence, if $w \geq \frac{1}{2}$, we choose it; otherwise, we go on, maximizing the probability of winning in the second to last step. Hence, we have $b_2 = \frac{1}{2}$. In this sense, the decision numbers are just those values for candidates' qualifications that make the agent indifferent between stopping an going on just in the same fashion as in the proof of Proposition 2.2.3.

In general: Let b_i denote the decision number at the $(n-i+1)^{st}$ draw. Suppose we are faced with the $(n-i)^{th}$ draw $W_{n-i} = w$. In order to obtain the optimal indifference value, we have to equate both parts of the Snell envelope reducing to the following consideration: Expected payoff from accepting draw $n-i$, the left hand side of the Snell envelope, is given by

$$\mathbb{P}(W_{n-i} = \max_{t \leq n} W_t | W_{n-i} = x = \max_{t \leq n-i} W_t) = w^i,$$

whereas expected payoff from going on, the right hand side of the Snell envelope, is calculated by the following considerations: First, we observe that the optimal decision numbers are increasing, i.e. decrease as we go on with drawing since the probability of drawing a larger number decreases. Hence, in later draws, we would choose any draw exceeding b_{i+1}. Assume $W_{n-i} = w = b_{i+1}$:

- If there is only one such draw exceeding w, following our strategy, we choose it.

- If two those occur, say $y \geq z \geq w$, we have $\mathbb{P}(y) = \frac{1}{2}$.

- If three occur, say $y \geq z \geq x \geq w$, we have $\mathbb{P}(y) = \frac{1}{3}$.

- etc.

Hence,

$$\mathbb{E}^{\mathbb{P}}[U_{i+1}|\mathcal{F}_i] = \sum_{k=1}^{i} \frac{1}{k} \binom{i}{k} w^{i-k}(1-w)^k.$$

Equating these probabilities, we obtain $\forall i \; b_{i+1} = w$ as solution to

$$\begin{aligned} \sum_{k=1}^{i} \frac{1}{k} \binom{i}{k} w^{i-1}(1-w)^k &= w^i \\ \Leftrightarrow \quad 1 &= \sum_{k=1}^{i} \frac{1}{k} \binom{i}{k} \left(\frac{1-w}{w}\right)^k, \end{aligned}$$

leading to the following proposition on optimal stopping numbers:

Proposition 2.2.6. *The problem is solved by the following optimal stopping rule:*

$$\tau^* := \min\{i \geq 1 \quad | \quad W_i \geq b_{n-i+1} \quad \wedge \quad W_i = \max_{k \leq i} W_k\},$$

where the sequence $(b_i)_{i=1,\dots,n}$ *is achieved as above.*

Proof. Again, in terms of the Snell envelope approach:

$$\begin{aligned} U_n &= \mathbb{P}(W_n = \max_{i \leq n} W_i | \mathcal{F}_n) = 1_{\{W_n = \max_{i \leq n} W_i\}}(W_n), \\ U_i &= \max\{\mathbb{P}(W_i = \max_{k \leq n} W_k | \mathcal{F}_i); \mathbb{E}[U_{i+1}|\mathcal{F}_i]\} \\ \tau^* &= \min\{i \geq 1 \quad | \quad \mathbb{P}(W_i = \max_{k \leq n} W_k | \mathcal{F}_i) \geq U_i\} \\ &= \min\{i \geq 1 \quad | \quad W_i \geq b_{n-i+1} \quad \wedge \quad W_i = \max_{k \leq i} W_k\}. \end{aligned}$$

The last equality is seen as follows: At $W_i(\omega) = x$, $i < n$, the Snell envelope is given by

$$\begin{aligned} U_i \;:=\; & \max\Big\{\mathbb{P}(W_i = \max_{k \leq n} W_k | W_i = \max_{k \leq i} W_k); \\ & \qquad \mathbb{E}[U_{i+1}| W_i = \max_{k \leq i} W_k)]\Big\} \mathbb{I}_{\{W_i = \max_{k \leq i} W_k\}} \\ & + \max\left\{ 0; \underbrace{\mathbb{E}[U_{i+1}| W_i < \max_{k \leq i} W_k)\;]}_{>0} \right\} \mathbb{I}_{\{W_i < \max_{k \leq i} W_k\}} \end{aligned}$$

and the first term

$$\begin{aligned} & \max\left\{\mathbb{P}(W_i = \max_{k\le n} W_k | W_i = \max_{k\le i} W_k); \mathbb{E}[U_{i+1} | W_i = \max_{k\le i} W_k)]\right\} \\ = & \max\left\{ w^{n-i} \quad ; \quad \sum_{k=1}^{n-i} \frac{1}{k} \binom{n-i}{k} w^{n-i-k}(1-w)^k \right\}. \end{aligned}$$

Hence,

$$\begin{aligned} \tau^* &= \min\left\{ t \ge 1 \,\middle|\, W_t^{n-t} \ge \sum_{k=1}^{n-t} \frac{1}{k} \binom{n-t}{k} W_t^{n-t-k}(1-W_t)^k \right. \\ & \qquad \left. \wedge \quad W_t = \max_{i\le t} W_i) \right\} \\ &= \min\left\{ t \ge 1 | W_t \ge b_{n-t+1} \quad \wedge \quad W_t = \max_{i\le t} W_i) \right\}. \end{aligned}$$

□

2.2.3 A Further Refinement

One objection to the Secretary problem is that only the best choice counts. Let us briefly consider the case when utility is given by the actual value of the draw. Due to the fact that the agent obtains strictly positive utility even from draws that are not candidates, the optimal stopping rule does not hinge on an applicant being a candidate and hence it might be even optimal to accept a non-candidate.

Let $(W_i)_{i\le n}$ be sequentially and independently drawn from a distribution with density f. Define the Snell envelope U recursively by

$$\begin{aligned} U_n &:= W_n, \\ U_i &:= \max\{W_i; \mathbb{E}[U_{i+1}|\mathcal{F}_i]\}, \quad 1 \le i < n. \end{aligned}$$

Then, it is optimal to stop at $\tau^* := \min\{i \ge 1 | U_i = W_i\}$. We have

$$\mathbb{E}[U_n|\mathcal{F}_{n-1}] = \mathbb{E}[W_n|\mathcal{F}_{n-1}] = \mathbb{E}[W_n] = \int w f(w) dw := b_1.$$

Hence, the value of the problem at draw $n-1$ is given by

$$V_{n-1} = U_{n-1} = \max\{W_{n-1}; b_1\}.$$

At draw $n-1$, accept W_{n-1} if and only if $W_{n-1} \geq b_1$. Let b_2 denote the value of the problem of length 2 when going on, i.e. $b_2 = \mathbb{E}[U_{n-1}|\mathcal{F}_{n-2}]$. Then, we accept W_{n-2} at draw $n-2$ if and only if $W_{n-2} \geq b_2$: The family $(b_i)_i$ is a family of decision numbers as well as the value of not accepting the current draw. Having a look at draw W_{n-2}, we have to decide whether to go on or to accept that draw. We accept, if the draw exceeds the expected value b_2 of going on. How do we obtain this value? When going on, i.e. after rejecting W_{n-2}, we accept W_{n-1} if exceeding b_1, i.e. with probability $P(W_{n-1} \geq b_1) = \int_{b_1}^{\infty} f(w)dw$. In that case we obtain the expected value of W_{n-1} conditional on exceeding b_1. We reject W_{n-1} if smaller than b_1, i.e. with probability $P(W_{n-1} \leq b_1) = \int_{-\infty}^{b_1} f(w)dw$, in which case we obtain the expected value of the last draw, b_1. Formally,

$$\begin{aligned}
b_2 &= \mathbb{E}[U_{n-1}|\mathcal{F}_{n-2}] \\
&= \mathbb{P}(W_{n-1} \geq b_1)\mathbb{E}[W_{n-1}|W_{n-1} \geq b_1] + \mathbb{P}(W_{n-1} < b_1)\underbrace{\mathbb{E}[W_n|W_{n-1} < b_1]}_{=\mathbb{E}[W_n]=b_1} \\
&= \left(\int_{b_1}^{\infty} f(w)dw\right)\mathbb{E}[W_{n-1}|W_{n-1} \geq b_1] + b_1\int_{-\infty}^{b_1} f(w)dw \\
&= \int_{b_1}^{\infty} f(w)dw\frac{\int_{b_1}^{\infty} wf(w)dw}{\int_{b_1}^{\infty} f(w)dw} + b_1\int_{-\infty}^{b_1} f(w)dw \\
&= \int_{b_1}^{\infty} wf(w)dw + b_1F(b_1).
\end{aligned}$$

Now, going on recursively, we set $b_i := \mathbb{E}[U_{n-i+1}|\mathcal{F}_{n-i}]$ or equivalently $b_{n-j+1} = \mathbb{E}[U_j|\mathcal{F}_{j-1}]$. Then, we accept W_{n-i-1} if and only if $W_{n-i-1} \geq b_{i+1}$. We obtain the following recursive relation:

$$\begin{aligned}
b_{i+1} &= \mathbb{E}[U_{n-i}|\mathcal{F}_{n-i-1}] \\
&= \mathbb{P}(W_{n-i} \geq b_i)\mathbb{E}[W_{n-i}|W_{n-i} \geq b_i]
\end{aligned}$$

$$+\mathbb{P}(W_{n-i} \leq b_i) \underbrace{\mathbb{E}[U_{n-i+1}|\mathcal{F}_{n-i} \wedge \{W_{n-i} \leq b_i\}]}_{\mathbb{E}[U_{n-i+1}|\mathcal{F}_{n-i}]=b_i}$$

$$= \int_{b_i}^{\infty} wf(w)dx + b_i F(b_i).$$

Since b_{n-j+1} is the expected value of rejecting the draw at $j-1$, the value function (equalling the Snell envelope) becomes $U_i = \max\{W_i, b_{n-i}\}$. For the optimal stopping time, we have

$$\tau^* = \min\{i \geq 1 | U_i = W_i\} = \min\{i \geq 1 | W_i \geq b_{n-i}\}$$

with $(b_i)_i$ achieved recursively as above.

2.3 The No-Information Problem with a Risky Number of Objects

The major contribution of this article is the extension of the Best-Choice problem to an ambiguous number of applicants taking care of model risk. As a first step, we relax the assumption of a known number of applicants and consider the case of a risky number, i.e. with a given probability distribution on the number of applicants. Now, in addition to the risk of missing out on the best applicant in the setup with a fixed number, the agent is faced with the danger of waiting too long and being surprised by having no further choices.

2.3.1 A Review

There are several approaches to introduce risk about the number of applicants. We briefly summarize two and elaborately discuss the model that underlies this article. From a mathematical point of view, these approaches are just alternative ways to pose the problem. However, from an economic standpoint, they are quite different: which ansatz to prefer is a decision based on faith in which parameters can possibly be known.

[Stewart, 81] discusses both approaches: First, one might just assume the number of applicants being a random variable N with known distribution $\mu \in \mathcal{M}(\mathbb{N})$. All other assumptions in the simple model preserve. In particular, applicants arrive at deterministic times. This approach is followed in many articles: [Presman & Sonin, 72] apply a Snell envelope approach to the Marcov chain of candidates. They show, that the optimal solution of the problem is qualitatively different from the fixed-horizon setting: Distinct *stopping islands* may emerge, i.e. choosing a candidate may be optimal within some interval of applicants followed by an interval, where accepting is not optimal, again followed by an interval, where stopping is optimal, and so on. Intuitively, *in course of the application process, data is gathered about the actual number of applicants that makes stopping at a candidate not optimal even though it would have been optimal at an earlier stage with less information.* The reason for multiple stopping islands is owed to the fact that the problem is not monotone any longer in case of risk. [Presman & Sonin, 72] show that we still may use non-randomized stopping rules[9] and, moreover, give sufficient conditions to ensure *single island* rules.

[Gianini-Pettitt, 79] use the same approach to treat the problem of minimizing the expected rank. [Rasmussen & Robbins, 75] and [Rasmussen, 75] also follow this approach for a bounded random variable N with known distribution and obtain results for increasing bound. However, they mistakenly obtain a single island rule to be optimal for all distributions, contradicting [Presman & Sonin, 72]. [Irle, 80] explicitly states a counterexample to this single-island-statement and shows an algorithm to compute stopping islands. Furthermore, a monotonicity condition is achieved as a sufficient condition for optimality of a single island stopping rule. [Petrucelli, 83] gives sufficient

[9]Here, we distinguish between non-randomized and randomized stopping times: The former ones are just adapted integer valued random variables, the latter ones are distribution valued random variables, i.e. at some point in time, they do not specify whether to stop or not but which distribution on stopping or not to choose.

and necessary conditions on the distribution of N to obtain optimality of single island rules. Moreover, [Petrucelli, 83] shows that virtually any family of sets in $\mathbb{N}$ can be achieved as stopping islands of an optimal stopping rule by appropriately selecting a distribution. [Petrucelli, 83] shows for bounded distributions that only finitely many stopping islands are possible, i.e. there exists an integer such that stopping will occur at the next candidate.

In [Samuel-Cahn, 96] and [Samuel-Cahn, 95], the author goes a slightly different route: The problem is assumed with a fixed number n of applicants but with a random freeze M. Hence, we have a problem with random number of applicants $N := n \wedge M$. For both settings, the full-information as well as the no-information model, a sufficient and necessary condition on the distribution of M implying optimality of a single island rule is derived. It is shown that the setup with random horizon is equivalent to the simple problem with horizon equalling the upper bound of the distribution and a discount on payoffs induced by the distribution of M. In case of an optimal single island rule, stopping occurs earlier than in the fixed horizon case du to the discounting. [DeGroot, 68] considers a partial-information model with random horizon.

A second approach doubts arrival times being deterministic: An applicant is selected within a fixed time horizon but arrival times are randomly distributed. [Gnedin, 96] applies a planar poisson process on $[0,1]\times]-\infty,0]$ to the full-information problem. [Bruss, 84] and [Bruss & Samuels, 87] use the following model: Let arrival times be independently and identically distributed on the fixed interval $[0,t]$, e.g. by a poisson process. Let the overall random number N of applicants be independent of the arrival times but with an unknown distribution. Then, knowledge of the arrival time distribution fully compensates for ignorance of the number of applicants: The probability of choosing the best is the same as for the simple problem, i.e. e^{-1}. Whereas in case of fixed arrival times and a known distribution of N, the probability of choosing the best is significantly decreased. Of course, we may again achieve multiplicity of stopping islands. [Stewart, 81] uses an intermediate route in assuming N being distributed via some prior and arrival times being i.i.d. exponential random variables with known parameter. Then, upon arrival of an applicant the belief about N is updated in a Bayesian manner. Given this posterior distribution of N, optimization takes place as in the very first approach. The idea underlying [Stewart, 81] is the incapability of achieving a correct prior distribution for N as also assumed in [Bruss & Rogers, 91]. Hence, [Stewart, 81] introduces the *non-informative* prior of N basically as some kind of “uniform distribution” on $\mathbb{N}$. Of course, this is not a proper distribution but the posterior is. The posterior has to be taken as an additional state variable in the value function V, whereas in [Rasmussen, 75] it is sufficient to just truncate the prior distribution. However, the idea to compare the value from stopping and the value from going on optimally is the same. [Stewart, 81] shows that the optimal rule is of single island type and if large values of N are likely, the optimal selection probability approaches that in the fixed horizon problem, e^{-1}. This observation is of particular importance: *In the deterministic case, misspecification of the number of applicants leads to severe consequences as model risk is a serious issue. Within Stewart's model an exact estimation of N is not needed but results are quite stable. Hence, model risk is considerably smaller in this setup*: [Stewart, 81] achieves robustness results showing that even for a relatively small number of applicants as well as for an erroneous specification of the parameter of exponential arrival rate (up to factor 2), the selection probability following the specified rule is still quite close to the optimal case.

The “formal Bayes rule” obtained in [Stewart, 78] coincides with the optimal rule in the infinite horizon problem. [Bruss & Samuels, 87] extend this insight and achieve that, for any loss function with finite risk in the infinite secretary problem, the rule that is optimal in the infinite secretary problem is formal Bayes in the sense of [Stewart, 78].

In the present article, we follow the first approach with *fixed arrival times* *but unknown horizon*, i.e. number of applicants. Based on this, we build a model for an ambiguous number of applicants. There are as many good reasons for the first as for the second approach. The main point for the latter is that reliability theory indicates arrival times of uncertain events being exponentially distributed. On the other hand, when considering financial markets or atomic bomb inspection programs, we know at which times we have a look at the market or at a potential repository but are not sure about the time horizon.

2.3.2 A Specific Model

Let us now extend the simple no-information Best-Choice problem to the case of a random number N of applicants with known distribution $\mu \in \mathcal{M}(\mathbb{N})$, $\mu(n) := \mu(N = n)$, where $\mathcal{M}(N)$ denotes the set of distributions on $\mathbb{N}$. As we want to maximize the probability of choosing the best applicant, being a candidate is a necessary condition to be accepted. Hence, the idea is to only consider the *candidate process* instead of the applicant process. This approach is found in [Presman & Sonin, 72].

Recall that Y_k denotes the relative rank of applicant k among the first k applicants, whereas $\bar{Y}_k$ denotes the absolute rank of applicant k among all. For $k > N$, we set $Y_k = \bar{Y}_k = \infty$. First, we intuitively obtain the relevant probabilities, then we rigorously introduce the probabilistic model at hand. Given $\mu \in \mathcal{M}(\mathbb{N})$, $k \in \mathbb{N}$, we have

$$\begin{aligned}
\mathbb{P}^\mu(Y_k = 1) &= \mu(N \geq k)\mathbb{P}^\mu(Y_k = 1|N \geq k) \\
&\quad + \mu(N < k)\mathbb{P}^\mu(Y_k = 1|N < k) \\
&= \mu(N \geq k)\mathbb{P}^\mu(Y_k = 1|N \geq k) \\
&= \mu(N \geq k)\mathbb{P}^\mu(Y_k = 1 \quad \wedge \quad N \geq k)\frac{1}{\mu(N \geq k)} \\
&= \mathbb{P}^\mu(\cup_{s=k}^{\infty}\{Y_k = 1 \quad \wedge \quad N = s\})
\end{aligned}$$

$$\begin{aligned}
&= \sum_{s=k}^{\infty} \mathbb{P}^{\mu}(Y_k = 1 \quad \wedge \quad N = s) \\
&= \sum_{s=k}^{\infty} \mathbb{P}^{\mu}(Y_k = 1 | N = s)\mu(s) \\
&= \frac{1}{k}\sum_{s=k}^{\infty} \mu(s) = \frac{1}{k}\mu(N \geq k).
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}^{\mu}(\bar{Y}_k = 1) &= \sum_{s=k}^{\infty} \frac{1}{s}\mu(s). \\
\mathbb{P}^{\mu}(\bar{Y}_k = 1 | Y_k = 1) &= \frac{\mathbb{P}^{\mu}(\bar{Y}_k = 1 \quad \wedge \quad Y_k = 1)}{\mathbb{P}^{\mu}(Y_k = 1)} = \frac{\mathbb{P}^{\mu}(\bar{Y}_k = 1)}{\mathbb{P}^{\mu}(Y_k = 1)} \\
&= \frac{\sum_{s=k}^{\infty} \frac{1}{s}\mu(s)}{\frac{1}{k}\mu(N \geq k)} = \frac{k}{\mu(N \geq k)} \sum_{s=k}^{\infty} \frac{\mu(s)}{s}.
\end{aligned}$$

The last term is the value from stopping at applicant k being a candidate. Note that this is just the μ-expectation of the value from stopping at applicant k being a candidate in the "simple" problem:

$$\mathbb{E}^{\mu}\left[\frac{k}{N}\right] = \mathbb{E}^{\mu}\left[\frac{k}{N}\middle| N \geq k\right] \frac{1}{\mu(N \geq k)} = \frac{k}{\mu(N \geq k)} \sum_{s=k}^{\infty} \frac{\mu(s)}{s}.$$

Following [Presman & Sonin, 72], we now separate the the payoff and the applicant process. The latter is then refined to the *candidate process* $(\xi_i)_i$, where $\xi_i = k$ means that the i^{th} candidate is the k^{th} applicant. We now compute the distribution of this process. From the problem with fixed number $n \geq k > l$ of applicants, i.e. $\mu(N = n) = 1$, we have

$$\begin{aligned}
&\mathbb{P}^{\mu}(Y_k = 1 \cap Y_{k-1} > 1 \cap \cdots \cap Y_{l+1} > 1 | Y_l = 1) \\
=\ &\frac{\mathbb{P}^{\mu}(Y_k = 1 \cap Y_l = 1 \cap \{Y_j > 1 | j \in \{l+1, \cdots, k-1\}\})}{\mathbb{P}^{\mu}(Y_l = 1)} \\
=\ &\frac{l}{k} \cdot \frac{1}{l} \cdot \frac{l}{l+1} \cdot \ldots \cdot \frac{k-2}{k-1} = \frac{l}{k(k-1)}
\end{aligned}$$

This is the transition probability $\mathbb{P}^{\mu}(\xi_i = k | \xi_{i-1} = l)$ of the homogenous Markov chain $(\xi_i)_i$ with state space $\{1, \ldots, n\}$ in case of a fixed $N = n$. For a random number N of applicants with distribution μ, the distribution $\mathbb{P}^{\mu}$ of

$(\xi_i)_i$ is fully characterized by the *initial distribution*

$$\mathbb{P}^\mu \circ \xi_1^{-1} = \mathbb{I}_{\{\xi_1=1\}}$$

and the *transition kernel* p^μ given by

$$\begin{aligned}
& p^\mu(l,k) := \mathbb{P}^\mu(\xi_i = k|\xi_{i-1} = l) \\
= \;& \mathbb{P}^\mu(Y_k = 1 \cap \{Y_j > 1 | j \in \{l+1,\cdots,k-1\}\}|Y_l = 1) \\
= \;& \frac{\mathbb{P}^\mu\left(Y_k = 1 \cap Y_l = 1 \cap \{Y_j > 1 | j \in \{l+1,\cdots,k-1\}\} \cap N \geq k)\right) + 0}{\mathbb{P}^\mu(Y_l = 1 \cap N \geq l) + 0} \\
= \;& \frac{\mathbb{P}^\mu\left(\begin{array}{l} Y_k = 1 \cap Y_l = 1 \\ \cap\{Y_j > 1 | j \in \{l+1,\cdots,k-1\}\} \end{array} \middle| N \geq k\right)\mu(N \geq k)}{\mathbb{P}^\mu(\{Y_l = 1 | N \geq l)\mu(N \geq l)} \\
= \;& \begin{cases} \frac{l\mu(N\geq k)}{k(k-1)\mu(N\geq l)} & l < k < \infty, \\ 0 & l \geq k, \end{cases} \qquad \forall k \geq 2, \\
& p^\mu(\infty,\infty) = 1, \quad p^\mu(l,\infty) = \mathbb{P}^\mu(\bar{Y}_l = 1 | Y_l = 1) = \frac{l}{\mu(N \geq l)} \sum_{s=l}^{\infty} \frac{\mu(s)}{s}.
\end{aligned}$$

Having entirely characterized the candidate process $(\xi_i)_i$ by its distribution, we turn to the appropriate *payoff function* given by the probability of a candidate being the best: given $\xi_i = k$, set $g^\mu(k) := p^\mu(k,\infty) = \frac{k}{\mu(N\geq k)} \sum_{s=k}^{\infty} \frac{\mu(s)}{s}$. Theorem 2.1 in [Presman & Sonin, 72] shows that it suffices to only consider the Markov chain of candidates to solve the Best-Choice problem and neglect all elements with zero payoff, the "non-candidates", as those will never be chosen.

Definition 2.3.1 (Best-Choice problem under risk)**.** *Following the approach in [Presman & Sonin, 72], the Best-Choice problem with a risky number of applicants for a distribution $\mu \in \mathcal{M}(\mathbb{N})$ is given by the candidate process $(\xi_i)_i$ with transition kernel $p^\mu(\cdot,\cdot)$ and payoff function $g^\mu(\cdot)$ as above.*

The Snell envelope for optimally stopping the process $(g^\mu(\xi_i))_i$ is, at $\xi_i =$

k, given by

$$\begin{aligned} U^{\mu}_{\xi_i} &:= \max\left\{ g^{\mu}(\xi_i) \quad ; \quad \mathbb{E}^{\mu}[U^{\mu}_{\xi_{i+1}}|\mathcal{F}_k] \right\} \\ &= \max\left\{ \frac{k}{\mu(N\geq k)}\sum_{s=k}^{\infty}\frac{\mu(s)}{s} \quad ; \quad \sum_{s=k+1}^{\infty} \underbrace{\frac{k}{s(s-1)}\frac{\mu(N\geq s)}{\mu(N\geq k)}}_{=:p(k,s)} U^{\mu}_{s} \right\}, \end{aligned}$$

where the first term is the probability of choosing the best when stopping at the current candidate and the second term the probability of choosing the best by going on and optimally decide whether to stop at the next candidate. Due to homogeneity, only the value k of the variable ξ_i is of interest. This is also the reason for our modest change of notation: *We now write $U^{\mu}_{\xi_i}$ instead of U^{μ}_i for the Snell envelope and thereby make explicit that the problem we consider now is not stated in terms of the applicant process any longer but in terms of the sub-process of candidates.*

E.g. in [Riedel, 09] it is mentioned that $U^{\mu}_{\xi_i}$ is the smallest supermartingale exceeding the payoff process $g^{\mu}(\xi_i) := \frac{\xi_i}{\mu(N\geq\xi_i)}\sum_{s=\xi_i}^{\infty}\frac{\mu(s)}{s}$ and hence, it is optimal to accept candidate i being applicant $\xi_i = k$, whenever $U^{\mu}_k \leq g^{\mu}(k)$. In other terms, the *optimal stopping set* Γ^{μ} is given by

$$\Gamma^{\mu} = \{k | U^{\mu}_k \leq g^{\mu}(k)\}$$

and the *smallest optimal stopping time* is

$$\tau^* := \min\{i|\xi_i \in \Gamma^{\mu}\} = \min\{i|U^{\mu}_{\xi_i} \leq g^{\mu}(\xi_i)\},$$

where stopping at $\tau = i$ means stopping at the i^{th} candidate, *not* at the i^{th} applicant. However, $k \in \Gamma^{\mu}$ means to stop at applicant k being a candidate. For more intuition on Γ^{μ} consider Section 2.4.2.

The main virtue of this article is to generalize the above model to the case of an ambiguous number of applicants and to find conditions for finitely many stopping islands, i.e. to generalize the results in [Presman & Sonin, 72].

2.4 The No-Information Problem with an Ambiguous Number of Objects

There are basically two approaches to model ambiguity: On one side, multiple priors or equivalently coherent risk measures, on the other side Choquet integrals. It is not just a matter of taste which ansatz to apply but what shall be modeled. The first route to model behavior under uncertainty is by virtue of *non-additive probabilities* or *capacities* applying Choquet integration for evaluation of "*expected utility*". As not followed here, we just have brief a look at the intuition of this approach. A valuable introduction to capacities and the related Choquet integral can be found in [Skulj, 01]. [Chateauneuf, 1991] gives economic content to the theory by presenting an "expected utility theorem" for capacities by virtue of Choquet integrals serving as expected utility. Furthermore, a connection to *multiple priors* expected utility on convex sets of additive probabilities is drawn and hence an equivalence to the multiple priors framework: An agent is ambiguity averse in the sense of [Gilboa & Schmeidler, 89] if and only if the ruling non-additive probability is convex. The gap of the total capacity to unity is a *measure for ambiguity*; measures for *convexity* of the capacity are in this sense measures for ambiguity aversion. A concave capacity is connected to an agent who is ambiguity loving. In case of an additive probability, there is no ambiguity or, more precisely, ambiguity does not matter to the agent. This is equivalent to the case of a unique prior in the multiple priors theory. Then the worst-case measure as obtained in [Chateauneuf, 1991] or in [Riedel, 09] for the distinct theories coincides with this unique prior additive probability. For a convex capacity, [Marinacci, 99] gives an explicit description of a worst-case measure, i.e. a probability distribution such that the integral with respect to that distribution and the Choquet integral coincide or, equivalently, the expected utilities coincide. [Gilboa, 87] generalizes the expected utility theorem for additive probabilities step by step to capacities and explicitly compares additive and non-additive theory. Further representations as well as equivalence results of convex capacities and multiple priors may be found in [Schmeidler, 86], [Schmeidler, 89] or [Yaari, 87]. In the latter reference, capacities are generated by monotone increasing distortions of additive probabilities. An introduction to Choquet integration and its relation to coherent risk measures can also be found in [Föllmer & Schied, 04].

The second approach is the *multiple priors* framework as introduced in [Gilboa & Schmeidler, 89]. They obtain a representation result for uncertainty averse agents' preferences in terms of *minimal expected payoff*, i.e. expected reward is calculated as the minimized expected value with respect to some set of prior distributions. [Epstein & Schneider, 03] extend this approach to a dynamic context where dynamic consistency leads to a recursive representation of utility: Assuming conditional preferences at each time-event pair to satisfy the axioms in [Gilboa & Schmeidler, 89] and the process of conditional preferences being dynamically consistent, the value function is obtained recursively. The notion of time-consistency, intuitively stating that the set of prior distributions is closed under pasting, is rigorously introduced in the next section.

The advantage of the approach in terms of capacities is to explicitly deliver a measure of ambiguity aversion. In our model below we make use of multiple priors. Therein an agent is assumed to be ambiguity averse and the expansion of the set of priors might possibly measure ambiguity averseness. [Riedel, 09] shows that this framework is adjuvant for optimal stopping problems. A further reason for the multiple priors framework is merely a question of belief: under the assumption of ambiguity aversion, is it easier to specify a unique capacity or to give a full range of additive probabilities that seem possible for the agent to rule the world? As we have seen above in the discussion of the literature on capacities, the Choquet approach and multiple priors framework are equivalent under mild conditions.

2.4.1 General Theory of Optimal Stopping with Multiple Priors

For multiple prior preferences, [Riedel, 09] derives a general theory of optimal stopping when the set of multiple priors is time-consistent. A recursive representation of the value function allows for a generalization of the Snell envelope approach in [Neveu, 75] and, hence, of the backward induction principle to ambiguous settings. As the Snell envelope is the smallest supermartingale dominating payoff in the risky case, we see that an appropriately generalized Snell envelope is the smallest *multiple prior supermartingale* with this property in the ambiguous case. As in the classical case, it is optimal to stop when the value of the *multiple prior Snell envelope* equals the payoff from stopping. Thereunto, [Riedel, 09] introduces a general theory of multiple prior (sub-/super-)martingales.

Intuition

We briefly recap the framework in [Riedel, 09]. Given an arbitrary underlying probability space $(\Omega, \mathcal{F}, \mathbb{P}_0)$ with filtration $(\mathcal{F}_t)_t$, a *multiple prior martingale* is a process $(M_t)_t$ that satisfies $M_t = \operatorname{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}}[M_{t+1}|\mathcal{F}_t]$ for some set $\mathcal{Q}$ of prior distributions all being assumed locally equivalent to $\mathbb{P}_0$. Hence, a minimax martingale is a submartingale for all distributions in $\mathcal{Q}$ and, in case of time-consistency, a martingale with respect to some *worst-case distribution* in $\mathcal{Q}$. In this sense, it is a *fair game* for an ambiguity averse agent who always expects nature to choose the worst distribution. Given time-consistency, the minimax Snell envelope is the *lower envelope* of the Snell envelopes with respect to priors in $\mathcal{Q}$ and it is the classical Snell envelope with respect to the worst case distribution in $\mathcal{Q}$. This amounts to the following main insight:

Remark 2.4.1. *Assuming time-consistency, the ambiguity averse agent behaves as the Bayesian expected utility maximizer given the worst case distribution in the set of priors $\mathcal{Q}$.*

This formalizes the precious intuition that ambiguity averse agents expect nature to be malevolent. Put another way: solving an optimal stopping problem reduces to finding the worst case distribution and then solving the problem as in the Bayesian setup.[10]

This insight simplifies the solution to the Best-Choice problem in the multiple priors framework: Whereas [Hill & Krengel, 91] need randomized stopping times in the uncertain case, [Abdel-Hamid et al., 82] as well as [Presman & Sonin, 72] have shown that it suffices to consider non-randomized rules in the Bayesian setup. Hence, we have:

Remark 2.4.2 (On randomized stopping rules)**.** *The optimal stopping time for the ambiguous Best Choice problem in case of time-consistent multiple priors is non-randomized.*[11]

Rigorous Set-up and Results

We now formally introduce the results in [Riedel, 09]. let $(\Omega, \mathcal{F}, \mathbb{P}_0)$ be a probability space with filtration $(\mathcal{F}_t)_t$. $\mathbb{P}_0$ serves a s a *reference distribution*. The time horizon might be finite or infinite. Given a bounded adapted payoff process $(X_t)_t$, the agent tries to maximize payoff by appropriately choosing a stopping time τ with respect to $(\mathcal{F}_t)_t$. The main assumption is that the distribution of $(X_t)_t$ is not entirely known but belongs to a convex, weakly compact set $\mathcal{Q}$ of measures that are (locally) equivalent to $\mathbb{P}_0$[12] or, equivalently, that the agent is an ambiguity averter in the sense of [Gilboa & Schmeidler, 89].

[10]This is actually not precisely the case for the ambiguous Best-Choice problem as we will see later: Given the worst case distribution the agent solves a problem that is just related to the Best-Choice problem under risk.

[11]Please recall the distinction between randomized and non-randomized stopping times: a non-randomized stopping time or just stopping time is a random variable τ such that $\{\tau \leq t\} \in \mathcal{F}_t$, whereas a randomized stopping time specifies a probability distribution whether to stop or not upon arrival at t.

[12]Note that $\mathcal{Q}$ consists of distributions of the payoff process $(X_t)_t$. The assumptions made so far ensure suprema and infima to be maxima and minima, respectively.

Hence, the agent has to solve

$$\max_{\tau} \min_{\mathbb{P} \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}}[X_\tau].$$

The following assumption is crucially needed for this approach to be feasible.

Assumption 2.4.3. *Let $\mathcal{Q}$ be time-consistent.*

The following definition is taken from [Riedel, 09]. Therein, equivalent definitions are discussed.

Definition 2.4.4 ([Riedel, 09], Assumption 4)**.** *A set of priors $\mathcal{Q}$ is said to be* time-consistent *if for all $\mathbb{P}$ and $\mathbb{Q}$ in $\mathcal{Q}$ and stopping times τ the "pasted" distribution $\mathbb{R}$ defined by virtue of*

$$\left.\frac{d\mathbb{R}}{d\mathbb{P}_0}\right|_{\mathcal{F}_t} := \begin{cases} p_t & \text{if } t \leq \tau, \\ \frac{p_\tau q_t}{q_\tau} & \text{else} \end{cases}$$

also belongs to $\mathcal{Q}$, where $\frac{d\mathbb{Q}}{d\mathbb{P}_0}$ denotes the Radon-Nikodym derivative with respect to $\mathbb{P}_0$ and p_i (resp. q_i) denotes the density process of $\mathbb{P}$ with respect to $\mathbb{P}_0$, i.e. $\forall\ i \in \mathbb{N}$

$$p_i := \left.\frac{d\mathbb{P}}{d\mathbb{P}_0}\right|_{\mathcal{F}_i}.$$

Intuitively, $\mathcal{Q}$ is assumed to be closed under pasting: at any time-event pair, combining marginals of $\mathbb{P} \in \mathcal{Q}$ with conditionals of other priors in $\mathcal{Q}$ has to be in $\mathcal{Q}$ again. This implies that $\mathcal{Q}$ is uniquely determined by the process of conditional one-step-ahead distributions. In the above definition, $\mathbb{R}$ is obtained as a distribution given by $\mathbb{P}$ up to time τ and $\mathbb{Q}$ thereafter

We now recall the mathematical concept crucial for our model:

Definition 2.4.5 ([Riedel, 09], Definition 1)**.** *Let $\mathcal{Q}$ be a time-consistent set of priors. Let $(M_t)_t$ be an adapted process with $\mathbb{E}^{\mathbb{P}}[M_t] < \infty$ $\forall \mathbb{P} \in \mathcal{Q}$ and $\forall t \in \mathbb{N}$. $(M_t)_t$ is called a* multiple prior (sub-, super-) martingale *with respect to $\mathcal{Q}$ if $\forall t \in \mathbb{N}$, it holds*

$$ess \inf_{\mathbb{P} \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}}[M_{t+1} | \mathcal{F}_t] = (\geq, \leq) M_t \quad a.s.$$

[Riedel, 09] shows that $(M_t)_t$ is a multiple prior submartingale if and only if $(M_t)_t$ is a submartingale for all $\mathbb{P} \in \mathcal{Q}$. $(M_t)_t$ is a multiple prior supermartingale if and only if there exists a $\mathbb{P}^* \in \mathcal{Q}$ such that $(M_t)_t$ is a $\mathbb{P}^*$-supermartingale. $(M_t)_t$ is a multiple prior martingale if and only if $(M_t)_t$ is a submartingale for all $\mathbb{P} \in \mathcal{Q}$ and there exists a $\mathbb{P}^* \in \mathcal{Q}$ such that $(M_t)_t$ is a $\mathbb{P}^*$-martingale. Hence, an ambiguity averse agent considers a game fair, if it is non-disadvantageous for all priors and fair for the worst case prior $\mathbb{P}^*$. For existence of this worst case distribution $\mathbb{P}^*$, time-consistency is crucially needed as it is achieved by pasting instantaneous worst case distributions recursively. It is shown in [Riedel, 09] that the Doob decomposition and the optional sampling theorem are still valid for minimax martingales given time-consistency. In a forthcoming article, we extend this notion and the respective results to the case of dynamic variational preferences or, equivalently, dynamic convex risk measures. The next theorem is the main result in [Riedel, 09]: Let first T be finite.

Definition 2.4.6 ([Riedel, 09], Theorem 1)**.** *The* multiple prior Snell envelope $U := (U_t)_t$ *of* $X := (X_t)_t$ *with respect to* $\mathcal{Q}$ *is defined recursively by virtue of* $U_T = X_T$ *and for all* $t < T$

$$U_t := \max\left\{ X_t, \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}}[U_{t+1}|\mathcal{F}_t] \right\}.$$

Theorem 2.4.7 ([Riedel, 09], Theorem 1)**.** *Let* $\mathcal{Q}$ *be time-consistent, then* U *is the smallest multiple prior supermartingale exceeding* X*.* U *is the value process of the optimal stopping problem under ambiguity, i.e.*

$$U_t = \operatorname*{ess\,sup}_{\tau\geq t} \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_t].$$

The smallest optimal stopping time is $\tau^* := \inf\{i \geq 0 | U_i = X_i\}$*.*

The proof follows an insightful idea: At each time-event pair, we calculate a one-step-ahead worst case distribution and then paste it with the worst case distribution already obtained from the following time period on.

Following this procedure recursively, we obtain a worst case distribution $\mathbb{P}^*$ of the payoff process $(X_t)_t$. As already stated, we have that the multiple prior Snell envelope U with respect to $\mathcal{Q}$ equals the Snell envelope $U^{\mathbb{P}^*}$ of the payoff process $(X_t)_t$ under $\mathbb{P}^*$. In this sense, the ambiguity averse agent behaves as the expected utility maximizer under a worst case distribution. We hence have a minimax theorem:

Proposition 2.4.8 ([Riedel, 09], Theorem 2)**.** *$(U_t)_t$ is the lower envelope of the Snell envelopes $(U_t^{\mathbb{P}})_t$ w.r.t. the priors $\mathbb{P} \in \mathcal{Q}$, and this envelope is attained by the worst case prior $\mathbb{P}^*$, i.e. $U_t = \operatorname{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} U_t^{\mathbb{P}} = U_t^{\mathbb{P}^*}$. More precisely, we have*

$$\begin{aligned} U_t = \operatorname*{ess\,sup}_{\tau\geq t} \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_t] &= \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} \operatorname*{ess\,sup}_{\tau\geq t} \mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_t] \\ &= \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} U_t^{\mathbb{P}} = U_t^{\mathbb{P}^*}. \end{aligned}$$

For sake of completeness, we state the results in [Riedel, 09] for the infinite horizon case, i.e. $T = \infty$.

Definition 2.4.9 ([Riedel, 09], Equation (6))**.** *The* value function $V := (V_t)_t$ *of the optimal stopping problem on $(X_t)_t$ is defined as*

$$V_t := \operatorname*{ess\,sup}_{\tau\geq t} \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_t].$$

Theorem 2.4.10 ([Riedel, 09], Theorem 3)**.** *V is the smallest multiple prior supermartingale with respect to $\mathcal{Q}$ exceeding X. V satisfies the Bellman equation*

$$V_t = \max\left\{X_t, \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}}[V_{t+1}|\mathcal{F}_t]\right\}$$

for all $t \geq 0$. The smallest optimal stopping time is given by $\tau^* := \inf\{i \geq 0 | V_i = X_i\}$ provided that $\tau^* < \infty$ a.s.

We can approximate infinite by finite horizon problems:

Proposition 2.4.11 ([Riedel, 09], Theorem 4)**.** *Denote by U^T the multiple prior Snell envelope of the optimal stopping problem of X with horizon T. Then $\lim_{T\to\infty} U_t^T = V_t$ for all $t \geq 0$, where $(V_t)_t$ denotes the infinite horizon value function.*

These results particularly show that the value function and the multiple prior Snell envelope coincide for an ambiguous problem. Moreover, the value function and the Snell envelope in the risky set-up are equal.

2.4.2 The Model

We introduce ambiguity on the number of applicants in the Best-Choice problem. In [Chudjakow & Riedel, 09], ambiguity comes into account by virtue of ambiguous orderings of the applicant process but with a fixed number of applicants, i.e. [Chudjakow & Riedel, 09] assume distinct sets of ordering distributions.

In [Hill & Krengel, 91], we see an extreme case of ambiguity: Basically nothing is known about the distribution of N, the number of applicants. However, we have to notice the formal difference between ambiguity therein and in the sense of [Riedel, 09]: In the former, ambiguity is introduced by non-uniqueness of priors $\mu \in \mathcal{M}(\mathbb{N})$ on the number of applicants. In the approach in [Riedel, 09], a prior is a distribution of the payoff process $(X_i)_i$. Thus, ambiguity in our ansatz comes into account by assuming a whole set $\mathcal{Q}$ of possible prior distributions of the payoff process.

In a first approach, the problem seems to be transforming a distribution $\mu \in \mathcal{M}(\mathbb{N})$ on the number N of applicants to a distribution of the payoff process $(X_i^\mu)_i$: Therefore, we first would have to appropriately define the payoff process $(X_i^\mu)_i$ given μ. Thereafter, it would suffice to give an initial distribution and a stochastic kernel to obtain a distribution of the payoff process as done in Section 2.3.2 mimicking [Presman & Sonin, 72]. We hence obtain a one-to-one mapping from the set of priors on the number of applicants to the set of priors on the payoff process and may solve the problem as in [Riedel, 09]. This approach directly generalizes the model in [Presman & Sonin, 72]. However, we will see that time-consistency turns out to be an integral problem. In a second approach, we find a remedy for the time-consistency issue when modeling ambiguity in terms of *assessments*.

Although the first approach does not immediately lead to a solution, we briefly consider it here as the calculations are the cornerstone for the second approach and we explicitly note impossibility of time-consistency in the first ansatz.

The Payoff Process and its Distribution for given $\mu \in \mathcal{M}(\mathbb{N})$

We briefly recall the setup leading to the appropriate payoff process in several steps. Let Y_k be the relative rank of applicant k within the first k applicants and $\bar{Y}_k$ its absolute rank among all. Set $Y_k = \infty$ for $k > N$. We again fix an underlying space $(\Omega, \mathcal{F}, \mathbb{P}_0)$ and define a filtration by virtue of $\mathcal{F}_k := \sigma(Y_1, \ldots, Y_k)$, $k \in \mathbb{N}$. Intuitively, $\mathcal{F}_k$ states whether applicant k is a candidate or not and in particular if $k \leq N$. *Payoff is unity if we have successfully chosen the best applicant and zero else.* However, the process

$$X_k := \left\{ \begin{array}{ll} 1 & \text{if } \bar{Y}_k = 1, \\ 0 & \text{else} \end{array} \right\} = \mathbb{I}_{\bar{Y}_k = 1}$$

is not $\mathcal{F}_k$-measurable since $\bar{Y}_k \notin \mathcal{F}_k$. Hence, the above definition does not yield an admissible payoff process. The intuitive reasoning is just that upon interviewing applicant k, we do not know if she is best among all. Otherwise, the problem would be equivalent to the parking problem, where the agent, upon observing an open lot, knows the utility that he gains from parking there. Hence, as in the classical case, the best the agent can do is to calculate the *conditional expected payoff* from accepting an applicant given the information available and use this as payoff process to be maximized. This payoff is equivalent to calculating the probability of an applicant being best among all. If the number of applicants is fixed at $N = n$, the payoff process $(X_k^{\delta_n})_k$ is

$$X_k^{\delta_n} := \mathbb{E}[\mathbb{I}_{\bar{Y}_k=1}|\mathcal{F}_k] = \mathbb{P}(\bar{Y}_k = 1|\mathcal{F}_k) = \begin{cases} \frac{k}{n} & \text{if } Y_k = 1, \\ 0 & \text{if } Y_k > 1. \end{cases}$$

By definition $X_k^{\delta_n} \in \mathcal{F}_k$. Hence, $(X_k^{\delta_n})_k$ is an admissible payoff process.

In case that the number of applicants N is a random variable with distribution $\mu \in \mathcal{M}(\mathbb{N})$, $\mu(N = s) =: \mu(s)$, the conditional probability that applicant k is best, and hence the payoff process, is given by

$$X_k^{\mu} := \mathbb{E}^{\mu}[\mathbb{I}_{\bar{Y}_k=1}|\mathcal{F}_k] = \begin{cases} \frac{k}{\mu(N\geq k)} \sum_{s=k}^{\infty} \frac{\mu(s)}{s} & \text{if } Y_k = 1, \\ 0 & \text{if } Y_k > 1. \end{cases}$$

The respective calculations are stated in Section 2.3.2.

As in [Presman & Sonin, 72] we separate the applicant process from the *payoff process* and w.l.o.g. reduce the former to the corresponding *candidate process* since non-candidates generate payoff zero and, hence, will never be chosen.

More formally: Consider the Markov chain $z_k := (Y_k, k)$ with payoff $g^{\mu}(z_k) = g^{\mu}(Y_k, k) := X_k^{\mu}$. Theorem 2.1 in [Presman & Sonin, 72] now allows for the following: Define the process $(\xi_i)_i$ by virtue of $\xi_1 = 1$, $\xi_k := \min\{n > \xi_{k-1} | g^{\mu}(z_n) > 0\}$, i.e. *$\xi_i$ is the arrival time of the i^{th} candidate.* We then set $Z_i := z_{\xi_i}$. It is shown that the stopping problems are equivalent, i.e. the maximal expected values from stopping $(z_k)_k$ equals that of $(Z_i)_i$. Hence, we may reduce our analysis to the candidate process Z_i. Since $Z_i = z_{\xi_i} = (1, \xi_i)$, we identify $Z_i \equiv \xi_i$ and $g^{\mu}(\xi_i) \equiv g^{\mu}(1, \xi_i)$. Hence, we have:

Remark 2.4.12. *The Best-Choice problem under risk is reduced to optimally stop the candidate process $(\xi_i)_i$ with corresponding payoff function $g^{\mu}(\xi_i) = X_{\xi_i}^{\mu} = \frac{\xi_i}{\mu(N\geq\xi_i)} \sum_{s=\xi_i}^{\infty} \frac{\mu(s)}{s}$.*

In order to solve the problem, we need to characterize the distribution $\mathbb{P}^{\mu}$ of $(\xi_i)_i$, which then of course also yields the distribution of $X_{\xi_i}^{\mu} = g^{\mu}(\xi_i)$.

Given μ, this is entirely achieved by the initial distribution

$$\mathbb{P}^{\mu} \circ \xi_1^{-1} = \mathbb{I}_{\{\xi_1 = 1\}},$$

as the first applicant is obviously a candidate, and the homogenous probability kernel (cf. Section 2.3.2)

$$\begin{aligned} p^{\mu}_{i-1}(l,k) &:= p^{\mu}(l,k) := \mathbb{P}^{\mu}(\xi_i = k | \xi_{i-1} = l) \\ &= \begin{cases} \frac{l\mu(N \geq k)}{k(k-1)\mu(N \geq l)}, & l < k < \infty, \\ 0 & l \geq k, \end{cases} \qquad \forall i \geq 2 \qquad (2.3) \\ p^{\mu}(\infty,\infty) &= 1, \\ p^{\mu}(l,\infty) &= \mathbb{P}^{\mu}(\bar{Y}_l = 1 | Y_l = 1) = \frac{l}{\mu(N \geq l)} \sum_{s=l}^{\infty} \frac{\mu(s)}{s}. \end{aligned}$$

$p^{\mu}(l,k)$ is the probability that the k^{th} applicant is a candidate given the foregoing candidate is applicant l. Note that this is the transition kernel of a homogenous Markov chain: intuitively, not the time of appearance of the candidate matters but the time of appearance of the applicant being that candidate.

The Payoff Process in an Ambiguous Setting – A First Approach

Let $\mu \in \tilde{\mathcal{Q}} \subset \mathcal{M}(\mathbb{N})$, the set of priors on $\mathbb{N}$. The aim in this section is to define an appropriate payoff process as well as the set $\mathcal{Q}$ of priors on that process corresponding to the set $\tilde{\mathcal{Q}}$ of priors on applicants.

Assumption 2.4.13. *Let $\tilde{\mathcal{Q}}$ be closed and convex. If $\mu_1, \mu_2 \in \tilde{\mathcal{Q}}$, then* $\sup\{n | n \in supp(\mu_1)\} = \sup\{n | n \in supp(\mu_2)\}$.

The last assumption ensures the corresponding set $\mathcal{Q}$ of distributions $\mathbb{P}^{\mu}$ of the candidate process $(\xi_i)_i$ via equation (2.3) being equivalent as imposed in [Riedel, 09].[13] Observing applicant $\xi_i = k$, the ambiguity averse agent

[13] This immediately follows from equation (2.3): if $\mu_1(N \geq k) = 0$ for some k, then the same has to hold for μ_2, otherwise, the candidate process corresponding to μ_2 puts positive probability on events that are null sets under the process corresponding to μ_1.

evaluates her minimax expected value from choosing her as

$$\begin{aligned} X_k^{\tilde{\mathcal{Q}}} &:= \min_{\mu\in\tilde{\mathcal{Q}}} \mathbb{E}^\mu[\mathbb{I}_{\bar{Y}_k=1}|\mathcal{F}_k] \\ &= \begin{cases} \min_{\mu\in\tilde{\mathcal{Q}}} \left\{\frac{k}{\mu(N\geq k)} \sum_{s=k}^\infty \frac{\mu(s)}{s}\right\} & \text{if } Y_k = 1, \\ 0 & \text{else.} \end{cases} \end{aligned}$$

This payoff is an immediate consequence of [Gilboa & Schmeidler, 89] in a static set up. By definition, $X_k^{\tilde{\mathcal{Q}}} \in \mathcal{F}_k$ and hence an admissible payoff process.

We have seen, that every $\mu \in \tilde{\mathcal{Q}} \subset \mathcal{M}(\mathbb{N})$ corresponds to a distribution $\mathbb{P}^\mu$ of the candidate process $(\xi_i)_i$ by virtue of equation (2.3), but with payoff function

$$g^{\tilde{\mathcal{Q}}}(\xi_i) := X_{\xi_i}^{\tilde{\mathcal{Q}}} = \min_{\mu\in\tilde{\mathcal{Q}}} \left\{ \frac{\xi_i}{\mu(N \geq \xi_i)} \sum_{s=\xi_i}^\infty \frac{\mu(s)}{s} \right\}. \tag{2.4}$$

Hence, $\tilde{\mathcal{Q}}$ corresponds to some set

$$\mathcal{Q} := \{\mathbb{P}^\mu := \mathbb{I}_{\{\xi_i=1\}} \otimes (p^\mu)^{\mathbb{N}} \quad | \quad \mu \in \tilde{\mathcal{Q}}\} \tag{2.5}$$

of priors $\mathbb{P}^\mu$ of $(\xi_i)_i$, where p^μ is defined in equation (2.3). Note, that μ is fixed in $\mathbb{P}^\mu$, i.e. does not switch to another prior on the number of applicants in course of time; this eventually will cause the time-consistency issues.

Remark 2.4.14 (Model I)**.** *Given $\tilde{\mathcal{Q}}$, we may now solve the optimal stopping problem of the candidate process $(\xi_i)_i$ with payoff $g^{\tilde{\mathcal{Q}}}$ as in equation (2.4) for an ambiguity averse agent facing $\mathcal{Q}$ from equation (2.5). In other words, we have the optimal stopping problem of the model $(\Omega, \mathcal{F}, \mathbb{P}_0, (\mathcal{F}_{\xi_i})_i, (X_{\xi_i}^{\tilde{\mathcal{Q}}})_i, \mathcal{Q})$ as in [Riedel, 09].*

Remark 2.4.15. *This model is an eligible generalization of the Best-Choice problem under risk, as it holds for any stopping time τ*

$$\inf_{\mathbb{P}^\mu\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu}\left[\min_{\mu\in\tilde{\mathcal{Q}}}\left\{\frac{\xi_\tau}{\mu(N\geq\xi_\tau)}\sum_{s=\xi_\tau}^\infty \frac{\mu(s)}{s}\right\}\right] = \inf_{\mathbb{P}^\mu\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu}\left[\mathbb{I}_{\{Y_j>1 \quad \forall j>\xi_\tau\}}\right].$$

This fact immediately follows from construction or, explicitly, from Lemma 1 in [Chudjakow & Riedel, 09].

When choosing a stopping time τ, we may calculate the (minimax) expected reward $\inf_{\mathbb{P}^\mu \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu}[X^{\tilde{\mathcal{Q}}}_{\xi_\tau}]$ and the agent's problem is

$$\sup_\tau \inf_{\mathbb{P}^\mu \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu}\left[X_{\xi_\tau}\right] = \sup_\tau \inf_{\mathbb{P}^\mu \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu}\left[\min_{\mu \in \tilde{\mathcal{Q}}} \left\{ \frac{\xi_\tau}{\mu(N \geq \xi_\tau)} \sum_{s=\xi_\tau}^{\infty} \frac{\mu(s)}{s} \right\}\right].$$

More formally, the (multiple prior) value $(V^{\tilde{\mathcal{Q}}}_{\xi_i})_i$ of the candidate process at candidate i is

$$V^{\tilde{\mathcal{Q}}}_{\xi_i} := \operatorname{ess\,sup}_{\tau \geq i} \operatorname{ess\,inf}_{\mathbb{P} \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}}[\underbrace{g^{\tilde{\mathcal{Q}}}(\xi_\tau)}_{X^{\tilde{\mathcal{Q}}}_{\xi_\tau}} | \mathcal{F}_{\xi_i}].$$

Again, we slightly misuse notation: We are now faced with optimally stopping the payoff process $(\bar{X}_i^{\tilde{\mathcal{Q}}})_i := (X^{\tilde{\mathcal{Q}}}_{\xi_i})_i$ adapted to the filtration $(\bar{\mathcal{F}}_i)_i := (\mathcal{F}_{\xi_i})_i$. To be entirely in line with the notation from the general theory, the value is actually given by $V^{\tilde{\mathcal{Q}}}_i = \operatorname{ess\,sup}_{\tau \geq i} \operatorname{ess\,inf}_{\mathbb{P} \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}}[\bar{X}^{\tilde{\mathcal{Q}}}_\tau | \bar{\mathcal{F}}_i]$. However, we consider the notation in terms of $(V^{\tilde{\mathcal{Q}}}_{\xi_i})_i$ more handy in our model. It furthermore makes explicit the relation of the reduced problem to the "simple" Best-Choice problem as we see that the value process of the candidate process is just a sub-process of the value process of the applicant process; the same holds true for the filtration. In this setup, a stopping time τ does not mean to stop at applicant τ but at candidate τ, i.e. at applicant ξ_τ. Note, in case of a unique μ, the above expression reduces to value function in [Presman & Sonin, 72].

Problems and their Removal

Before we go on, we have to answer two questions: Are all properties satisfied in order to apply the theory in [Riedel, 09]? Does the set-up make sense from an economic point of view?

Answering the first question is equivalent to posing the question whether we can identify properties of $\tilde{\mathcal{Q}}$ in order for $\mathcal{Q}$ to be time-consistent. As we will see in Proposition 2.4.16, constructing $\mathcal{Q}$ as above, we cannot obtain

$\mathcal{Q}$ to be time-consistent. Of course, from a mathematical point of view, we could introduce something like a *time-consistent hull* of $\mathcal{Q}$:

$$\text{TC}(\mathcal{Q}) := \left\{ \mathbb{I}_{\{\xi_i=1\}} \otimes_{i=1}^{\infty} p_i^{\mu_i} \middle| p_i^{\mu_i} \text{ as in equation (2.3) for some } \mu_i \in \tilde{\mathcal{Q}} \right\}.$$

This approach has two major disadvantages: First, it only allows for simply pasting kernels from distributions in $\mathcal{Q}$. However, we have to change the internal structure of kernels for a meaningful formalization of the notion of time-consistency in this context, since kernels do not just incorporate a marginal distribution of μ at the respective candidate but a probability induced by that μ of all future applicants, in particular of applicants beyond the time of pasting. More formally, if we paste at candidate t, then, being at candidate $k < t$, the kernel used at k incorporates the respective measure μ_1 also for times beyond t via the term $\mu_1(N \geq k)$, where μ_2 is the generating measure. Secondly, a pasted distribution in $TC(\mathcal{Q})$ does not correspond to a distribution in $\tilde{\mathcal{Q}}$ in general, i.e. there are distributions in $TC(\mathcal{Q})$ that cannot be induced by a single distribution in $\tilde{\mathcal{Q}}$. In particular, we might achieve a worst case distribution that is not induced by a prior in $\tilde{\mathcal{Q}}$.

Having obtained the set of priors $\mathcal{Q}$ on the candidate process $(\xi_i)_i$ from the set of priors $\tilde{\mathcal{Q}}$ on the number of applicants by virtue of equation (2.5), recall that time-consistency in terms of Definition 2.4.3 assumes $\mathcal{Q}$ to be closed under pasting.

Proposition 2.4.16. *If $\mu_1 \neq \mu_2 \in \tilde{\mathcal{Q}}$ with corresponding priors $\mathbb{P}_1 \neq \mathbb{P}_2 \in \mathcal{Q}$, $1 \leq t \leq \max\{n | n \in supp(\mu_i)\}$, and we define $\mathbb{P}_3$ by virtue of $\mathbb{P}_3 := \mathbb{I}_{\xi_1=1} \otimes p_1 \otimes \ldots \otimes p_1 \otimes p_2 \otimes \ldots$, where p_i are the respective kernels, i.e. $\mathbb{P}_3$ is obtained by pasting kernels at candidate t. Then, there does not exist any $\mu \in \mathcal{M}(\mathbb{N})$ generating $\mathbb{P}_3$ via equation (2.3).* In particular, $\mathcal{Q}$ generated from $\tilde{\mathcal{Q}}$ by virtue of equation (2.5) cannot be time-consistent as $\mathbb{P}_3 \notin \mathcal{Q}$.

Proof. Assume, there exists $\tilde{\mu} \in \tilde{\mathcal{Q}}$ s.t. $p^{\tilde{\mu}}(l,k)$ is generated from $\tilde{\mu}$ as in equation (2.3), i.e.

$$p_{i-1}^{\tilde{\mu}}(l,k) = \frac{l}{k(k-1)} \frac{\tilde{\mu}(N \geq k)}{\tilde{\mu}(N \geq l)} \quad \forall l < k < \infty, \quad i \geq 2$$

and

$$\begin{aligned} p^{\tilde{\mu}}_{i-1}(l,k) &= \frac{l}{k(k-1)} \frac{\mu_1(N \geq k)}{\mu_1(N \geq l)} \quad \forall l < k < \infty, \quad t > i \geq 2, \\ p^{\tilde{\mu}}_{i-1}(l,k) &= \frac{l}{k(k-1)} \frac{\mu_2(N \geq k)}{\mu_2(N \geq l)} \quad \forall l < k < \infty, \quad i \geq t. \end{aligned}$$

Set $l = 1$ and obtain

$$\frac{1}{k(k-1)} \mu_1(N \geq k) = p^{\tilde{\mu}}_1(1,k) = \frac{1}{k(k-1)} \tilde{\mu}(N \geq k) \quad k > 1,$$

implying $\tilde{\mu} = \mu_1$. Likewise, we see $\tilde{\mu} = \mu_2$, Contradicting $\mu_1 \neq \mu_2$. In particular we would have $\mathbb{P}_1 = \mathbb{P}_2 = \mathbb{P}_3$. □

In order to generate a time-consistent model, we now consider the following definition motivated by the proof of Proposition 2.4.16:

Definition 2.4.17. *For $\mu_1, \mu_2 \in \mathcal{Q}$, $t \in \mathbb{N}$, let $\tilde{\mu} \in \mathcal{Q}$ defined as*

$$\tilde{\mu}(i) := \frac{1}{\mu_1(N < t) + \mu_2(N \geq t)} \begin{cases} \mu_1(i) & \text{if } i < t, \\ \mu_2(i) & \text{if } i \geq t. \end{cases}$$

The corresponding kernel is given by

$$\begin{aligned} p^{\tilde{\mu}}(l,k) &= \frac{l}{k(k-1)} \frac{\tilde{\mu}(N \geq k)}{\tilde{\mu}(N \geq l)} \\ &= \frac{l}{k(k-1)} \begin{cases} \frac{\mu_2(N \geq k)}{\mu_2(N \geq l)} & k > l \geq t, \\ \frac{\mu_2(N \geq k)}{\mu_1(t > N \geq l) + \mu_2(N \geq t)} & k \geq t > l, \\ \frac{\mu_1(t > N \geq k) + \mu_2(N \geq t)}{\mu_1(t > N \geq l) + \mu_2(N \geq t)} & t > k > l, \\ 0 & \text{else.} \end{cases} \end{aligned}$$

However, having a look at this kernel, we immediately observe the following problem: Given a stopping time τ, i.e. stop at candidate τ, i.e. stop at applicant $\xi_\tau = t$. Consider the case $\xi_\tau = t > l = \xi_{i-1}$, in particular $\tau > i - 1$. Hence, we have that $\mathbb{P}(\xi_i = k | \xi_{i-1} = l)$ is not $\mathcal{F}_{\xi_{i-1}}$-measurable but $\mathcal{F}_{\xi_\tau}$-measurable by the above formula, contradicting the general properties that conditional probabilities have to satisfy. *Hence, this is not an admissible density process since we need future information.*

In the next section, we come up with an appropriate notion taking care of time-consistency as well as measurability problems.

Ambiguity in terms of Assessments

As seen, a straightforward generalization of [Presman & Sonin, 72]'s model leads to time-*in*consistency, non-measurability as well as to properties that are hard to justify in economic terms. We now tackle these issues. The problem of time-consistency arises because $\tilde{\mathcal{Q}}$ does not incorporate any notion of time, whereas $\mathcal{Q}$ does. Hence, we consider the following definition:

Definition 2.4.18. *A sequence* $\mu := (\mu_1, \mu_2, \ldots) \in \mathcal{M}(\mathbb{N})^{\mathbb{N}}$ *is called an* assessment*.*

Notation 2.4.19. *In order to keep notation simple, we stick to our old notation though the content has changed: Now,* μ *denotes an assessment and not an element in* $\mathcal{M}(\mathbb{N})$*, whereas* μ_i *is the generic notation for these distributions.* $\tilde{\mathcal{Q}}$ *now denotes a set of assessments, not of simple distributions any longer and* Q *defined below is the set of priors corresponding to the set* $\tilde{\mathcal{Q}}$ *of assessments.*

Intuitively, given an assessment $(\mu_i)_i$, μ_k denotes the distribution on the number of applicants, the agent thinks to be correct upon observing applicant k. We do not assume $\mu_k(i) = 0$ for $i < k$; in particular, μ_k is in general not the distribution conditional on $N \geq k$. Recall that the aim is to find a time-consistent set $\mathcal{Q}$ of distributions of $(\xi_i)_i$. Hence, let us now assume that the agent has a set $\tilde{\mathcal{Q}}$ *consisting of assessments.* Assume that an assessment induces a distribution of $(\xi_i)_i$ via the kernels

$$\mathbb{P}^*(\xi_i = k|\xi_{i-1} = l) := \begin{cases} \frac{l}{k(k-1)} \frac{\mu_i(k)+\mu_{i+1}(k+1)+\ldots}{\mu_{i-1}(l)+\mu_i(l+1)+\ldots} & \text{if } k > l, \\ 0 & \text{else,} \end{cases}$$

or alternatively

$$\mathbb{P}^{**}(\xi_i = k|\xi_{i-1} = l) := \begin{cases} \frac{l}{k(k-1)} \frac{\mu_k(k)+\mu_{k+1}(k+1)+\ldots}{\mu_l(l)+\mu_{l+1}(l+1)+\ldots} & \text{if } k > l, \\ 0 & \text{else.} \end{cases}$$

Note, that the first kernel does not only depend on k and l but also on i. Now, pasted kernels correspond to some assessment and in order to achieve time-consistency, this is assumed to be in $\tilde{\mathcal{Q}}$. Note, that the second kernel is the one induced by $\tilde{\mu} \in \mathcal{M}(\mathbb{N})$ defined by $\tilde{\mu}(i) := \frac{1}{\sum_{j\geq 1}\mu_j(j)} \sum_{j\geq 1} \mathbb{I}_{j=i}\mu_i(i)$, $\mu_i(k) := \mu_i(N = k)$. However, having a look at this approach we immediately observe two aspects: Pasting assessments, we may easily run into the same measurability problems as before. Furthermore, both kernels do not have to be probability kernels. Even more severe, the first approach does not make sense, because we evaluate the probability of a k^{th} applicant existing in terms of the measure at the i^{th} candidate. As for the second alternative, does it really make sense to evaluate the probability that $N = j$ by $\mu_j(j)$ for $j \geq l$ being at applicant l, where we have assesment μ_l? We don't think so and hence, we define the transistion probability in another way:

Definition 2.4.20. *Given assessment $\mu := (\mu_i)_i$, define the kernel*

$$\mathbb{P}^{\mu}(\xi_i = k|\xi_{i-1} = l) = p^{\mu}(l,k) := \begin{cases} \frac{l}{k(k-1)} \frac{\mu_l(N\geq k)}{\mu_l(N\geq l)} & \text{if } l < k < \infty, \\ 0 & l \geq k, \end{cases} \tag{2.6}$$

$$p^{\mu}(\infty,\infty) := 1,$$

$$p^{\mu}(l,\infty) := \frac{l}{\mu_l(N \geq l)} \sum_{s=l}^{\infty} \frac{\mu_l(s)}{s}.$$

Note that $\mathbb{P}^{\mu}(\xi_i = k|\xi_{i-1} = l) \in \mathcal{F}_{\xi_{i-1}}$.

Assumption 2.4.21. *Given a set $\tilde{\mathcal{Q}}$ of assessments, set $\tilde{\mathcal{Q}}_k := \{\mu_k|(\mu_i)_i \in \tilde{\mathcal{Q}}\}$. For every k, let $\tilde{\mathcal{Q}}_k$ be convex and closed. Moreover, if $\mu^1, \mu^2 \in \tilde{\mathcal{Q}}$, then* $\sup\{i|\mu^1_k(i) > 0\} = \sup\{i|\mu^2_k(i) > 0\}\ \forall k$.

Definition 2.4.22. *For $\tilde{\mathcal{Q}}$, the set of assessments, we define the set of priors of $(\xi_i)_i$ as $\mathcal{Q} := \{\mathbb{P}^{\mu} = \mathbb{I}_{\xi_1=1} \otimes (p^{\mu})^{\mathbb{N}}|\mu = (\mu_i)_i \in \tilde{\mathcal{Q}}\}$,[14] where p^{μ} is obtained as in equation (2.6).*

Note that $\tilde{\mathcal{Q}}$ now denotes a set of assessments and not of simple elements in $\mathcal{M}(\mathbb{N})$. $\mathcal{Q}$ still denotes the set of priors on $(\xi_i)_i$ but now induced by multiple assessments. $\tilde{\mathcal{Q}}_k$ contains elements in $\mathcal{M}(\mathbb{N})$, the k-projections of

[14] Convex and compact by the foregoing assumption.

the respective assessments. Again, the latter part of the assumption induces $\mathcal{Q}$ consisting of equivalent distributions, the former allows for the following payoff process: Given a set of assessments $\tilde{\mathcal{Q}}$, upon observing applicant k, we have the (multiple prior) payoff

$$\begin{aligned} X_k^{\tilde{\mathcal{Q}}} &:= \min_{\mu_k \in \tilde{\mathcal{Q}}_k} \mathbb{E}[\mathbb{I}_{\bar{Y}_k=1} | \mathcal{F}_k] \\ &= \begin{cases} \min_{\mu_k \in \tilde{\mathcal{Q}}_k} \left\{ \frac{k}{\mu_k(N \geq k)} \sum_{s=k}^{\infty} \frac{\mu_k(s)}{s} \right\} & \text{if } Y_k = 1, \\ 0 & \text{else.} \end{cases} \end{aligned}$$

By definition, $X_k^{\tilde{\mathcal{Q}}} \in \mathcal{F}_k$ and hence an admissible payoff process. Thus, we consider the candidate process $(\xi_i)_i$ with payoff function

$$g^{\tilde{\mathcal{Q}}}(\xi_i) := X_{\xi_i}^{\tilde{\mathcal{Q}}} = \min_{\mu_{\xi_i} \in \tilde{\mathcal{Q}}_{\xi_i}} \left\{ \frac{\xi_i}{\mu_{\xi_i}(N \geq \xi_i)} \sum_{s=\xi_i}^{\infty} \frac{\mu_{\xi_i}(s)}{s} \right\}. \tag{2.7}$$

Remark 2.4.23 (The Correct Model $(\Omega, \mathcal{F}, \mathbb{P}_0, (\mathcal{F}_{\xi_i})_i, (X_{\xi_i}^{\tilde{\mathcal{Q}}})_i, \mathcal{Q})$)**.** *We are now in the context of [Riedel, 09] and may solve the optimal stopping problem of the candidate process $(\xi_i)_i$ with payoff function $g^{\tilde{\mathcal{Q}}}$ as in equation (2.7), i.e. stopping the payoff process $(X_{\xi_i}^{\tilde{\mathcal{Q}}})_i$, for an ambiguity averse agent facing priors in $\mathcal{Q}$ from Definition 2.4.22 with transition kernel in Definition 2.4.20.*

Proposition 2.4.24. *$\mathcal{Q}$ is time-consistent if and only if $\tilde{\mathcal{Q}}$ satisfies the following property: Given $\mu^1, \mu^2 \in \tilde{\mathcal{Q}}$ and a stopping time τ, then $\mu^3 := (\mu_1^1, \ldots, \mu_{\xi_\tau - 1}^1, \mu_{\xi_\tau}^2, \ldots) \in \tilde{\mathcal{Q}}$.*

Proof. Let $\mathbb{P}^1$ be the distribution corresponding to assessment μ^1 and $\mathbb{P}^2$ to μ^2. Then, we the have as density process of $(\xi_i)_i$ for the respective assessments:

$$\begin{aligned} p_i^j &:= \left.\frac{d\mathbb{P}^j}{d\mathbb{P}_0}\right|_{\mathcal{F}_i} = \frac{d(\mathbb{I}_{\xi_1=1} \otimes (p^j)^{i-1})}{d(\mathbb{I}_{\xi_1=1} \otimes (p^0)^{i-1})}, \\ p_i^j(\xi_1 = l_1, \xi_2 = l_2, \ldots \xi_i = l_i) &= \frac{\mathbb{I}_{\{1\}}(l_1) p^j(l_1, l_2) \ldots p^j(l_{i-1}, l_i)}{\mathbb{I}_{\{1\}}(l_1) p^0(l_1, l_2) \ldots p^0(l_{i-1}, l_i)}, \end{aligned}$$

$\forall l_1 < l_2 < \ldots < l_i$. Now consider a stopping time τ and set

$$r_i := \left.\frac{d\mathbb{R}}{d\mathbb{P}_0}\right|_{\mathcal{F}_i} := \begin{cases} p_i^1 & i \leq \tau, \\ \frac{p_\tau^1 p_i^2}{p_\tau^2} & i > \tau. \end{cases}$$

$$r_i(\xi_1 = l_1, \ldots \xi_i = l_i) = \begin{cases} \frac{\mathbb{I}_{\{1\}}(l_1)p^1(l_1,l_2)\ldots p^1(l_{i-1},l_i)}{\mathbb{I}_{\{1\}}(l_1)p^0(l_1,l_2)\ldots p^0(l_{i-1},l_i)} & i \leq \tau, \\ \frac{\mathbb{I}_{\{1\}}(l_1)p^1(l_1,l_2)\ldots p^1(l_{\tau-1},l_\tau)p^2(l_\tau,l_{\tau+1})\ldots p^2(l_{i-1},l_i)}{\mathbb{I}_{\{1\}}(l_1)p^0(l_1,l_2)\ldots p^0(l_{i-1},l_i)} & i > \tau, \end{cases}$$

$\forall l_1 < l_2 < \ldots < l_i$. We immediately see that $\mathbb{R}$ is induced by any assessment of the form $\mu^3 := (\mu_1^1, \ldots, \mu_{l_{\tau-1}}^1, \mu_{l_{\tau-1}}^{a_1}, \ldots \mu_{l_{\tau-1}}^{a_{l_\tau - l_{\tau-1}-1}}, \mu_{l_\tau}^2, \ldots)$ with $a_j \in \{1, 2\}$, $1 \leq j \leq l_\tau - l_{\tau-1} - 1$. However, since this has to hold for all τ and since all μ are equivalent to μ^0 in the sense that the induced distributions of the candidate process have to be equivalent and hence ξ_i can take all values $k \geq i$ with positive probability, we have that $\mathcal{Q}$ is time-consistent, if and only if

$$\mu^3 := (\mu_1^1, \ldots, \mu_{l_{\tau-1}}^1, \mu_{l_{\tau-1}+1}^1, \ldots, \mu_{l_\tau-1}^1, \mu_{l_\tau}^2, \mu_{l_\tau+1}^2 \ldots) \in \tilde{\mathcal{Q}}$$

for all stopping times τ. □

Example 2.4.25. *$\mathcal{Q}$ is time-consistent if $\tilde{\mathcal{Q}}$ is the independent product of its projections, i.e. $\tilde{\mathcal{Q}} = \tilde{\mathcal{Q}}_1 \otimes \tilde{\mathcal{Q}}_2 \otimes \ldots$*

Remark 2.4.26. *If $\mu_1 = \mu_2 = \ldots$, and $|\tilde{\mathcal{Q}}| = 1$, we are back in the case of [Presman & Sonin, 72]. Our first approach (the time-consistent hull) is achieved by the assumption $\tilde{\mathcal{Q}}_1 = \tilde{\mathcal{Q}}_2 = \ldots$ and the independence assumption.*

To keep the model simple, we pose the following assumption:

Assumption 2.4.27. *We assume that $\tilde{\mathcal{Q}}$ is of the form $\tilde{\mathcal{Q}}_1 = \tilde{\mathcal{Q}}_2 = \ldots$ and that $\tilde{\mathcal{Q}}$ satisfies the assumptions for $\mathcal{Q}$ being time-consistent as given in Proposition 2.4.3,*[15] *i.e. $\tilde{\mathcal{Q}}$ being an independent product of its projections.*

One may object that we might have $\mu_i(k) > 0$ for $k < i$ though it seems counterintuitive given the intuition of an assessment. We might also have assumed $\tilde{\mathcal{Q}}_k$ to only enclose the respective distributions appropriately updated,

[15] Note, that time-consistency is not automatically satisfied in the indistinguishable case: Indeed, set $\mathcal{Q}_i = \{\mu_i^1, \mu_i^2\}$, $\mu_1^j = \mu_2^j := \mu^j$, $(\mu^1, \mu^1), (\mu^2, \mu^2) \in \tilde{\mathcal{Q}}$ but $(\mu^1, \mu^2), (\mu^2, \mu^1) \notin \tilde{\mathcal{Q}}$.

i.e. contingent on observing k applicants. However, this does not change the payoff process $(g^{\tilde{\mathcal{Q}}}(\xi_i))_i$ or the distribution of $(\xi_i)_i$ since these contingencies are "averaged out" in the respective formulae.

2.4.3 Results

Considering our formulation of the ambiguous best choice problem, we are in context of optimally stopping the stochastic process $(\xi_i)_i$ with payoff function $g(\xi_i)$ as in equation (2.7) and transition kernel from Definition 2.4.20, where the agent faces the set of priors $\mathcal{Q}$ as in Definition 2.4.22, induced by a set $\tilde{\mathcal{Q}}$ of assessments satisfying Assumption 2.4.27. The value process $(V^{\tilde{\mathcal{Q}}}_{\xi_i})_i$ at candidate i being applicant ξ_i is given by

$$V^{\tilde{\mathcal{Q}}}_{\xi_i} = \operatorname*{ess\,sup}_{\tau \geq i} \operatorname*{ess\,inf}_{\mathbb{P} \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}}[g^{\tilde{\mathcal{Q}}}(\xi_\tau) | \mathcal{F}_{\xi_i}].$$

From Theorem 2.4.10, we know that the value function equals the multiple prior Snell envelope

$$\begin{aligned}
U^{\tilde{\mathcal{Q}}}_{\xi_i} &= \max\left\{ g^{\tilde{\mathcal{Q}}}(\xi_i); \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}}[U^{\tilde{\mathcal{Q}}}_{\xi_{i+1}} | \mathcal{F}_{\xi_i}] \right\} \\
&= \max\left\{ g^{\tilde{\mathcal{Q}}}(\xi_i); \operatorname*{ess\,inf}_{p(\xi_i,\cdot)} \sum_{s=\xi_i+1}^{\infty} p(\xi_i, s) U^{\tilde{\mathcal{Q}}}_s \right\} \\
&= \max\left\{ g^{\tilde{\mathcal{Q}}}(\xi_i); \operatorname*{ess\,inf}_{\mu_{\xi_i}\in\tilde{\mathcal{Q}}_{\xi_i}} \sum_{s=\xi_i+1}^{\infty} \frac{\xi_i}{s(s-1)} \frac{\mu_{\xi_i}(N \geq s)}{\mu_{\xi_i}(N \geq \xi_i)} U^{\tilde{\mathcal{Q}}}_s \right\} \\
&= \max\left\{ \min_{\mu_{\xi_i}\in\tilde{\mathcal{Q}}_{\xi_i}} \frac{\xi_i}{\mu_{\xi_i}(N \geq \xi_i)} \sum_{s=\xi_i}^{\infty} \frac{\mu_{\xi_i}(s)}{s}; \right. \\
&\qquad \left. \min_{\mu_{\xi_i}\in\tilde{\mathcal{Q}}_{\xi_i}} \sum_{s=\xi_i+1}^{\infty} \frac{\xi_i}{s(s-1)} \frac{\mu_{\xi_i}(N \geq s)}{\mu_{\xi_i}(N \geq \xi_i)} U^{\tilde{\mathcal{Q}}}_s \right\}.
\end{aligned}$$

We now set

$$\tau^* := \min\{i \geq 1 | U^{\tilde{\mathcal{Q}}}_{\xi_i} = g^{\tilde{\mathcal{Q}}}(\xi_i)\}.$$

By Theorem 2.4.10, τ^* is the smallest optimal stopping time. Recall that $\tau(\omega) = m$ means to "stop at candidate m". However, to comply with the classical problem, we want to have a stopping strategy telling us "stop at applicant m, given she is a candidate". Hence, we set

$$\Gamma^{\tilde{\mathcal{Q}}} := \{k | U_k^{\tilde{\mathcal{Q}}} = g^{\tilde{\mathcal{Q}}}(k)\}$$

and we see that

$$\tau^* = \min\{i > 0 | \xi_i \in \Gamma^{\tilde{\mathcal{Q}}}\}.$$

$\Gamma^{\tilde{\mathcal{Q}}}$ *is the set of all arrival times of applicants (not of candidates) that are optimally chosen if being a candidate.* τ^* is the first candidate in $\Gamma^{\tilde{\mathcal{Q}}}$.

Remark 2.4.28. *The specific structure of* $\Gamma^{\tilde{\mathcal{Q}}}$ *is the solution to our problem and the multiple prior Snell envelope entirely characterizes* $\Gamma^{\tilde{\mathcal{Q}}}$*.*

Remark 2.4.29 (Instantaneous Worst-Case Assessment)**.** *Having a look at the left hand side of the Snell envelope,* $g^{\tilde{\mathcal{Q}}}(k)$*, at applicant* k *being a candidate, the agent has to calculate the* instantaneous payoff *from stopping by minimizing expected payoff with respect to the set of* k*-projections of assessments. As for a given distribution of the number of applicants, the instantaneous payoff is just the probability that no further candidate will follow, the instantaneous payoff in the multiple priors set-up is just given by the minimum of this probability with respect to all possible distributions of numbers of applicants.*

Recall that we assume all orderings of agents being equally likely. Hence, the probability of the current candidate being the last is minimal for the distribution that puts weight on large numbers of applicants. *Hence, for every applicant* k*, there corresponds a distribution* $\tilde{\mu}_k \in \tilde{\mathcal{Q}}_k$ *such that*

$$\min_{\mu_k \in \tilde{\mathcal{Q}}_k} \left\{ \frac{k}{\mu_k(N \geq k)} \sum_{s=k}^{\infty} \frac{\mu_k(s)}{s} \right\} = \frac{k}{\tilde{\mu}_k(N \geq k)} \sum_{s=k}^{\infty} \frac{\tilde{\mu}_k(s)}{s}.$$

Hence we may define the instantaneous worst-case assessment $\tilde{\mu}$ *by virtue of components* $(\tilde{\mu}_i)_i$ *minimizing the instantaneous payoff at applicant* i*. From the structure of the minimization problem and assuming* $\tilde{\mathcal{Q}}_k = \tilde{\mathcal{Q}}_m$*, we see* *that* $\tilde{\mu}_k = \tilde{\mu}_m$ *for every* k, m, *i.e. the instantaneously minimizing assessment is constant. It is immediate, that there is no problem in calculating the instantaneous worst case assessment* $(\tilde{\mu}_k)_k$ *for the instantaneous payoff* $(g^{\tilde{\mathcal{Q}}}(k))_k$ *in advance as this is, irrespective of whatever might happen, the distribution that puts on average most weight on higher values.*

Notation 2.4.30. *By the foregoing remark, we may hence write* $g^{\tilde{\mathcal{Q}}}(k) = g^{\tilde{\mu}_k}(k)$, *where* $\tilde{\mu} = (\tilde{\mu}_i)_i$ *denotes the instantaneous worst-case assessment.*

A first – unfruitful – Approach to a Solution

The first idea to the solution of the problem is to use the minimax theorem in order to interchange the infimum and the supremum in the problem's value function. This would allow for solving the inner maximization problem as in [Presman & Sonin, 72] for every assessment under consideration and then obtain the worst case assessment in terms of that solution with minimal payoff to the agent. However, as the instantaneous payoff also depends on the distribution, we will show that this approach is not eligible for the ambiguous Best-Choice problem. Formally, we have

$$\begin{aligned}
V_{\xi_i}^{\tilde{\mathcal{Q}}} &= \operatorname*{ess\,sup}_{\tau \geq i} \operatorname*{ess\,inf}_{\mathbb{P}^\mu \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu}[g^{\tilde{\mathcal{Q}}}(\xi_\tau) | \mathcal{F}_{\xi_i}] \\
&= \operatorname*{ess\,sup}_{\tau \geq i} \operatorname*{ess\,inf}_{\mu \in \tilde{\mathcal{Q}}} \mathbb{E}^{\mathbb{P}^\mu}[g^{\tilde{\mathcal{Q}}}(\xi_\tau) | \mathcal{F}_{\xi_i}] \\
&= \operatorname*{ess\,sup}_{\tau \geq i} \operatorname*{ess\,inf}_{\mu \in \tilde{\mathcal{Q}}} \mathbb{E}^{\mathbb{P}^\mu} \left[\min_{\mu_{\xi_\tau} \in \tilde{\mathcal{Q}}_{\xi_\tau}} \frac{\xi_\tau}{\mu_{\xi_\tau}(N \geq \xi_\tau)} \sum_{s=\xi_\tau}^{\infty} \frac{\mu_{\xi_\tau}(s)}{s} \middle| \mathcal{F}_{\xi_i} \right] \\
&\overset{\text{in general}}{\neq} \operatorname*{ess\,sup}_{\tau \geq i} \operatorname*{ess\,inf}_{\mu \in \tilde{\mathcal{Q}}} \mathbb{E}^{\mathbb{P}^\mu} \left[\frac{\xi_\tau}{\mu_{\xi_\tau}(N \geq \xi_\tau)} \sum_{s=\xi_\tau}^{\infty} \frac{\mu_{\xi_\tau}(s)}{s} \middle| \mathcal{F}_{\xi_i} \right] \\
&\overset{\text{MiniMax}}{=} \operatorname*{ess\,inf}_{\mu \in \tilde{\mathcal{Q}}} \operatorname*{ess\,sup}_{\tau \geq i} \mathbb{E}^{\mathbb{P}^\mu} \left[\frac{\xi_\tau}{\mu_{\xi_\tau}(N \geq \xi_\tau)} \sum_{s=\xi_\tau}^{\infty} \frac{\mu_{\xi_\tau}(s)}{s} \middle| \mathcal{F}_{\xi_i} \right],
\end{aligned}$$

The second to last inequality destroys our simple approach to the probelm: We cannot just reduce the ambiguous problem to the risky one, i.e. solve the inner problem in the last line as in [Presman & Sonin, 72] for every assessment on its own and then apply the worst of these risky solutions to solve the ambiguous one. We will make this more concrete in Remark 2.4.32 below as the upper inequality shows the multiple prior Snell envelope not to be the lower envelope of the individual risky Best-Choice problems' Snell envelopes. We however argue that this does not contradict [Riedel, 09] as we consider a family of induced risky problem, the Snell envelopes of which are enveloped from below by the multiple prior Snell envelope.

Remark 2.4.31 (On the Schizophrenia of Agents)**.** *It is not just formally obvious that the line of equations does not hold in general but also intuitively. Before we apply the minimax theorem in the above line of equations, we combine the minimal instantaneous distribution with the worst case dynamic distribution. However, we have to distinguish these terms: the instantaneous worst case distribution is just the minimizer in the instantaneous payoff* $g(k)$ *at applicant* k *being a candidate. Of course, due to homogeneity, we can at time zero calculate the assessment minimizing the instantaneous payoff, i.e.* $(\tilde{\mu}_k)_k$ *s.t.* $\tilde{\mu}_k \in \arg\min g^{\tilde{Q}}(k)$ *for every* k*. We call* $(\tilde{\mu}_k)$ *the instantaneous worst case assessment as each component gives the worst case distribution for the instantaneous payoff at the respective candidate. On the other hand, we calculate the worst case distribution of the candidate process. This is given by some worst case assessment* $(\bar{\mu}_k)_k$ *that induces the worst case distribution for the payoff in terms of the kernels in equation (2.3).*

Our approach above would now imply these worst case assessments (the instantaneous and the dynamic one) to coincide. *However, this is not true as might immediately be seen in case of prior assessments consisting of distributions that induce the best choice problem to still be monotone, as e.g. families of uniform distributions. Observing applicant* k *being a candidate, the instantaneous worst case distribution* $\tilde{\mu}_k$ *on the number of applicants would prefer high values as this would minimize the probability of the respective candidate chosen being the best and hence minimize the instantaneous* *payoff.*[16] *However, in the monotonic case, as the value function, and hence the right hand side of the Snell envelope is increasing, the worst case assessment puts most weight on lower values of the candidate process. Hence, in general* $(\tilde{\mu}_k)_k \neq (\bar{\mu}_k)_k$*.*

This behavior seems quite schizophrenic on first sight: At applicant k being a candidate, the agent beliefs that nature will choose a different distributions contingent on her decision to stop or not. We, however, do not consider this observation as unintuitive: Having decided on stopping or going further, the agent's view of what might happen in worst case changes drastically.

Solution to the Ambiguous Problem

We have seen the multiple prior Snell envelope to be given by

$$\begin{aligned} U_{\xi_i}^{\tilde{\mathcal{Q}}} &= \max\left\{ \min_{\mu_{\xi_i}\in\tilde{\mathcal{Q}}_{\xi_i}} \left(\frac{\xi_i}{\mu_{\xi_i}(N \geq \xi_i)} \sum_{s=\xi_i}^{\infty} \frac{\mu_{\xi_i}(s)}{s} \right); \right. \\ & \qquad \left. \min_{\mu_{\xi_i}\in\tilde{\mathcal{Q}}_{\xi_i}} \left(\sum_{s=\xi_i+1}^{\infty} \frac{\xi_i}{s(s-1)} \frac{\mu_{\xi_i}(N \geq s)}{\mu_{\xi_i}(N \geq \xi_i)} U_s^{\tilde{\mathcal{Q}}} \right) \right\} \end{aligned}$$

and the optimal stopping time as $\tau^* := \min\{i|\xi_i \in \Gamma^{\tilde{\mathcal{Q}}}\}$, where the stopping set $\Gamma^{\tilde{\mathcal{Q}}} := \{k|g^{\tilde{\mathcal{Q}}}(k) = U_k^{\tilde{\mathcal{Q}}}\}$. Hence, we can write

$$\begin{aligned} \tau^* &= \min\{i|g^{\tilde{\mathcal{Q}}}(\xi_i) = U_{\xi_i}^{\tilde{\mathcal{Q}}}\} \\ &= \min\left\{ i \,\middle|\, g^{\tilde{\mathcal{Q}}}(\xi_i) \geq \min_{\mu\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu}\left[U_{\xi_{i+1}}^{\tilde{\mathcal{Q}}} \,\middle|\, \mathcal{F}_{\xi_i} \right] \right\} \\ &= \min\left\{ i \,\middle|\, g^{\tilde{\mathcal{Q}}}(\xi_i) \geq \min_{\mu_{\xi_i}\in\tilde{\mathcal{Q}}_{\xi_i}} \left(\sum_{s=\xi_i+1}^{\infty} \frac{\xi_i}{s(s-1)} \frac{\mu_{\xi_i}(N \geq s)}{\mu_{\xi_i}(N \geq \xi_i)} U_s^{\tilde{\mathcal{Q}}} \right) \right\} \end{aligned}$$

Remark 2.4.32 (Major Problem for the solution)**.** *In [Riedel, 09]'s theory of optimal stopping under ambiguity, instantaneous payoff did not depend on*

[16] Of course, this monotonic behavior of instantaneous payoff g always holds and, hence, the instantaneous worst case assessment $(\tilde{\mu}_k)_k$ is calculated in any case. Of course, appropriate assumptions have to be required as, otherwise, nature would choose a distribution favoring infinitely many applicants and hence set the payoff to zero.

priors. Hence, in that case the multiple prior Snell envelope is the lower envelope of the individual Snell envelopes with respect to the single priors. In the Snell envelope of the risky Best-Choice problem, however, the instantaneous payoff g, the left hand side of the Snell envelope, also depends on the distribution. Hence, the multiple prior Snell envelope of the ambiguous Best-Choice problem is not the lower envelope of the individual Snell envelopes of the respective risky Best-Choice problems.

However the way we solve this apparent contradiction to [Riedel, 09] is by artificially introducing an induced risky problem *by virtue of the instantaneous payoff g already as the minimal instantaneous payoff with respect to priors. In that respect,* the multiple prior Snell envelope of the ambiguous Best-Choice Problem is the lower envelope of the the Snell envelopes of the optimal stopping problems with artificial payoff g given the respective priors. *But it is important to keep in mind that these optimal stopping problems are not the risky Best-Choice problems as the payoff of the induced risky problems is given by the minimized expectation and hence in general not equal to the payoff of the risky Best-Choice problems.*

More formally *it holds:*

$$
\begin{aligned}
& \mathit{ess}\sup_{\tau\geq i} \mathit{ess}\inf_{\mu\in\tilde{\mathcal{Q}}} \mathbb{E}^{\mathbb{P}^\mu}\left[\min_{\mu_{\xi_\tau}\in\tilde{\mathcal{Q}}_{\xi_\tau}} \frac{\xi_\tau}{\mu_{\xi_\tau}(N\geq\xi_\tau)}\sum_{s=\xi_\tau}^{\infty}\frac{\mu_{\xi_\tau}(s)}{s}\middle|\mathcal{F}_{\xi_i}\right] \\
= \quad & \mathit{ess}\inf_{\mu\in\tilde{\mathcal{Q}}} \mathit{ess}\sup_{\tau\geq i} \mathbb{E}^{\mathbb{P}^\mu}\left[\min_{\mu_{\xi_\tau}\in\tilde{\mathcal{Q}}_{\xi_\tau}} \frac{\xi_\tau}{\mu_{\xi_\tau}(N\geq\xi_\tau)}\sum_{s=\xi_\tau}^{\infty}\frac{\mu_{\xi_\tau}(s)}{s}\middle|\mathcal{F}_{\xi_i}\right] \\
\overset{\text{in general}}{\neq} \quad & \mathit{ess}\inf_{\mu\in\tilde{\mathcal{Q}}} \mathit{ess}\sup_{\tau\geq i} \mathbb{E}^{\mathbb{P}^\mu}\left[\frac{\xi_\tau}{\mu_{\xi_\tau}(N\geq\xi_\tau)}\sum_{s=\xi_\tau}^{\infty}\frac{\mu_{\xi_\tau}(s)}{s}\middle|\mathcal{F}_{\xi_i}\right].
\end{aligned}
$$

The multiple prior Snell envelope of the optimal stopping problem with payoff $g^{\tilde{\mathcal{Q}}}$ *given by* $g^{\tilde{\mathcal{Q}}}(k) := \min_{\mu_k\in\tilde{\mathcal{Q}}_k}\left\{\frac{k}{\mu_k(N\geq k)}\sum_{s=k}^{\infty}\frac{\mu_k(s)}{s}\right\}$ *and set of assessments* $\tilde{\mathcal{Q}}$*, i.e. the ambiguous Best-Choice problem (first line above), is not the lower envelope of the individual Snell envelopes of the risky problems with payoff* g^μ *defined by virtue of* $g^\mu(k) := \frac{k}{\mu_k(N\geq k)}\sum_{s=k}^{\infty}\frac{\mu_k(s)}{s}$ *for assessments* $\mu\in\tilde{\mathcal{Q}}$*, i.e. the risky Best-Choice problems (third line). It is however the lower envelope* *of the individual Snell envelopes of the risky problems with payoff* $g^{\tilde{\mathcal{Q}}}$ *and distributions given by* $\mu \in \tilde{\mathcal{Q}}$ *(second line).*

The Finite Problem

In order to obtain a feeling for solving the problem, we first consider the almost surely finite case. This will already make several aspects explicit.

Assumption 2.4.33. *Given a fixed* $T \in \mathbb{N}$*, we have for all* $\mu \in \tilde{\mathcal{Q}}$ $\operatorname{supp}(\mu_i) \subset [0, T]$ *for all* i*.*

Given this assumption, we have $\max\{\operatorname{supp}(\mu_i)\} = T$ for all i. Recall that $\xi_i = \infty$ if there does not exists an i^{th} candidate and, hence, in particular if there does not exist an i^{th} applicant. Furthermore, $g(\infty) = 0$.

Have in mind that for all i, $\xi_{i+1} > \xi_i$ a.s., in particular $\xi_i \geq i$, and hence the *effective state spaces* of $(\xi_i)_i$ are of the form

$$\begin{aligned}
\xi_1 &= 1 \\
\xi_2 &\in \{2, 3, \ldots, T, \infty\} \quad \mu_2 - a.s. \forall \mu_2 \in \mathcal{Q}_2 \\
&\vdots \\
\xi_{i+1} &\in \{\xi_i + 1, \ldots, T, \infty\} \subset \{i + 1, \ldots, T, \infty\} \quad \mu_{i+1} - a.s. \forall \mu_{i+1} \in \mathcal{Q}_{i+1} \\
&\vdots \\
\xi_T &\in \{T, \infty\} \quad \mu_T - a.s. \forall \mu_T \in \mathcal{Q}_T
\end{aligned}$$

We can now compute:

$$U^{\tilde{\mathcal{Q}}}_{\xi_T} = g^{\tilde{\mathcal{Q}}}(\xi_T) = 1_{\{\xi_T = T\}}$$

and for $U^{\tilde{\mathcal{Q}}}_{\xi_{T-1}}$

$$g^{\tilde{\mathcal{Q}}}(\xi_{T-1}) = \begin{cases} 0 & \text{if } \xi_{T-1} = \infty, \\ 1 & \text{if } \xi_{T-1} = T, \\ \min_{\mu_{T-1}} \left(\frac{T-1}{\mu_{T-1}(N \geq T-1)} \left(\frac{\mu_{T-1}(T-1)}{T-1} + \frac{\mu_{T-1}(T)}{T} \right) \right) & \text{if } \xi_{T-1} = T - 1 \end{cases}$$

and, as one-step ahead conditional expected minimax payoff

$$\begin{aligned}
&\min_{\mu\in\mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu}\left[U^{\tilde{\mathcal{Q}}}_{\xi_T}\middle|\mathcal{F}_{\xi_{T-1}}\right]\\
=&\begin{cases} 0 & \text{if } \xi_{T-1}=\infty,\\ 0 & \text{if } \xi_{T-1}=T,\\ \min_{\mu_{T-1}}\left(p^\mu(T-1,T)U^{\tilde{\mathcal{Q}}}_T + p^\mu(T-1,\infty)U^{\tilde{\mathcal{Q}}}_\infty\right) & \text{if } \xi_{T-1}=T-1. \end{cases}\\
=&\begin{cases} 0 & \text{if } \xi_{T-1}=\infty,\\ 0 & \text{if } \xi_{T-1}=T,\\ \min_{\mu_{T-1}}\left(\frac{1}{T}\frac{\mu_{T-1}(N\geq T)}{\mu_{T-1}(N\geq T-1)}\right) & \text{if } \xi_{T-1}=T-1. \end{cases}
\end{aligned}$$

Hence, we stop at $T-1$ if and only if either

$$\begin{aligned}
\xi_{T-1} &= \infty,\\
\xi_{T-1} &= T,
\end{aligned}$$

or, in case $\xi_{T-1}=T-1$

$$\begin{aligned}
&\min_{\mu_{T-1}}\left(\frac{T-1}{\mu_{T-1}(N\geq T-1)}\left(\frac{\mu_{T-1}(T-1)}{T-1}+\frac{\mu_{T-1}(T)}{T}\right)\right)\\
\geq\;& \min_{\mu_{T-1}}\frac{\mu_{T-1}(N\geq T)}{T\mu_{T-1}(N\geq T-1)},
\end{aligned}$$

where the left hand side of the inequality equals

$$\begin{aligned}
&\min_{\mu_{T-1}}\left(\frac{T-1}{\mu_{T-1}(N\geq T-1)}\left(\frac{\mu_{T-1}(T-1)}{T-1}+\frac{\mu_{T-1}(T)}{T}\right)\right)\\
=\;& \min_{\mu_{T-1}}\left(1-\frac{\mu_{T-1}(T)}{T(\mu_{T-1}(T-1)+\mu_{T-1}(T))}\right)\\
=\;& \min_{\mu_{T-1}}\left(1-\frac{\mu_{T-1}(T)}{T\mu_{T-1}(N\geq T-1)}\right).
\end{aligned}$$

Hence, upon observing $\xi_{T-1}=T-1$, the agent stops the process if and only if

$$\min_{\mu_{T-1}}\left(1-\frac{\mu_{T-1}(T)}{T\mu_{T-1}(N\geq T-1)}\right)\geq \min_{\mu_{T-1}}\frac{\mu_{T-1}(N\geq T)}{T\mu_{T-1}(N\geq T-1)}.$$

Two observations are worthwhile to note: First, $\frac{\mu_{T-1}(T)}{T\mu_{T-1}(N\geq T-1)}$ is the probability that ξ_{T-1} is the second to last candidate, i.e. there will be the best applicant among all to follow at T given the candidate at $T-1$. In this sense *the left hand side is the probability that the candidate at $T-1$ is the best among all, the right hand side the probability that a better applicant is still to follow and hence observed at T.*

In terms of the intuition of coherent risk or multiple prior preferences, the minimization problem that nature has to solve on both sides of the inequality is immediate: on the left hand side, nature has to minimize the probability of the chosen candidate to be the best, on the right hand side she has to minimize the probability that, if candidate $T-1$ is not chosen, there still follows a candidate at T, i.e. nature wants that the candidate at $T-1$ was actually the best and the agent realizes that stopping is too late. Formally:

$$\begin{aligned} \mathbb{P}^{\mu}\left[\bar{Y}_{T-1}=1 \middle| Y_{T-1}=1\right] &= 1-\frac{\mu_{T-1}(T)}{T\mu_{T-1}(N \geq T-1)}, \\ \mathbb{P}^{\mu}\left[Y_{T}=1 \middle| Y_{T-1}=1\right] &= \frac{\mu_{T-1}(T)}{T\mu_{T-1}(N \geq T-1)}, \end{aligned}$$

where $\bar{Y}$ denotes the absolute and Y the relative rank.

Secondly, it is immediate that, observing candidate $T-1$, these minimization problems are conflicting: As set out, on the right hand side, nature minimizes the probability of a better candidate to be chosen, i.e. at $T-1$ to minimize the probability that there is a candidate at T, whereas the left hand side is equivalent to maximize this probability, as then the chosen candidate at $T-1$ is not the best applicant. More formally, the left hand side is equivalent to the problem

$$\max_{\mu_{T-1}} \frac{\mu_{T-1}(N \geq T)}{T\mu_{T-1}(N \geq T-1)}.$$

Hence, at $T-1$, observing $\xi_{T-1}=T-1$, for the immediate payoff function g, we obtain a minimizing assessment $\tilde{\mu}$ s.t. $\tilde{\mu}_{T-1}(T-1)=\mu^{l}(T-1)$ and $\tilde{\mu}_{T-1}(T)=\mu^{u}(T)$, where μ^{u} denotes the assessment putting most weight on T and μ^{l} the one putting least weight. On the other hand, the worst case measure from $T-1$ onwards is abteined by $\bar{\mu}$ s.t. $\bar{\mu}_{T-1}(T-1)=\mu^{u}(T-1)$

and $\bar{\mu}_{T-1}(T) = \mu^l(T)$, i.e. exactly the opposite. Thus, the Snell envelope, upon observing $\xi_{T-1} = T - 1$[17] takes the form

$$U_{T-1}^{\tilde{\mathcal{Q}}} = \left\{ 1 - \frac{\mu_{T-1}^u(T)}{T\mu_{T-1}^u(N \geq T-1)}; \frac{\mu_{T-1}^l(T)}{T\mu_{T-1}^u(N \geq T-1)} \right\}.$$

At this stage, we observe the difference of our Snell envelope in the ambiguous case and the one in [Presman & Sonin, 72] in the risky case: in the risky set up, there is the same distribution on both sides, in our ambiguous approach, there is an instantaneous worst case assessment on the left and a dynamic worst case assessment on the right hand side and those do not coincide. Explicit solutions can now be achieved by going on further with the backward induction principle given explicit characteristics of the set of assessments under consideration. We, however, do not want to achieve this here but have a look in theoretical results on the set Γ of stopping islands.

The General Problem

Again, we note that the problem is entirely solved by characterizing Γ, the stopping set. In general, i.e. when the support is not assumed bounded, the Snell envelope of the problem is given by[18]

$$U_{\xi_i}^{\tilde{\mathcal{Q}}} = \max \left\{ g^{\tilde{\mathcal{Q}}}(\xi_i); \min_{\mu \in \mathcal{Q}} \mathbb{E}^{\mathbb{P}^\mu} \left[U_{\xi_{i+1}}^{\tilde{\mathcal{Q}}} \middle| \mathcal{F}_{\xi_i} \right] \right\},$$

which, for $\xi_i = k$ takes the form

$$U_k^{\tilde{\mathcal{Q}}} = \max \left\{ \min_{\mu_k \in \tilde{\mathcal{Q}}_k} \sum_{s=k}^{\infty} \frac{k}{s} \frac{\mu_k(s)}{\mu_k(N \geq k)}; \min_{\mu_k \in \tilde{\mathcal{Q}}_k} \sum_{s=k+1}^{\infty} \frac{k}{s(s-1)} \frac{\mu_k(N \geq s)}{\mu_k(N \geq k)} U_s^{\tilde{\mathcal{Q}}} \right\}.$$

In [Chudjakow & Riedel, 09], the approach to ambiguity is again leading to monotone problems but does not cover the case of an ambiguous number of

[17]Due to homogeneity of the process, this is the same value for any $\xi_i = T - 1$ as it does not matter if it is the first or $(T-1)^{st}$ candidate at applicant $T - 1$ or whatever in between, i.e. be it ξ_1, ξ_2, ... does not matter.

[18]More precisely, in the infinite case, it is the value function satisfying the Bellman equation.

applicants. As in the risky Best-Choice problem, the main problem here is the lack of monotonicity leading to a multiplicity of stopping islands. Hence, it is not possible in our case to find a worst case distribution for the payoff process in terms of stochastic dominance as it is done in several examples in [Riedel, 09]. We will now emphasis on a theoretical result: The question is, whether we can find conditions to ensure finitely many stopping islands in case of not necessarily bounded support of priors. The following theorem shows that there exists a *final stopping island up to infinity* and, hence, there can only be finitely many stopping islands. It generalizes the main result in [Presman & Sonin, 72] to an ambiguous number of applicants.

We have already introduced the difference between the instantaneous worst case assessment $(\tilde{\mu}_k)_k$ and the dynamic worst case assessment $(\bar{\mu}_k)_k$ in the foregoing paragraph:

$$\tilde{\mu}_k \in arg \underbrace{\min_{\mu_k \in \tilde{\mathcal{Q}}_k} \sum_{s=k}^{\infty} \frac{k}{s} \frac{\mu_k(s)}{\mu_k(N \geq k)}}_{=g^{\tilde{\mathcal{Q}}}(k)}$$

$$\bar{\mu}_k \in arg \min_{\mu_k \in \tilde{\mathcal{Q}}_k} \sum_{s=k+1}^{\infty} p^{\mu_k}(k,s) U^{\tilde{\mathcal{Q}}}(s).$$

Let us know pose two definitions as in [Presman & Sonin, 72]:

$$c_k^{\mu_k} := g^{\tilde{\mathcal{Q}}}(k) - \sum_{s=k+1}^{\infty} p^{\mu_k}(k,s) g^{\tilde{\mathcal{Q}}}(s)$$

and define the operator

$$\mathbb{Q}^{\mu_k} g^{\tilde{\mathcal{Q}}}(k) := \max \left\{ g^{\tilde{\mathcal{Q}}}(k); \sum_{s=k+1}^{\infty} p^{\mu_k}(k,s) g^{\tilde{\mathcal{Q}}}(s) \right\}$$

In the proof of the following main theorem, generalizing Theorem 3.1 in [Presman & Sonin, 72], we inevitably use the fact, that the multiple prior Snell envelope is the lower envelope of the Snell envelopes of the induced risky problems.

Theorem 2.4.34. (a) *If $\Gamma^{\tilde{\mathcal{Q}}}$ consists of finitely many stopping islands, then there exists some k^* such that $c_k^{\bar{\mu}_k} \geq 0$ for all $k \geq k^*$.*

(b) *If there exists $(\mu_k^*)_k$ such that $c_k^{\mu_k^*} \geq 0$ for all $k \geq k^*$, then $\Gamma^{\tilde{\mathcal{Q}}}$ exists of finitely many stopping islands; in particular, $[k^*; \infty] \subset \Gamma^{\tilde{\mathcal{Q}}}$.*

(c) *Given k^* from part (b), if for all $\mu_{k^*-1} \in \tilde{\mathcal{Q}}_{k^*-1}$ it holds $c_{k^*-1}^{\mu_{k^*-1}} < 0$, then $k^* - 1 \notin \Gamma^{\tilde{\mathcal{Q}}}$.*

Proof. ad (a): Note that $\Gamma^{\tilde{\mathcal{Q}}}$ possesses finitely many stopping islands if there exists a "last" stopping island up to infinity. Let $[k^*, \infty] \subset \Gamma^{\tilde{\mathcal{Q}}}$, then for all $k \geq k^*$ we have

$$\begin{aligned} U_k^{\tilde{\mathcal{Q}}} &= g^{\tilde{\mathcal{Q}}}(k) \quad \text{by definition of } \Gamma^{\tilde{\mathcal{Q}}} \\ &\geq \min_{\mu_k} \sum_{s=k+1}^{\infty} p^{\mu_k}(k,s) U_s^{\tilde{\mathcal{Q}}} \quad \text{by definition of } U^{\tilde{\mathcal{Q}}} \\ &= \sum_{s=k+1}^{\infty} p^{\bar{\mu}_k}(k,s) U_s^{\tilde{\mathcal{Q}}} \quad \text{by definition of } \bar{\mu}_k \\ &\geq \sum_{s=k+1}^{\infty} p^{\bar{\mu}_k}(k,s) g^{\tilde{\mathcal{Q}}}(s). \end{aligned}$$

ad (b): Let μ_k^* be such that $c_k^{\mu_k^*} \geq 0$ for all $k \geq k^*$, then

$$g^{\tilde{\mathcal{Q}}}(k) \geq \sum_{s=k+1}^{\infty} p^{\mu_k^*}(k,s) g^{\tilde{\mathcal{Q}}}(s)$$

and hence

$$\mathbb{Q}^{\mu_k^*} g^{\tilde{\mathcal{Q}}}(k) = \max\left\{ g^{\tilde{\mathcal{Q}}}(k); \sum_{s=k+1}^{\infty} p^{\mu_k^*}(k,s) g^{\tilde{\mathcal{Q}}}(s) \right\} = g^{\tilde{\mathcal{Q}}}(k).$$

As $(\xi_i)_i$ is increasing we have that $p^{\mu_k^*}(k,s) = 0$ for all $s \leq k$ and it follows inductively that the payoff process is idempotent with respect to Q, i.e.

$$(\mathbb{Q}^{\mu_k^*})^n g^{\tilde{\mathcal{Q}}}(k) = g^{\tilde{\mathcal{Q}}}(k) \quad \forall n \quad \forall k \geq k^*.$$

Let $U^{\mu_k^*}$ denote the Snell envelope of the induced risky problem under distribution μ_k^* but still with payoff $g^{\tilde{\mathcal{Q}}}$. Then, we know from the general theory of optimal stopping of Markov chains:

$$U_k^{\mu_k^*} = \lim_{n\to\infty} \left(\mathbb{Q}^{\mu_k^*}\right)^n g^{\tilde{\mathcal{Q}}}(k) = g^{\tilde{\mathcal{Q}}}(k) \quad \forall k \geq k^*.$$

As the multiple prior Snell envelope of our ambiguous problem ist the lower envelope of these Snell envelopes, we have

$$g^{\tilde{\mathcal{Q}}}(k) = U_k^{\mu_k^*} \geq \min_{\mu_k \in \tilde{\mathcal{Q}}_k} U_k^{\mu_k} = U_k^{\tilde{\mathcal{Q}}} \quad \forall k \geq k^*.$$

and hence, as by definition of $U^{\tilde{\mathcal{Q}}}$ we have $U_k^{\tilde{\mathcal{Q}}} \geq g^{\tilde{\mathcal{Q}}}(k)$ for all k,

$$U_k^{\tilde{\mathcal{Q}}} = g^{\tilde{\mathcal{Q}}}(k) \quad \forall k \geq k^*.$$

This implies $k \in \Gamma^{\tilde{\mathcal{Q}}}$ for all $k \geq k^*$.

ad (c): If now $c_{k^*-1}^{\mu_{k^*-1}} < 0$ for all $\mu_{k^*-1} \in \tilde{\mathcal{Q}}_{k^*-1}$, then

$$\begin{aligned}
& \mathbb{Q}^{\mu_{k^*-1}} g^{\tilde{\mathcal{Q}}}(k^*-1) > g^{\tilde{\mathcal{Q}}}(k^*-1) \\
\Rightarrow \quad & U_{k^*-1}^{\mu_{k^*-1}} > g^{\tilde{\mathcal{Q}}}(k^*-1) \quad \forall \mu_{k^*-1} \in \tilde{\mathcal{Q}}_{k^*-1} \\
\Rightarrow \quad & U_{k^*-1}^{\tilde{\mathcal{Q}}} = \min_{\mu_{k^*-1}} U_{k^*-1}^{\mu_{k^*-1}} > g^{\tilde{\mathcal{Q}}}(k^*-1) \\
\Rightarrow \quad & k^*-1 \notin \Gamma^{\tilde{\mathcal{Q}}}.
\end{aligned}$$

□

Of course, part (a) of the foregoing theorem is quite difficult to check. However, for applications, parts (b) and (c) are the interesting ones. Assertion (b) particularly holds for $(\tilde{\mu}_k)_k$, the instant worst-case assessment.

2.5 Conclusions

Having elaborated the "simple" and the risky Best-Choice problem, we came up with an adequate generalization to an ambiguous number of applicants. To solve this problem, we made use of the theory of optimal stopping with respect to multiple priors as set out in [Riedel, 09]: Agents assess expected reward in terms of multiple prior preferences or, equivalently, coherent risk measures.

When solving the problem, we have seen that a direct generalization of the risky to the ambiguous set-up is not feasible as time-consistency is impossibly achieved in that model. The problem, however, was seen to be solvable in terms of multiple assessments with a time-consistency assumption. By virtue of multiple prior Snell envelopes we have achieved conditions for the solution to consist of finitely many stopping islands. Furthermore, these stopping islands were entirely characterized by the multiple prior Snell envelope.

As we have seen, the major problem in solving the model is that the multiple prior Snell envelope is not the lower envelope of the individual Snell envelopes of the respective risky problems as the respective distribution is also incorporated in the instantaneous payoff. It is, however, the lower envelope of the Snell envelopes of the respective risky problems with artificially introduced instantaneous payoff g.

So far, we have not considered learning in our model as introduced in [Epstein & Schneider, 07]. On the contrary, an agent observing applicant, say, 10 may still put positive probability on having only, say, 5 applicants. However, updating assessments won't change the results as updating is "averaged out" in expected reward. Nevertheless, future research should introduce learning to this model: One might think of state dependent projections of sets of assessments, narrowing or widening contingent on available information. In this sense, assessments would emerge from learning.

Furthermore, ambiguous arrival times with a fixed horizon, extending [Stewart, 81] or [Bruss, 84] to the ambiguous case, should be considered; e.g. in case of Poisson arrival times, one might introduce multiplicity of parameters. Moreover, the case of ambiguity on the quality of applicants and an updating approach of beliefs over this set of priors is to elaborate in case of the partial information Best-Choice problem. After several observations, the set of priors on agents' quality is refined since several priors seem too unlikely; an approach as used in [Epstein & Schneider, 07].

A further extension is to consider the secretary problem when uncertainty over the number of applicants is not given as here but in terms of dynamic variational preferences or convex risk measures. A general theory for optimal stopping problems in that context can be found in the next chapter.

Chapter 3

Optimal Stopping with Dynamic Variational Preferences

3.1 Introduction

In our everyday life we face a broad variety of *optimal stopping problems*: We accept bids for our used car to sell or stop the process of potential marriage partners not knowing whether a more appropriate partner is still to come. On financial markets, agents try to maximize profits from American options. Hence, optimal stopping problems are not just of value for theoretical considerations but of great virtue in applications. All examples have in common that, on an abstract level, an agent has to find an optimal stopping time for some stochastic payoff process. The classical solution to this problem, as inter alia given in [Neveu, 75], assumes the agent to possess a unique subjective prior ruling the payoff process and to maximize expected payoff. In an *uncertain* environment however, there might not be a unique prior distribution: On incomplete financial markets, we might be faced with multiple equivalent martingale measures not being sure which one is ruling the world. Hence, with multiple possible distributions, a solution to the problem by virtue of simple expected utility maximization with respect to some subjective prior cannot be eligible: An alternative notion of "expected reward" has to be used. In this article, we hereto choose *dynamic variational preferences.*

[Riedel, 09] considers the problem to optimally stop an adapted payoff process $(X_t)_{t\in\mathbb{N}}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in\mathbb{N}})$ when expected reward is induced by *multiple prior preferences* or, equivalently, by *coherent risk measures.* By virtue of a *robust representation* theorem, expected reward for chosen stopping time τ is then given by a minimal expectation of the form $\inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[X_\tau]$ on a set of priors $\mathcal{Q}$. Several reasons for considering optimal stopping problems in terms of multiple prior preferences are stated therein: An ambiguity averse agent might not be able to completely determine the distribution governing the payoff process $(X_t)_t$ and hence apply this worst-case approach. Equivalently, when considering the problem from point of view of risk assessment, the above minimized expectation is, modulo a minus sign, the robust representation of coherent risk measures as seen in [Riedel, 04] or [Föllmer & Schied, 04]. We let the matter of justification rest at this point but mention the following example: In case $(X_t)_t$ is viewed as the payoff process of an American option in an incomplete financial market, a unique real world measure may induce several risk neutral martingale measures and, hence, a robust approach to expected payoff maximization with $\mathcal{Q}$ as the set of risk neutral measures seems appropriate.

As mentioned, the approach in [Riedel, 09] is based on multiple prior preferences introduced in [Gilboa & Schmeidler, 89], applied to a dynamic framework in [Epstein & Schneider, 03]. It can equivalently be stated in context of coherent risk measures introduced in [Artzner et al., 99] and applied to a dynamic setting in [Riedel, 04]. However, [Föllmer & Schied, 04] point out the limitations of the coherent approach: Due to *homogeneity* coherent risk measures do not account for *liquidity risk.* Secondly, the robust representation shows coherent risk measures to assess risk quite conservatively. Hence, the coherent approach is generalized to *convex risk measures* relaxing the homogeneity and sub-additivity assumption to a convexity condition resulting in a more liberal assessment of risk; in a dynamic context elaborately discussed in [Föllmer & Penner, 06] and [Cheridito et al, 06]. Furthermore, several fundamental risk measures are not coherent but convex as inter alia *entropic risk.* Equivalently, [Maccheroni et al., 06a] generalize the multiple priors approach to so called *variational preferences* and to *dynamic variational preferences* in [Maccheroni et al., 06b]. In a more general setup dynamic risk adjusted values or (concave) utilities are introduced in [Cheridito et al, 06] for stochastic processes. [Maccheroni et al., 06b] show dynamic multiple prior preferences to be a special class of dynamic variational preferences; [Cheridito et al, 06] show dynamic coherent risk measures to be a special class of dynamic convex risk measures.

For both approaches, the one in terms of variational preferences as well as the one in terms of convex risk measures, robust representations in terms of minimal penalized expected payoff (or maximal penalized expected loss) are achieved. These approaches are equivalent in the sense that the robust representations coincide up to a factor of -1. Under the assumption of risk neutrality but uncertainty aversion, a discount factor of unity and without intermediate payoffs, expected reward π_t for stopping the process $(X_t)_t$ with stopping strategy τ induced by dynamic variational preferences at time t is given by a robust representation of the form

$$\pi_t(X_\tau) = \operatorname*{ess\,inf}_{\mathbb{Q}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t(\mathbb{Q})\right), \tag{3.1}$$

for some *dynamic penalty* $(\alpha_t)_t$. The equivalent dynamic convex risk measure is then given as $\rho_t = -\pi_t$. Having in mind the robust representation of dynamic multiple prior preferences, it is immediate that these are a special case of dynamic variational preferences when the penalty is trivial, i.e. only achieves values null and infinity. In the same token, this holds for coherent risk measures as a special case of convex ones. It is beyond the scope of this article to discuss the axioms of variational preferences or convex risk measures, respectively, leading to the robust representation. We just take the representation as given and build our theory upon that. $(\alpha_t)_t$, formally derived by a Fenchel-Legendre transform, might be interpreted as an *ambiguity index*; this is inter alia done in [Maccheroni et al., 06a] and [Maccheroni et al., 06b]. We might intuitively think of $(\alpha_t)_t$ as an inverse likelihood of a distribution to be the ruling one: the larger the penalty, the less likely the agent assumes the respective distribution to be the true underlying one. Stated differently: Given two agents, one characterized by $(\alpha_t)_t^1$, the other by $(\alpha_t)_t^2$. If $(\alpha_t)_t^1 \geq (\alpha_t)_t^2$, then the first agent is less ambiguity averse. Equivalently, $(rho_t^1)_t$ assess risk more liberally. Throughout, we assume robust representation in terms of *minima* penalty $(\alpha_t)_t^{\min}$ uniquely characterizing the variational preference.

For dynamic models, the first question is how preferences or risk measures at distinct time periods are interrelated. An assumption that serves as a link between time periods is *time-consistency*, defined by virtue of $\pi_t = \pi_t(\pi_{t+1})$. Robust representation results showing equivalence of time-consistency and a condition on dynamic minimal penalty $(\alpha_t^{\min})_t$, called *no-gain condition*, are obtained in [Cheridito et al, 06], [Föllmer & Penner, 06], and [Maccheroni et al., 06b]. The basic idea is to represent minimal penalty as a sum of contingent penalties and a one-step-ahead penalty, thus connecting penalties in different time periods. Hence, the great advantage of the approach via dynamic variational preferences is that time consistency as a property of dynamic minimal penalty function $(\alpha_t^{\min})_t$ leads to a recursive robust representation in terms of minimal penalized expected utility. This property will elaborately be discussed in the next section. As shown in [Maccheroni et al., 06b], the no-gain condition on $(\alpha_t^{\min})_t$ reduces to *stability* of the set of priors in context of multiple prior preferences.

By virtue of the above expected reward in terms of minimal penalized expectation, results in this article constitute a generalization of the results in

[Riedel, 09] by applying optimal stopping to dynamic convex risk measures or dynamic variational preferences under the assumption of time-consistenty. In terms of a recursion formula we obtain a worst-case distribution for expected reward induced by dynamic variational preferences. It is however important that we do not obtain the elegant intuition in [Riedel, 09] that the agent behaves as expected utility maximizer with respect to the worst-case distribution since the penalty is not trivial and, hence, does not vanish. We make use of a *Snell envelope approach* to solve the problem at hand by showing equality of the value function and an appropriately generalized Snell envelope, called *variational Snell envelope*, for a finite horizon. In the infinite horizon case, we show the Bellman principle to hold for the value function. These results allow us to obtain an optimal stopping strategy recursively: We observe that the smallest optimal stopping time obeys well-known characteristics. A further result is a minimax theorem for optimal stopping under convex risk. In order to achieve our results, we introduce the notion of *variational (super-, sub-) martingales* and an accompanying variational martingale theory.

We then consider two prominent examples of dynamic convex risk measures. First, we have a look at dynamic *entropic risk measures* (or dynamic multiplier preferences). We state a robust representation of these measures and obtain quite intuitive results on the worst case measure for a specific kind of payoff processes. Secondly, we consider dynamic convex generalizations of *average value at risk* (AVaR) as introduced in [Cheridito & Li, 09]. As the natural dynamic extension of these risk measures is not time-consistent, we first achieve a dynamic version directly in terms of the definition of time-consistency. Secondly, we achieve a time-consistent version of generalized AVaR by virtue of a recursive construction in terms of the minimal penalty from the static version of generalized AVaR assumed to satisfy the no-gain condition when applied to a dynamic framework. As we see in the examples, when considering non-trivial penalty functions applications become more complex: in particular, independence, inevitably used in simple examples in [Riedel, 09], does not hold any longer. Nevertheless, the second example constitutes a tangible alternative to widely used VaR taking into account liquidity risk, satisfying time-consistency, and avoiding the problem of risk accumulation caused by VaR.

To prevent confusion, the reader should have the following in mind: Even though we stop a payoff process $(X_t)_t$, we do not need dynamic risk measures in all generality for stochastic processes $(X_t)_t$ as set out in [Cheridito et al, 06]. We only consider dynamic risk of a random variable X_τ. Hence, it suffices to consider the results from the theory of dynamic risk measures for end-period payoffs as in [Föllmer & Penner, 06].

[Schied, 07] applies an approach to optimal behavior on financial markets neglecting time-consistency. Agents maximize minimal penalized intertemporal utility as given above. Making use of convex conjugates, he achieves a minimax theorem similar to ours without using time-consistency. Hence, no constructive recursion for worst-case measures is achieved in that setup. However, we are convinced that time-consistency is not only a crucial property from a theoretical point of view to obtain explicit results but also intuitively justifiable.

Again, let us note the (mathematical) equivalence of dynamic convex risk measures and dynamic variational preferences. Slovenly, both are the same modulo a minus sign in terms of robust representation. The notion of time-consistency as well as necessary and sufficient conditions for it to be satisfied are in both approaches basically identical. Just the interpretation differs: The risk measure of a (financial) position reflects the amount of the numeraire needed to make the position acceptable or might be seen as penalized worst expected loss, whereas variational preferences are used to assess utilities. Hence, throughout the article we identify convex risk measures with variational preferences and coherent risk measures with multiple priors preferences. Particularly in the last section where we state examples, dynamic variational preferences are directly given in terms of dynamic convex risk measures.

The article is structured as follows: The next section defines the model, gathers the relevant assumptions and then states the optimal stopping problem in terms of a value function. This directly leads to the definition of variational supermartingales and an accompanying theory in Section 3. Section 4 contains the main results with proofs. Section 5 discusses some interesting examples. Thereafter, we conclude.

3.2 The Model

We now come up with a model to optimally stop a payoff process $(X_t)_{t \leq T}$, $T \in \mathbb{N} \cup \{\infty\}$, in discrete time. For this purpose let $(\Omega, \mathcal{F}, \mathbb{P}_0)$ be an underlying probability space with filtration $(\mathcal{F}_t)_{t \leq T}$, $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F} = \sigma\left(\bigcup_{t \leq T} \mathcal{F}_t\right)$, modeling the information process for the agent. $\mathbb{P}_0$ serves as a reference distribution. Consider an adapted process $(X_t)_{t \leq T}$ assumed to be essentially bounded.[1] If not stated otherwise, equalities are meant to hold $\mathbb{P}_0$-a.s. Let $\mathcal{M}^e(\mathbb{P}_0)$ denote the set of all probability measures on $(\Omega, \mathcal{F})$ that are *locally equivalent* to $\mathbb{P}_0$, i.e. for every t, $\mathbb{Q} \approx \mathbb{P}_0$ on $\mathcal{F}_t$. In particular, if $T < \infty$, this is just the set of all distributions equivalent on $\mathcal{F} = \mathcal{F}_T$. As we see in [Föllmer & Penner, 06], the assumption of locally equivalent distributions is justified from a mathematical point of view as the robust representation allows for only considering these distributions under suitable assumptions on convex risk measures. Intuitively, equivalence of distributions implies that the agent, not sure which distribution is the correct one, at least agrees upon which events are possible, i.e. have positive mass under all distributions, and which are not, i.e. have mass zero. We will elaborate on local equivalence being appropriate further in Chapter 4. Recall that a stopping time τ is an integer valued random variable such that $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \leq T$.

[1] This crucial assumption is mainly used for convenience. In the proofs it can be seen that weaker notions of integrability might be sufficient.

For $\omega \in \Omega$, we set $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$. Let $L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P}_0)$ be the space of all essentially bounded $\mathcal{F}$-measurable random variables. Analog, for $t \leq T$, let $L^\infty(\mathcal{F}_t) := L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}_0)$ be the space of all essentially bounded $\mathcal{F}_t$-measurable random variables.

3.2.1 Robust Representation of Time-Consistent Variational Preferences

For the payoff process $(X_t)_{t\leq T}$ an agent chooses a stopping time τ with respect to filtration $(\mathcal{F}_t)_{t\leq T}$ in order to maximize expected reward.

How do Agents assess Utility?

Given a stopping time τ, we first have to answer the following question:

Remark 3.2.1 (Initial Question)**.** *Given the agent is not able to entirely assess the ruling distribution of the payoff process and is uncertainty averse but risk neutral, how does expected reward look like?*

The assumption in *expected utility theory* would be that the agent has a subjective probability distribution, say $\mathbb{Q}$, of the payoff process and assesses expected reward by $\mathbb{E}^{\mathbb{Q}}[X_\tau]$.[2] [Riedel, 09] assumes the agent not being sure about the appropriate distribution of $(X_t)_t$ but knowing that it belongs to some convex set $\mathcal{Q} \subset \mathcal{M}^e(\mathbb{P}_0)$ with reference distribution $\mathbb{P}_0$. However, all elements in $\mathcal{Q}$ are assumed being equally probable. Then, *multiple prior* expected reward is given by $\inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[X_\tau]$.

In this article, we go a step further by assuming that an agent determines expected reward from stopping time τ in terms of *dynamic variational preferences* as introduced in [Maccheroni et al., 06b] or, equivalently, by a *dynamic convex risk measure* as in [Föllmer & Penner, 06]. As shown in

[2]We have implicitly assumed the agent to be risk-neutral as we will do throughout the article. Hence, we may choose the identity as Bernoulli state utility.

[Maccheroni et al., 06b] as well as in [Cheridito et al, 06], the agent then assesses *variational expected reward* at time t from stopping at τ by

$$\pi_t(X_\tau) = \operatorname*{ess\,inf}_{\mathbb{Q}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t(\mathbb{Q})\right), \tag{3.2}$$

where $(\alpha_t)_{t\leq T}$ denotes the *dynamic penalty*, also called dynamic ambiguity index in [Maccheroni et al., 06b]. This robust representations is obtained from the axioms of dynamic variational preferences. The penalty is achieved in terms of a Fenchel-Legendre transform. However, throughout this article, we take the robust representation as given and build our theory upon that; we do not consider the axiomatic approach to dynamic variational preferences. Equivalently, the axioms of dynamic convex risk measures $(\rho_t)_t$ lead to a robust representation satisfying $\rho_t = -\pi_t$.

Before stating appropriate assumptions and rigorous definitions, let us make a short note on the penalty's intuition: As set out in the introduction, the approach of assessing expected reward in term of minimal penalized expected utility emerges from the (dynamic) variational preferences axioms in [Maccheroni et al., 06b], as well as the convex risk measure axioms in [Cheridito et al, 06]. Robust representation results therein justify representing expected reward in the above manner. [Maccheroni et al., 06a] and [Maccheroni et al., 06b], as well as [Rosazza Gianin, 06] in the time-consistent case, incorporate a broad discussion of the penalty α_t: The penalty function is a measure for ambiguity aversion of an agent: If $\alpha_t^1 \geq \alpha_t^2$ for all t and all distributions, then agent 1 is less ambiguity averse than agent 2. Interpreted in another way, the penalty represents the subjective likelihood of a distribution to be the ruling one: The higher the value of α_t, the less likely the agent considers the respective distribution. In terms of a game against nature, α_t is usually interpreted as a cost nature has to bear for choosing a specific probability at time t. the penalty is – under the assumption of risk neutrality but ambiguity aversion – the characterization of the agent's preferences; unique as long as it is the *minimal penalty function*. Distinct examples of dynamic convex risk measures and dynamic variational preferences will be given later. As an extreme case, consider a distribution $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$ such that, for all t, $\alpha_t(\mathbb{Q}) = 0$ and ∞ for all $\mathbb{P} \neq \mathbb{Q}$: We achieve expected utility theory with subjective prior $\mathbb{Q}$. As shown in [Maccheroni et al., 06b], multiple prior expected reward with $\mathcal{Q} \subset \mathcal{M}^e(\mathbb{P}_0)$ is a special case of variational expected reward where $\alpha_t = 0$ on $\mathcal{Q}$ and ∞ else. In this sense, the present article constitutes a generalization of the approach in [Riedel, 09].

We now state a rigorous definition of the penalty $(\alpha_t)_{t\leq T}$ and appropriate assumptions for the above expected reward $(\pi_t)_{t\leq T}$ to be well defined as a robust representation of dynamic (time-consistent) variational preferences. There are several justifications for our definition of penalty: As seen in the respective literature as e.g. [Cheridito et al, 06] or [Föllmer & Penner, 06], our assumptions yield a representation of convex risk measures or variational preferences in terms of a penalty $(\alpha_t)_t$ satisfying the properties below.

Notation 3.2.2. *Define the set $\mathcal{M}$ of distributions in $\mathcal{M}^e(\mathbb{P}_0)$ by*

$$\mathcal{M} := \{\mathbb{Q} \quad | \quad \mathbb{Q}|_{\mathcal{F}_t} \approx \mathbb{P}_0|_{\mathcal{F}_t} \forall t, \quad \alpha_0(\mathbb{Q}) < \infty\},$$

where "$\approx$" means two probability distributions to be equivalent. Given the distribution $\mathbb{Q} \in \mathcal{M}$, $\mathbb{Q}|_{\mathcal{F}_t}$ denotes the restriction of $\mathbb{Q}$ to $\mathcal{F}_t$, i.e. the distribution of the process up to time t. As usual $\mathbb{Q}(\cdot|\mathcal{F}_t)$ denotes the conditional probability distribution of the process given history up to time t.

The following definitions are obtained from [Föllmer & Penner, 06] and [Maccheroni et al., 06b]:

Definition 3.2.3 (Dynamic Penalty & Time-Consistency)**.** (a) *We call a family $(\alpha_t)_t$ a* dynamic penalty *if each α_t satisfies:*

- *α_t is a mapping $\alpha_t : \mathcal{M} \to L^1_+(\mathcal{F}_t)$: For each $\mathbb{Q} \in \mathcal{M}$, $\alpha_t(\mathbb{Q})$ is an $\mathcal{F}_t$-measurable random variable with values in $\mathbb{R}_+$.*[3][4]

[3] More elaborately, for all $\omega \in \Omega$, $\alpha_t(\cdot)(\omega)$ is a function on the $\mathcal{F}_t$-bayesian updated distributions in $\mathcal{M}$, i.e. the *effective domain* satisfies $\text{effdom}(\alpha_t(\cdot)(\omega)) \subset \{\mathbb{Q}(\cdot|F_t) : \mathbb{Q} \in \mathcal{M}, \omega \in F_t \in \mathcal{F}_t\}$. Hence, when writing $\alpha_t(\mathbb{Q})$ we actually have in mind $\alpha_t(\mathbb{Q}(\cdot|\mathcal{F}_t))$.

[4] It can be seen in [Föllmer & Penner, 06], Lemma 3.5, that this domain of a penalty is

- *For all* $t \geq 0$, α_t *is* grounded, *i.e.* $\text{ess inf}_{\mathbb{Q} \in \mathcal{M}} \alpha_t(\mathbb{Q}) = 0$.
- α_t *is closed and convex,*[5] *i.e. convex as a mapping on* $\mathcal{M}$ *and closed in the sense that images of closed sets are again closed.*

(b) *At* t*, define the* acceptance set *by* $\mathcal{A}_t := \{X \in L^\infty | \rho_t(X) \leq 0\}$. *Then, we define the* minimal penalty $(\alpha_t^{\min})_t$ *by*

$$\alpha_t^{\min}(\mathbb{Q}) := \operatorname*{ess\,sup}_{X \in \mathcal{A}_t} \mathbb{E}^{\mathbb{Q}}[-X|\mathcal{F}_t].$$

for all $\mathbb{Q} \in \mathcal{M}$.[6]

(c) *Let* p_t *(resp.* q_t*) denote the density process of* $\mathbb{P}$ *(resp.* $\mathbb{Q}$*) with respect to* $\mathbb{P}_0$*, i.e.* $p_t := \frac{d\mathbb{P}}{d\mathbb{P}_0}\Big|_{\mathcal{F}_t}$*, where* $\frac{d\mathbb{P}}{d\mathbb{P}_0}$ *denotes the Radon-Nikodym derivative with respect to* $\mathbb{P}_0$*. For a stopping time* θ *define the "pasted distribution"* $\mathbb{P} \otimes_\theta \mathbb{Q}$ *by virtue of*

$$\frac{d(\mathbb{P} \otimes_\theta \mathbb{Q})}{d\mathbb{P}_0}\Bigg|_{\mathcal{F}_t} := \begin{cases} p_t & \text{if } t \leq \theta, \\ \frac{p_\theta q_t}{q_\theta} & \text{else.} \end{cases}$$

(d) *We call a dynamic penalty* $(\alpha_t)_t$ time-consistent *if it satisfies the following* no-gain *condition: for all* $t \geq 0$ *and* $\mathbb{Q}$ *we have*

$$\alpha_t(\mathbb{Q}) = \mathbb{E}^{\mathbb{Q}}\left[\alpha_{t+1}(\mathbb{Q})|\mathcal{F}_t\right] + \operatorname*{ess\,inf}_{\mathbb{P} \in \mathcal{M}} \alpha_t(\mathbb{Q} \otimes_{t+1} \mathbb{P}). \tag{3.3}$$

Notation 3.2.4. *Taking into account that* α_t *only depends on Bayesian updates, we simplify notation when appropriate and write*

$$\alpha_t(\mathbb{Q} \otimes_{t+1} \mathbb{P}) = \alpha_t\left(\frac{q_1 \cdots q_{t+1} p_{t+2} \cdots}{q_1 \cdots q_t}\right) = \alpha_t(q_{t+1} p_{t+2} \ldots).$$

well defined in case of relevant time-consistent dynamic convex risk measures as relevance allows to only consider the set of locally equivalent distributions in the robust representation and time-consistency in conjunction with relevance implies $\alpha_t(\mathbb{Q}) < \infty$ for all t. We call a dynamic convex risk measure $(\rho_t)_{t \leq T}$ *relevant*, if $\mathbb{P}_0[\rho_t(-\epsilon \mathbb{I}_A) > 0] > 0$ for all t, $\epsilon > 0$ and $A \in \mathcal{F}$ such that $\mathbb{P}_0[A] > 0$.

[5]This assumption is well defined by [Föllmer & Schied, 04], Remark 4.16.

[6]$(\alpha_t^{\min})_{t \leq T}$ is a penalty function in terms of (a).

Assumption 3.2.5. *Throughout this article we assume the agent to assess risk in terms of a relevant time-consistent dynamic convex risk measure* $(\rho_t)_{t\leq T}$ *on the set of essentially bounded* $\mathcal{F}$*-measurable random variables as in [Föllmer & Penner, 06] or, equivalently, assess utility in terms of time-consistent dynamic variational preferences* $(\pi_t)_{t\leq T}$ *for end-period payoffs as in [Maccheroni et al., 06b].* Note that we identify dynamic variational preferences with its robust representation of induced payoff. *Furthermore, we assume continuity from below for* $(\rho_t)_{t\leq T}$*, i.e. for all* $(X_n)_n \subset L^\infty$ *such that* $X_n \nearrow X$ *for some* $X \in L^\infty$*, we have* $\rho_t(X_n) \searrow \rho_t(X)$*. Equivalently, we assume continuity from below of* $(\pi_t)_{t\leq T}$*, i.e.* $\pi_t(X_n) \nearrow \pi_t(X)$ *for the above sequence.*

Definition 3.2.6. $(\rho_t)_t$ *is called* time-consistent *if it satisfies* $\rho_t = \rho_t(-\rho_{t+1})$ *for all* $t < T$*. Equivalently,* $\pi_t = \pi_t(\pi_{t+1})$*.*[7]

Remark 3.2.7. *[Cheridito et al, 06] and [Föllmer & Penner, 06] show that, under Assumption 3.2.5,* $(\rho_t)_{t\leq T}$ *and* $(\pi_t)_{t\leq T}$ *have a robust representation of the form*

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

with some dynamic penalty $(\alpha_t)_{t\leq T}$*. Furthermore, it is shown that this robust representation holds true in terms of, the minimal penalty* $(\alpha_t^{\min})_{t\leq T}$*, satisfying the no-gain condition (3.3) by the time-consistency assumption.*

Remark 3.2.8. *By virtue of the Fenchel-Legendre Transform, the minimal*

[7] In general, time-consistency is defined as: $\rho_t = \rho_t(-\rho_{t+s})$, $t, s \leq T$, $t + s \leq T$. In this sense, our definition of time-consistency is a special case, called "one-step time-consistency" in [Cheridito et al, 06]. However, for the proofs in this article, our definition is sufficient and, of course, always satisfied in the general case of time-consistency. On the other hand, one-step time-consistency implies general time-consistency under our continuity assumptions by Proposition 4.5 in [Cheridito et al, 06]. Hence, our definition of time-consistency in terms of "one-step time-consistency" is equivalent to the general notion of time-consistency.

penalty can be written as

$$\alpha_t^{\min}(\mathbb{Q}) = \operatorname*{ess\,sup}_{X \in L^\infty}(\mathbb{E}^{\mathbb{Q}}[-X|\mathcal{F}_t] - \rho_t(X))$$

for all $\mathbb{Q} \in \mathcal{M}$*. The term "minimal" is justified as the robust representation of* $(\rho_t)_{t\leq T}$ *or* $(\pi_t)_{t\leq T}$ *might allow for multiple penalties* $(\alpha_t)_{t\leq T}$*, but the minimal one satisfies*

$$\alpha_t^{\min}(\mathbb{Q}) \leq \alpha_t(\mathbb{Q})$$

for all $\mathbb{Q} \in \mathcal{M}$ *and* $(\alpha_t)_{t\leq T}$ *in the robust representation of* $(\rho_t)_{t\leq T}$ *or* $(\pi_t)_{t\leq T}$*.*

The minimal penalty uniquely characterizes the agent's preferences or, equivalently, risk attitude by virtue of the robust representation.

Assumption 3.2.9. *We assume robust representation in terms of the minimal penalty throughout this article.*

Remark 3.2.10. (a) *The no-gain condition on the minimal penalty* $(\alpha_t^{\min})_t$ *is equivalent to time-consistency of* $(\pi_t)_{t\leq T}$ *or* $(\rho_t)_{t\leq T}$*. Connecting distinct periods via the penalty function, this property leads to a recursive structure of penalty and hence of the value function of the optimal stopping problem. We will make this explicit later on.*

(b) *As stated in [Föllmer et al., 09], Remark* 1.1*, continuity from below of* π_t *or* ρ_t *implies continuity from above of either one. Continuity from above is equivalent to the existence of a robust representation of* π_t *(or* ρ_t*) in terms of minimal penalized expected payoff; continuity from below of* π_t *(or* ρ_t*) induces the worst case distribution to be achieved. We hence could change the* sup *into a* max*.* π_t *is continuous from above (below) if and only if the convex risk measure* ρ_t *is continuous from above (below).*

The intuition of equation (3.3), the no-gain condition, is the following: We might think that nature has to pay a penalty for choosing a specific distribution at time t: $\alpha_t^{\min}$. Nature may now accomplish the task of choosing a probability in two ways: On the left hand side of equation (3.3), it uses the time-consistent way by just choosing a probability $\mathbb{Q}$, pay the appropriate amount and do nothing in the next period and go with the conditional distribution $\mathbb{Q}(\cdot|\mathcal{F}_t)$. However, the right hand side describes the possibly time-inconsistent way of choosing a probability: It chooses today a distribution $\mathbb{P}$ that inuces the same distribution today as $\mathbb{Q}$ but may differ from tomorrow on and pays the amount $\alpha_t^{\min}(\mathbb{Q} \otimes_{t+1} \mathbb{P})$. In the second step, i.e. after realization of $\mathcal{F}_{t+1}$, nature may deviate and, conditionally on $\mathcal{F}_t$, choose a distribution $\mathbb{Q}$. If this time-inconsistent way of choosing a distribution is not less costly, we call $(\alpha_t^{\min})_t$ time-consistent. Equation (3.3) particularly tells us that the cost of choosing $\mathbb{Q}$ at time t can be decomposed into the sum of expected cost of choosing $\mathbb{Q}$'s conditionals at time $t+1$ and the cost of inducing $\mathbb{Q}|_{\mathcal{F}_{t+1}}$ as a so-called *one-period-ahead* marginal distribution of the payoff process at time t.

The no-gain condition on $(\alpha_t^{\min})_t$ is the generalization of the time-consistency condition in [Riedel, 09]: As shown in [Maccheroni et al., 06b], if (α_t) is trivial, i.e. only assumes values in $\{0, \infty\}$, the no-gain condition is equivalent to stability of the set of priors $\mathcal{Q} := \{\mathbb{Q} \in \mathcal{M} : \alpha_t^{\min}(\mathbb{Q}) = 0\}$. This also holds true in the not necessarily finite case as shown in e.g. in [Cheridito et al, 06].

In course of this section, we explicitly show time-consistency results when assuming a robust representation of dynamic convex risk measures or dynamic variational preferences in terms of minimal penalty.

Explicit Answer to the Initial Question

The following assumption answers the question how agents assess utility in the present set-up.

Assumption 3.2.11 (Main Assumption on Preferences)**.** *To sum up, for given τ we assume expected reward $(\pi_t)_{t \leq T}$ being continuous from below and possessing the robust representation as in Remark 3.2.7: for all t*

$$\pi_t(X_\tau) = \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right)$$

with dynamic minimal penalty $(\alpha_t^{\min})_{t\leq T}$ *assumed to be time-consistent, i.e. satisfying equation (3.3). This is equivalent to Assumption 3.2.5 but in terms of robust representation.*

Again, due to continuity from below, we can write the robust representation as ess min *instead of* ess inf*.*

In terms of dynamic variational preferences, time consistency is given by the recursion formula $\pi_t(\pi_{t+1}) = \pi_t$, which, as elaborately discussed below, in our case becomes for $\tau \geq t+1$

$$\begin{aligned} &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}\left[\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P})\right)\middle| \mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q})\right). \end{aligned}$$

Remark 3.2.12. *The following assumption is equivalent to* π_t *(or equivalently* ρ_t*) being continuous from below:*

$$\left\{ \frac{d\mathbb{P}}{d\mathbb{P}_0}\bigg|_{\mathcal{F}_t} \;\bigg|\; \alpha_t(\mathbb{P}) < c \right\},$$

for each $c \in \mathbb{R}$, $t \in \mathbb{N}$, *being relatively weakly compact in* $L^1(\Omega, \mathcal{F}, \mathbb{P}_0)$.[8]

Proof. Theorem 1.2 in [Föllmer et al., 09] states the assertion in an unconditional setting. Due to the properties of conditional expectations, the assertion also holds in our dynamic set-up. □

Remark 3.2.13 (Robust Representation as in Remark 3.2.7)**.** *We have now justified the representation in Remark 3.2.7. Relevance in conjunction with time-consistency allows us to only consider locally equivalent distributions in the robust representation and ensures* $\mathcal{M}$ *being non-empty as shown in [Föllmer & Penner, 06].*

The second part, continuity from below, then induces the worst case distribution to be attained in the coherent case, cp. [Föllmer & Schied, 04], Corollary 4.35, and Lemma 9 *and* 10 *in [Riedel, 09], and the minimal distribution to be achieved in our approach as will be seen in Proposition 3.3.6.*

[8] Or, assuming $\alpha_t^{\min}$ to be lsc, then just weakly compact. Due to time-consistency, we have $\alpha_t^{\min}(\mathbb{Q}) < \infty$ for all t whenever there exists one such t.

Remark 3.2.14 (Conditional Cash Invariance). *One of the axioms of dynamic variational preferences (and dynamic convex risk measures) is* conditional cash invariance*. In conjunction with a normalization assumption, this property becomes: for all $t \leq T$ and $\mathcal{F}_t$-measurable X, we have $\pi_t(X) = X$. As we do not consider the axiomatic approach, we immediately derive this property from the robust representation as $\alpha_t^{\min}$ is assumed to be grounded:*

$$\begin{aligned}\pi_t(X) &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &= X + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}) = X.\end{aligned}$$

The next result justifies to define time-consistency in terms of the penalty as it results in time-consistency of dynamic variational preferences $(\pi_t)_{t\leq T}$. Proposition 4.5 in [Cheridito et al, 06] shows in case of continuity from below our definition of time consistency, $\pi_t = \pi_t(\pi_{t+1})$, to be equivalent to the general definition, $\pi_t = \pi_t(\pi_{t+s})$. The proof of Proposition 3.2.15 is a special case of the proof of Theorem 4.22 in [Cheridito et al, 06]. It is explicitly stated here as it generates fruitful insights.

Proposition 3.2.15. *The no-gain condition, equation (3.3), implies time-consistency of dynamic variational preferences $(\pi_t)_{t\leq T}$, i.e. $\pi_t = \pi_t(\pi_{t+1})$ for $t < T$. More precisely, we have for all $(X_t)_{t\leq T}$ and $\tau \leq T$*

$$\begin{aligned}\pi_t(X_\tau) &= X_\tau \mathbb{I}_{\{\tau\leq t\}} + \pi_t(\pi_{t+1}(X_\tau))\mathbb{I}_{\{\tau\geq t+1\}} \\ &= \pi_t(X_\tau \mathbb{I}_{\{\tau\leq t\}} + \pi_{t+1}(X_\tau)\mathbb{I}_{\{\tau\geq t+1\}}) \\ &= \pi_t(\pi_{t+1}(X_\tau)).\end{aligned}$$

Proof. (i) $\tau \leq t$: In this case, X_τ is $\mathcal{F}_t$-measurable and in particular $\mathcal{F}_{t+1}$-measurable. Hence, by conditional cash invariance, we have

$$\pi_t(X_\tau) = X_\tau = \pi_{t+1}(X_\tau)$$

and hence $\pi_t(X_\tau) = \pi_t(\pi_{t+1}(X_\tau))$.

(ii) $\tau \geq t+1$: "$\leq$": If, for all $\mathbb{Q} \in \mathcal{M}$, we have

$$\alpha_t^{\min}(\mathbb{Q}) \leq \mathbb{E}^{\mathbb{Q}}\left[\alpha_{t+1}^{\min}(\mathbb{Q})|\mathcal{F}_t\right] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P}),$$

then, as $\operatorname{ess\,inf}_{\mathbb{R}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{R}) \leq \alpha_t^{\min}(\mathbb{Q})$, also

$$\alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P}) \leq \mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}}\left[\alpha_{t+1}^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})|\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}).$$

Now, consider $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}$ and $B \in \mathcal{F}$. Set $\frac{d\mathbb{Q}_3}{d\mathbb{P}_0} := \mathbb{I}_B \frac{d\mathbb{Q}_1}{d\mathbb{P}_0} + \mathbb{I}_{B^c}\frac{d\mathbb{Q}_1}{d\mathbb{P}_0}$. Then $\mathbb{Q}_3 \in \mathcal{M}$ and by the local property of minimal penalty, [Föllmer & Penner, 06], Lemma 3.3, we have $\alpha_t^{\min}(\mathbb{Q}_3) = \mathbb{I}_B\alpha_t^{\min}(\mathbb{Q}_1) + \mathbb{I}_{B^c}\alpha_t^{\min}(\mathbb{Q}_2)$. Define B as

$$B := \{\mathbb{E}^{\mathbb{Q}_2}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q}_2) \geq \mathbb{E}^{\mathbb{Q}_1}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q}_1)\}.$$

Then

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}_3}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q}_3) \\ = \; & \min\left\{\mathbb{E}^{\mathbb{Q}_1}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q}_1); \mathbb{E}^{\mathbb{Q}_2}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q}_2)\right\} \end{aligned}$$

showing the set

$$\left\{\mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P}) : \mathbb{P} \in \mathcal{M}\right\}$$

to be downward directed. Hence, there exists a sequence $(\mathbb{P}_n)_n \subset \mathcal{M}$ such that

$$\mathbb{E}^{\mathbb{P}_n}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P}_n) \searrow \pi_{t+1}(X_\tau).$$

As $(\alpha_t^{\min})_{t\leq T}$ is assumed to satisfy equation (3.3) and $(\pi_t)_{t\leq T}$ is assumed to be relevant, pasted distributions again have finite penalty. Hence, $\mathcal{M}$ is closed under pasting and we obtain for all $\mathbb{Q} \in \mathcal{M}$ and such $\mathbb{P}_n$:

$$\begin{aligned} \pi_t(X_\tau) &= \operatorname*{ess\,inf}_{\mathbb{P},\mathbb{Q}} \left(\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})\right) \\ &\leq \mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}_n}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P}_n) \\ &\leq \underbrace{\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}_n}[X_\tau|\mathcal{F}_t]}_{=\mathbb{E}^{\mathbb{Q}}[\mathbb{E}^{\mathbb{P}_n}[X_\tau|\mathcal{F}_{t+1}]|\mathcal{F}_t]} \\ &+ \underbrace{\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}_n}\left[\alpha_{t+1}(\mathbb{Q}\otimes_{t+1}\mathbb{P}_n)|\mathcal{F}_t\right]}_{=\mathbb{E}^{\mathbb{Q}}[\alpha_{t+1}(\mathbb{P}_n)|\mathcal{F}_t]} + \alpha_t^{\min}(\mathbb{Q}) \\ &= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{P}_n}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P}_n)\middle|\,\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}), \end{aligned}$$

i.e. for all $\mathbb{Q} \in \mathcal{M}$ we have

$$\pi_t(X_\tau) \quad \leq \quad \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{P}_n}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P}_n)\right|\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}).$$

Hence, letting $n \to \infty$, we achieve for all $\mathbb{Q} \in \mathcal{M}$

$$\pi_t(X_\tau) \leq \mathbb{E}^{\mathbb{Q}}\left[\pi_{t+1}(X_\tau)\right|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}).$$

Applying the essential infimum to this expression yields

$$\pi_t(X_\tau) \leq \pi_t(\pi_{t+1}(X_\tau)).$$

"$\geq$": Assuming

$$\alpha_t^{\min}(\mathbb{Q}) \geq \mathbb{E}^{\mathbb{Q}}\left[\alpha_{t+1}^{\min}(\mathbb{Q})|\mathcal{F}_t\right] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})$$

for all $\mathbb{Q} \in \mathcal{M}$, we obtain

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \\ \geq\ & \mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \mathbb{E}^{\mathbb{Q}}\left[\alpha_{t+1}^{\min}(\mathbb{Q})|\mathcal{F}_t\right] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P}) \\ \geq\ & \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}}\left[\mathbb{E}^{\mathbb{Q}}\left[X_\tau|\mathcal{F}_{t+1}\right] + \alpha_{t+1}^{\min}(\mathbb{Q})\right|\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})\right) \\ \geq\ & \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}}\left[\pi_{t+1}(X_\tau)\right|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})\right) \\ \geq\ & \pi_t(\pi_{t+1}(X_\tau)). \end{aligned}$$

Applying the essential infimum, we achieve

$$\pi_t(X_\tau) \geq \pi_t(\pi_{t+1}(X_\tau)).$$

□

As in [Maccheroni et al., 06b] we have the following result on the recursive structure of expected reward π_t at time t. However, we achieve this result for more general probability spaces but under the assumption of end-period payoffs, risk neutrality and a discount factor of unity.

Corollary 3.2.16. *For time-consistent dynamic minimal penalty* $(\alpha_t^{\min})_t$*, the time-t conditional expected reward from choosing stopping time* $\tau \leq T$ *satisfies*

$$\pi_t(X_\tau) = X_\tau \mathbb{I}_{\{\tau \leq t\}} + \operatorname*{ess\,inf}_{\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int \pi_{t+1}(X_\tau) d\mu + \gamma_t(\mu) \right) \mathbb{I}_{\{\tau \geq t+1\}},$$

where

$$\gamma_t(\mu) := \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{Q}) \quad \forall \mu \in \mathcal{M}|_{\mathcal{F}_{t+1}},$$

and $\mathcal{M}|_{\mathcal{F}_{t+1}}$ *denotes the set of all distributions in* $\mathcal{M}$ *restricted on* $\mathcal{F}_{t+1}$ *conditional on* $\mathcal{F}_t$*. To have this expression well-defined, we set* $\operatorname{ess\,inf}_{\mathbb{P} \in \mathcal{M}} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}) := \operatorname{ess\,inf}_{\mathbb{P} \in \mathcal{M}} \alpha_t^{\min}(\mathbb{Q} \otimes_{t+1} \mathbb{P})$ *with* $\mathbb{Q} \in \mathcal{M}$ *such that* $\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t) = \mu$*.*

Proof. By conditional cash invariance, we have

$$\begin{aligned} \pi_t(X_\tau) &= \pi_t(X_\tau \mathbb{I}_{\{\tau \leq t\}} + \pi_{t+1}(X_\tau) \mathbb{I}_{\{\tau \geq t+1\}}) \\ &= X_\tau \mathbb{I}_{\{\tau \leq t\}} + \pi_t(\pi_{t+1}(X_\tau)) \mathbb{I}_{\{\tau \geq t+1\}}. \end{aligned}$$

As π_{t+1} is $\mathcal{F}_{t+1}$-measurable we have, whenever $\tau \geq t+1$,

$$\begin{aligned} \pi_t(\pi_{t+1}(X_\tau)) &= \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[\pi_{t+1}(X_\tau)|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right) \\ &= \operatorname*{ess\,inf}_{\mathbb{R}, \mathbb{P} \in \mathcal{M}} \left(\underbrace{\mathbb{E}^{\mathbb{R} \otimes_{t+1} \mathbb{P}}[\pi_{t+1}(X_\tau)|\mathcal{F}_t]}_{\mathbb{E}^{\mathbb{R}|_{\mathcal{F}_{t+1}}}[\pi_{t+1}(X_\tau)|\mathcal{F}_t]} + \alpha_t^{\min}(\mathbb{R} \otimes_{t+1} \mathbb{P}) \right) \\ &= \operatorname*{ess\,inf}_{\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}, \mathbb{P} \in \mathcal{M}} \left(\mathbb{E}^{\mu}[\pi_{t+1}(X_\tau)|\mathcal{F}_t] + \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}) \right) \\ &= \operatorname*{ess\,inf}_{\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\mathbb{E}^{\mu}[\pi_{t+1}(X_\tau)|\mathcal{F}_t] + \underbrace{\operatorname*{ess\,inf}_{\mathbb{P} \in \mathcal{M}} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P})}_{=:\gamma_t(\mu)} \right). \end{aligned}$$

□

γ_t might be viewed as nature's penalty when choosing the one-period-ahead marginal μ. Hence, it is called *one-period-ahead penalty* in analogy to [Maccheroni et al., 06b]. In terms of γ_t, equation (3.3) becomes

$$\alpha_t^{\min}(\mathbb{Q}) = \mathbb{E}^{\mathbb{Q}}[\alpha_{t+1}^{\min}(\mathbb{Q})|\mathcal{F}_t] + \gamma_t(\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)). \tag{3.4}$$

Remark 3.2.17 (Bellman Principle for Nature). *Given $\tau \le T$, Corollary 3.2.16 can be rephrased as*

$$\pi_t(X_\tau) \quad = \quad \underset{\mathbb{Q}|_{\mathbb{F}_{t+1}} \in \mathcal{M}|_{\mathcal{F}_{t+1}}}{\text{ess inf}} \left(\mathbb{E}^{\mathbb{Q}|_{\mathbb{F}_{t+1}}}[\pi_{t+1}(X_\tau)|\mathcal{F}_t] + \gamma_t(\mathbb{Q}|_{\mathbb{F}_{t+1}}) \right) :$$

Indeed, this is immediately seen as $X_\tau \mathbb{I}_{\{\tau \le t\}}$ is $\mathcal{F}_t$-measurable, γ_t is grounded, and the conditional expectation is the unconditional one with respect to the conditional distribution.

Intuitively, this constitutes a Bellman principle for nature's choice of a worst-case distribution:[9] *Given the optimal (worst-case) distribution from time $t+1$ on, represented by its value π_{t+1}, nature chooses a minimizing one-period ahead conditional distribution $\mathbb{Q}|_{\mathbb{F}_{t+1}}$. Note, that the above expression is basically the same as the robust representation but in terms of a one-step-ahead problem. This insight is particularly adjuvant when constructing a worst-case distribution in Proposition 3.3.6 in terms of pasted one-period ahead conditional distributions.*

3.2.2 The Agent's Problem

Let $(X_t)_{t \le T}$, $T \in \mathbb{N} \cup \{\infty\}$, be a payoff process adapted to the filtered "reference space" $(\Omega, \mathcal{F}, \mathbb{P}_0, (\mathcal{F}_t)_{t \le T})$ and expected reward $(\pi_t)_{t \le T}$ with robust representation by virtue of time-consistent minimal penalty $(\alpha_t)_{t \le T}$. Recall that $\mathcal{M} := \{\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0) : \alpha_0^{\min}(\mathbb{Q}) < \infty\}$.

The agent solves the following problem by finding an appropriate stopping time τ with respect to $(\mathcal{F}_t)_{t \le T}$:

$$\sup_{0 \le \tau \le T} \inf_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_0] + \alpha_0^{\min}(\mathbb{Q}) \right) \tag{3.5}$$

among all stopping times that are universally finite, i.e.

$$\inf_{\mathbb{Q} \in \mathcal{M}} \mathbb{Q}[\tau < \infty] = 1.$$

[9] This should not be mixed up with the Bellman principle in the next chapter's theorems on optimal stopping: there, we achieve Bellman equations for the optimal stopping decision of the agent, not for the worst-case distribution decision of nature.

Definition 3.2.18 (Value Function, Snell Envelope)**.** (a) *For the problem at hand, the* value (function) $(V_t)_{t\leq T}$ *at time* $t \leq T$ *is given by*

$$V_t := \operatorname*{ess\,sup}_{T\geq\tau\geq t} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right). \tag{3.6}$$

(b) *For finite* T*, define the* variational Snell envelope $(U_t)_{t\leq T}$ *of* $(X_t)_{t\leq T}$ *with respect to dynamic minimal penalty* $(\alpha_t^{\min})_t$ *recursively by* $U_T := X_T$ *and*

$$U_t := \max\left\{X_t, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\} \quad \text{for } t < T. \tag{3.7}$$

(c) *Define the stopping time*

$$\tau^* := \inf\{t \geq 0 | U_t = X_t\}. \tag{3.8}$$

Due to time-consistency of $(\pi_t)_{t\leq T}$ the variational Snell envelope can also be written as:

$$\begin{aligned}
U_t &= \max\{X_t, \pi_t(U_{t+1})\} \\
&= \max\left\{X_t, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\} \\
&= \max\left\{X_t; \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int \pi_{t+1}(U_{t+1})d\mu + \gamma_t(\mu)\right)\right\} \\
&= \max\left\{X_t; \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int U_{t+1}d\mu + \gamma_t(\mu)\right)\right\}
\end{aligned}$$

Subsequently, we show that the value function and the variational Snell envelope coincide when T is finite. In the infinite time-horizon case, we show the Bellman principle to hold for the value function allowing for recursive solutions. Furthermore, it follows that τ^* is an optimal stopping time, i.e. a solution to the initial problem. Note, that the variational Snell envelope coincides with the multiple prior Snell envelope in case of multiple prior preferences as introduced in [Riedel, 09]. It coincides with the "good old" Snell envelope as e.g. set out in [Neveu, 75] in case of a unique subjective prior.

3.3 Variational Supermartingales

From the approach to optimal stopping in terms of Snell envelopes as e.g. set out in [Neveu, 75] or more generally with multiple prior Snell envelopes as in [Riedel, 09], we know that the value function satisfies some kind of supermartingale property.[10] The sleight of hand is always showing the value function to be "some kind" of martingale until the optimal stopping time and "some kind" of supermartingale thereafter. Hence, in order to solve the agent's problem, we have to come up with an appropriate notion of martingale for dynamic variational preferences: the following definition generalizes the notion of multiple prior (sub-, super-) martingales in [Riedel, 09]:

Definition 3.3.1. *Given a time-consistent dynamic minimal penalty function* $(\alpha_t^{\min})_{t\in\mathbb{N}}$*, let* $(M_t)_{t\in\mathbb{N}}$ *be an* $(\mathcal{F}_t)_{t\in\mathbb{N}}$*-adapted process with* $\mathbb{E}^{\mathbb{Q}}[|M_t|] < \infty$ *for all* $t \leq T$ *and all* $\mathbb{Q} \in \mathcal{M}$*.* $(M_t)_{t\in\mathbb{N}}$ *is called a* variational (sub-, super-) martingale *with respect to* $(\alpha_t^{\min})_{t\in\mathbb{N}}$ *if the following relation holds for* $t < T$*:*

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = (\geq, \leq) M_t.$$

In Lemma 6, [Riedel, 09] achieves a quite elegant way to characterize the concepts of multiple prior (sub-, super-) martingales with respect to some time-consistent set $\mathcal{Q}$ of distributions in terms of (sub-, super-) martingales with respect to a worst-case distribution $\mathbb{P}^* \in \mathcal{Q}$. However, this result is owed to the simple structure of α_t in the multiple priors case. Under variational preferences, we do not achieve such an elegant lemma, but nevertheless can state a similar result for variational supermartingales as being a supermartingale "modulo penalty" with respect to some worst-case distribution $\mathbb{Q}^* \in \mathcal{M}$. However, non-triviality of the minimal penalty in case of variational preferences is the reason why the intuition of an agent behaving as expected utility maximizer under the worst case distribution does *not* carry over from

[10] Having a look in the respective chapters in [Neveu, 75], it can be seen that the term Snell envelope is not explicitly used therein; the solution procedure, however, is identical.

[Riedel, 09]. As in Riedel's Lemma 6, the worst-case distribution is achieved recursively: At each time t, the *worst-case conditional one-step-ahead distribution* is chosen. In [Riedel, 09], time-consistency is needed to ensure the recursively pasted distribution to be again in the set of priors $\mathcal{Q}$. By definition of $\mathcal{M}$ and equation (3.3), we obviously have that pasted distributions are again in $\mathcal{M}$: $\alpha_{t+1}^{\min}(\mathbb{Q}) < \infty$ implies $\alpha_t^{\min}(\mathbb{Q}) < \infty$. However, the most important part in our construction of a worst-case $\mathbb{Q}^*$ is that, given equation (3.3), pasting of worst-case one-step-ahead distributions is consistent with being of worst-case type given equation (3.3) *Having achieved a worst-case distribution from $t+1$ onwards, we paste this with the one-step-ahead worst-case conditional distribution from t to $t+1$ and achieve the worst-case distribution from time t onwards.*

For the next result analog to Lemma 6 in [Riedel, 09], we need several lemmata directly generalizing the respective ones in [Riedel, 09] (Lemmata 9 and 10) to dynamic variational preferences applying interim results from [Föllmer & Penner, 06]. Throughout we assume the minimal penalty to satisfy equation (3.3).

Lemma 3.3.2. *For all $\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}$ there exists $\mathbb{P}^* \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})$ such that $\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}^*) = \text{ess inf}_{\mathbb{P}\in\mathcal{M}(\cdot|\mathcal{F}_{t+1})} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P})$.*

Proof. By the weak compactness assumption on the set of density processes (equivalent to continuity from below), it is sufficient to show that there exists a sequence $(\mathbb{P}_n)_n \subset \mathcal{M}(\cdot|\mathcal{F}_{t+1})$ such that

$$\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_n) \searrow \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}(\cdot|\mathcal{F}_{t+1})} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}).$$

Hence, it suffices to show that for all $\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}$, the set

$$\{\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_n) : \mathbb{P} \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})\}$$

is downward directed, i.e. for every $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})$, there exists a $\mathbb{P}_3 \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})$ sucht that

$$\min\left\{\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_1), \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_2)\right\} = \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_3).$$

Indeed, set $A := \{\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_1) < \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_2)\}$ and define $\mathbb{P}_3$ by virtue of

$$\frac{d\mathbb{P}_3}{d\mathbb{P}_0} := \mathbb{I}_A \frac{d\mathbb{P}_1}{d\mathbb{P}_0} + \mathbb{I}_{A^C} \frac{d\mathbb{P}_2}{d\mathbb{P}_0}.$$

By Lemma 3.3 in [Föllmer & Penner, 06], we have

$$\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_3) = \mathbb{I}_A \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_1) + \mathbb{I}_{A^C} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_2)$$

since $\mu \otimes_{t+1} \mathbb{P}_3 = (\mu \otimes_{t+1} \mathbb{P}_1)\mathbb{I}_A + (\mu \otimes_{t+1} \mathbb{P}_2)\mathbb{I}_{A^C}$. Hence, we have

$$\min\left\{\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_1), \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_2)\right\} = \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_3),$$

which concludes the proof. □

Lemma 3.3.3. *Let $Z \in L^\infty(\Omega, \mathcal{F}_T, \mathbb{P}_0)$. Then, for any stopping time τ, the set*

$$\left\{\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}) : \mathbb{Q} \in \mathcal{M}, \mathbb{P}^\tau|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}\right\}$$

is downward directed, i.e. for any $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}$ with $\mathbb{Q}_1|_{\mathcal{F}_\tau} = \mathbb{Q}_2|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$, there exists $\mathbb{Q}_3 \in \mathcal{M}$ with $\mathbb{Q}_3|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$ such that

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}_3}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_3) \\ = \quad & \min\left\{\mathbb{E}^{\mathbb{Q}_1}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_1); \mathbb{E}^{\mathbb{Q}_2}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_2)\right\}. \end{aligned}$$

Proof. Let $\mathbb{Q}_1$ and $\mathbb{Q}_2$ be chosen as above. Consider some arbitrary set $B \in \mathcal{F}_\tau$ and define $\mathbb{Q}_3$ by virtue of

$$\frac{d\mathbb{Q}_3}{d\mathbb{P}_0} := \mathbb{I}_B \frac{d\mathbb{Q}_1}{d\mathbb{P}_0} + \mathbb{I}_{B^C} \frac{d\mathbb{Q}_2}{d\mathbb{P}_0}.$$

We have $\mathbb{Q}_3 \in \mathcal{M}$, $\mathbb{Q}_3|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$, and by [Föllmer & Penner, 06], Lemma 3.3, we have the so called *local propery* for the minimal penalty

$$\alpha_\tau^{\min}(\mathbb{Q}_3) = \mathbb{I}_B \alpha_\tau^{\min}(\mathbb{Q}_1) + \mathbb{I}_{B^C} \alpha_\tau^{\min}(\mathbb{Q}_2) < \infty.$$

Now, define $B \in \mathcal{F}_\tau$ as

$$\begin{aligned} B := \{\omega \in \Omega |\, & \mathbb{E}^{\mathbb{Q}_2}[Z|\mathcal{F}_\tau](\omega) + \alpha_\tau^{\min}(\mathbb{Q}_2)(\omega) \\ & \geq \mathbb{E}^{\mathbb{Q}_1}[Z|\mathcal{F}_\tau](\omega) + \alpha_\tau^{\min}(\mathbb{Q}_1)(\omega)\right\}. \end{aligned}$$

Then, by definition of $\mathbb{Q}_3$ and the local property, we have

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}_3}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_3) \\ = & \left(\mathbb{E}^{\mathbb{Q}_1}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_1)\right)\mathbb{I}_B + \left(\mathbb{E}^{\mathbb{Q}_2}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_2)\right)\mathbb{I}_{B^C} \\ = & \min\left\{\mathbb{E}^{\mathbb{Q}_1}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_1); \mathbb{E}^{\mathbb{Q}_2}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_2)\right\}, \end{aligned}$$

which completes the proof. □

Lemma 3.3.4. *Let $Z \in L^\infty(\Omega, \mathcal{F}_s, \mathbb{P}_0)$, $s \leq T$, and τ a stopping time.*[11] *Then there exists $\mathbb{P}^\tau \in \mathcal{M}$ sucht that $\mathbb{P}^\tau|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$ and*

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q})\right) = \mathbb{E}^{\mathbb{P}^\tau}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{P}^\tau)\mathbb{I}_{\{s>\tau\}}.$$

Proof. In case $\tau \geq s$, the assertion obviously holds true by conditional cash invariance: Both sides of the equation equal Z.[12]

Hence, we consider the case $\tau < s$. To show: $\exists\ (\mathbb{P}_m)_m \subset \mathcal{M}$ with $\mathbb{P}_m|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$ such that

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q})\right) = \lim_{m\to\infty} \mathbb{E}^{\mathbb{P}_m}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{P}_m).$$

By the weak closedness assumption, such a sequence $(\mathbb{P}_m)_m$ then weakly converges to some $\mathbb{P}_\infty \in \mathcal{M}$ that satisfies

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q})\right) = \mathbb{E}^{\mathbb{P}_\infty}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{P}_\infty).$$

Setting $\mathbb{P}_\infty =: \mathbb{P}^\tau$ then concludes the proof.

It leaves to prove existence of a sequence $(\mathbb{P}_m)_m \subset \mathcal{M}$ with the above properties: As in the proof of Lemma 10 in [Riedel, 09], Bayes rule as well as the dependence of α_τ only on the $\mathcal{F}_\tau$-conditional distribution allows us to restrict attention to $\mathbb{Q} \in \mathcal{M}$ such that $\mathbb{Q} = \mathbb{P}_0$ on $\mathcal{F}_t$. This is made explicit

[11] We actually state the assertion in a more general fashion than needed: For our results it suffices to have a fixed stopping period $t \in \mathbb{N}$.

[12] Of course, the assertion in the lemma would still be correct without the indicator function attached to the penalty. Then, in case $\tau \geq s$, the minimizing distribution $\mathbb{P}^\tau$ is just the one for which $\alpha_\tau^{\min}(\mathbb{P}^\tau) = 0$. However, our form makes more explicit that the α-term vanishes in that case.

in Corollary 2.4 in [Föllmer & Penner, 06]. Hence, existence of the sequence is assured if we can show the set

$$\left\{\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}) \quad : \quad \mathbb{Q} \in \mathcal{M}, \quad \mathbb{P}^\tau|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}\right\}.$$

to be downward directed as achieved in Lemma 3.3.3. □

Corollary 3.3.5 (from Lemma 3.3.4)**.** *For all* $Z \in L^\infty(\Omega, \mathcal{F}_{t+1}, \mathbb{P}_0)$*,* $\exists \mu^* \in \mathcal{M}|_{\mathcal{F}_{t+1}}$ *such that*

$$\operatorname*{ess\,inf}_{\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\mathbb{E}^\mu[Z|\mathcal{F}_t] + \gamma_t(\mu)\right) = \mathbb{E}^{\mu^*}[Z|\mathcal{F}_t] + \gamma_t(\mu^*).$$

Proposition 3.3.6. *Let* $(M_t)_{t\in\mathbb{N}}$ *be an adapted process and* $(\alpha_t^{\min})_{t\in\mathbb{N}}$ *a time-consistent minimal dynamic penalty function.*
(a) *If* $(M_t)_{t\in\mathbb{N}}$ *is a* $\mathbb{Q}$*-submartingale for all* $\mathbb{Q} \in \mathcal{M}$*, then* $(M_t)_{t\in\mathbb{N}}$ *is a variational submartingale with respect to* $(\alpha_t^{\min})_t$*.*
(b) $(M_t)_{t\in\mathbb{N}}$ *is a variational supermartingale with respect to* $(\alpha_t^{\min})_{t\in\mathbb{N}}$ *if and only if there exist a* $\mathbb{Q}^* \in \mathcal{M}$ *such that* $(M_t)_{t\in\mathbb{N}}$ *is a* $\mathbb{Q}^*$*-supermartingale "modulo penalty", i.e.*

$$\mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*) \leq M_t.$$

In particular, $(M_t)_{t\in\mathbb{N}}$ *is a* $\mathbb{Q}^*$*-supermartingale, i.e.* $\mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] \leq M_t$*.*

Proof. ad (a): Let $(M_t)_{t\in\mathbb{N}}$ be a submartingale for every $\mathbb{Q} \in \mathcal{M}$, i.e.

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] \geq M_t \quad \forall \mathbb{Q} \in \mathcal{M} \\
\Rightarrow\ &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\} \\
&\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}) \\
&= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] \geq M_t.
\end{aligned}$$

This shows (a).

ad (b): "⇐" Let $\mathbb{Q}^* \in \mathcal{M}$ be such that $M_t \geq \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*)$. Then obviously, $M_t \geq \operatorname{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\}$ and hence $(M_t)_t$

is a variational supermartingale w.r.t. $(\alpha_t)^{\min}_{t\in\mathbb{N}}$ as well a $\mathbb{Q}^*$-supermartingale: $M_t \geq \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*) \geq \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t]$.

"$\Rightarrow$" By making use of Corollary 3.2.16, we will explicitly construct a *worst-case distribution* $\mathbb{Q}^* \in \mathcal{M}$ that satisfies

$$\begin{aligned} M_t &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &= \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*) \end{aligned}$$

for $t < T$. Let $\mathcal{M}(\cdot|\mathcal{F}_t)$ denote the set of all distributions in $\mathcal{M}$ conditional on $\mathcal{F}_t$ and $\mathcal{M}|_{\mathcal{F}_t}$ as defined in Corollary 3.2.16. We use that, due to continuity from below, the infima in the robust representation of preferences are achieved and, hence, are actually minima. We nevertheless state the equations in terms of infima as this is common in the respective literature. We have

$$\begin{aligned} M_t &\geq \pi_t(M_{t+1}) \\ &= \operatorname*{ess\,inf}_{\mathbb{Q}|_{\mathcal{F}_{t+1}}\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\mathbb{E}^{\mathbb{Q}|_{\mathcal{F}_{t+1}}}[\underbrace{\pi_{t+1}(M_{t+1})}_{=M_{t+1}}|\mathcal{F}_t] + \gamma_t(\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t))\right) \\ &\qquad \text{by Corollary 3.2.16} \\ &= \operatorname*{ess\,inf}_{\mathbb{Q}|_{\mathcal{F}_{t+1}}\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\mathbb{E}^{\mathbb{Q}|_{\mathcal{F}_{t+1}}}[M_{t+1}|\mathcal{F}_t] + \gamma_t(\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t))\right) \\ &= \mathbb{E}^{\mathbb{Q}^*|_{\mathcal{F}_{t+1}}}[M_{t+1}|\mathcal{F}_t] + \gamma_t(\mathbb{Q}^*|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)) \\ &\qquad \text{with } \mathbb{Q}^*|_{\mathcal{F}_{t+1}} \text{ as achieved in Corollary 3.3.5} \\ &= \mathbb{E}^{\mathbb{Q}^*|_{\mathcal{F}_{t+1}}}[M_{t+1}|\mathcal{F}_t] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}(\cdot|\mathcal{F}_{t+1})} \alpha_t^{\min}(\mathbb{Q}^*|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t) \otimes_{t+1} \mathbb{P}) \\ &\qquad \text{by definition of } \gamma_t \\ &= \mathbb{E}^{\mathbb{Q}^*|_{\mathcal{F}_{t+1}}\otimes_{t+1}\mathbb{Q}^*(\cdot|\mathcal{F}_{t+1})}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t) \otimes_{t+1} \mathbb{Q}^*(\cdot|\mathcal{F}_{t+1})) \\ &\qquad \text{by Lemma 3.3.2 and Bayes rule on the first summand} \\ &= \mathbb{E}^{\mathbb{Q}^*(\cdot|\mathcal{F}_t)}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*(\cdot|\mathcal{F}_t)) \\ &\qquad \text{by definition of } \otimes_{t+1} \\ &= \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*), \end{aligned}$$

where $\mathbb{Q}^*(\cdot|\mathcal{F}_t) := \mathbb{Q}^*|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t) \otimes_{t+1} \mathbb{Q}^*(\cdot|\mathcal{F}_{t+1})$ is the pasting of the $\mathbb{Q}^*|_{\mathcal{F}_s}$'s,

$s \geq t$, and $\mathbb{Q}^*$ the respective recursive pasting down to time 0. The last equality makes use of the fact that the penalty only depends on conditional distributions[13] and that the conditional expectation is the unconditional one with respect to the conditional distribution. □

In the foregoing proposition, we see that a variational submartingale with respect to some minimal penalty $(\alpha_t^{\min})_{t\in\mathbb{N}}$ does not need to be a submartingale with respect to some $\mathbb{Q} \in \mathcal{M}$. This insight limits the mathematical theory obtained later. Luckily however, our economic results only rely on the properties of variational supermartingales.

Remark 3.3.7. *As seen in the lemmata, the foregoing assertion can be generalized to: $\exists\ \mathbb{Q}^* \in \mathcal{M}$ such that $\forall t, s$ we have*

$$\mathbb{E}^{\mathbb{Q}^*}[M_s|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*)\mathbb{I}_{\{s>t\}} \leq M_t.$$

Indeed, if $s \leq t$, due to projection property of conditional expectation, the left hand side reduces to M_s as, in that case, M_s is $\mathcal{F}_t$-measurable, and $\alpha_t^{\min}$ is assumed to be grounded.

In the same token as [Riedel, 09], we generalize standard results for supermartingales to our notion of variational supermartingales. First, we show the fundamental Doob Decomposition in martingale theory to still be valid in our framework. Thereafter, we show an Optional Sampling theorem for variational supermartingales.

Proposition 3.3.8 (Doob Decomposition)**.** *Let $(S_t)_{t\in\mathbb{N}}$ be a variational supermartingale with respect to time-consistent minimal penalty $(\alpha_t^{\min})_{t\in\mathbb{N}}$. Then there exists a variational martingale $(M_t)_{t\in\mathbb{N}}$ with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ and a predictable non-decreasing process $(A_t)_{t\in\mathbb{N}}$, $A_0 = 0$, such that $S_t = M_t - A_t$ for all t and this decomposition is unique.*

[13] I.e. the effective domain of the dynamic minimal penalty is the set of conditionals and, hence, our intuitive notation here is justified.

Proof. (a) Uniqueness: Let $S = M - A$ as above. Then

$$\begin{aligned} & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t + A_{t+1} - A_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1} - M_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) - M_t = 0, \end{aligned}$$

as M was assumed to be a variational martingale. Since $\alpha_t^{\min}$ is uniquely given (as ρ_t is assumed to be relevant) and A is assumed to be predictable, we have

$$A_{t+1} = A_t - \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right).$$

This shows uniqueness of A and thus also of M.

(b) Existence: Define $(A_t)_{t\in\mathbb{N}}$ by virtue of $A_0 = 0$ and

$$A_{t+1} := A_t - \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right).$$

Then, $A_{t+1} \in \mathcal{F}_t$, i.e. $(A_t)_{t\in\mathbb{N}}$ is predictable and, moreover, it is non-decreasing. Set $M_t := S_t + A_t$. It is left to show that $(M_t)_{t\in\mathbb{N}}$ is a variational martingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$:

$$\begin{aligned} & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) - M_t \\ = & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1} - M_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t + A_{t+1} - A_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = & A_{t+1} - A_t + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = & 0, \end{aligned}$$

where the last equality follows by definition of $(A_t)_{t\in\mathbb{N}}$ and the second to last because of its predictability. □

Proposition 3.3.9 (Optional Sampling)**.** *Let $(S_t)_{t\in\mathbb{N}}$ be a variational supermartingale with respect to the time-consistent minimal penalty $(\alpha_t^{\min})_{t\in\mathbb{N}}$ and $\sigma \leq \tau$ be universally finite stopping times. Then*

$$S_\sigma \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right).$$

Proof. We know from Proposition 3.3.6 that there exists a "worst case" distribution $\mathbb{P}^* \in \mathcal{M}$ such that

$$S_t \geq \mathbb{E}^{\mathbb{P}^*}[S_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*).$$

Whereas the proof of optional sampling with multiple priors in [Riedel, 09] is immediate as the minimal penalty vanishes for the worst case distribution, we have to mimic the proof of the original optional sampling theorem and carry with us the penalty. The proof is accomplished in two steps:

(i) First, we show that for fixed $N \in \mathbb{N}$ a stopped "supermartingale modulo penalty" $(S_{N\wedge t})_{t\in\mathbb{N}}$ is again one such. I.e.[14]

$$S_{N\wedge t} \geq \mathbb{E}^{\mathbb{P}^*}[S_{N\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>t\}}. \tag{3.9}$$

Indeed, we have

$$\begin{aligned}
S_{N\wedge t} &= S_0 + \sum_{k=1}^{t} \mathbb{I}_{\{N\geq k\}}(S_k - S_{k-1}) \\
&\geq S_0 + \sum_{k=1}^{t} \mathbb{I}_{\{N\geq k\}}(S_k - S_{k-1}) \\
&\quad + \mathbb{I}_{\{N\geq t+1\}}(\mathbb{E}^{\mathbb{P}^*}[S_{t+1} - S_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*)) \\
&= \mathbb{E}^{\mathbb{P}^*}\left[S_0 + \sum_{k=1}^{t} \mathbb{I}_{\{N\geq k\}}(S_k - S_{k-1}) + \mathbb{I}_{\{N\geq t+1\}}(S_{t+1} - S_t)|\mathcal{F}_t\right] \\
&\quad + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>t\}} \\
&= \mathbb{E}^{\mathbb{P}^*}[S_{N\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>t\}},
\end{aligned}$$

where the inequality holds with equality for variational martingales.

(ii) Note: By (i), we have for a variational martingale $(M_t)_{t\in\mathbb{N}}$

$$\mathbb{E}^{\mathbb{P}^*}[M_{N\wedge t}] = \mathbb{E}^{\mathbb{P}^*}[M_{N\wedge(t+1)} + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>t\}}]$$

[14] It might, at first sight, seem quite confusing that there is an indicator function adjacent to the penalty in equation (3.9) as already stated in Remark 3.3.7. However, the intuition is that a time $t > N$, i.e. when the process has already been stopped, its value is known since S_N is $\mathcal{F}_t$-measurable and nature does not have to be penalized any more as it does not choose any distribution.

and in particular

$$\mathbb{E}^{\mathbb{P}^*}[M_0] = \mathbb{E}^{\mathbb{P}^*}\left[M_{N\wedge t} + \sum_{i=0}^{t-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}}\right] \quad \forall N, t.$$

Moreover, it holds

$$\begin{aligned} & \lim_{t\to\infty} \mathbb{E}^{\mathbb{P}^*}[M_{N\wedge t} + \sum_{i=0}^{t-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}}] \\ = \; & \mathbb{E}^{\mathbb{P}^*}[M_N] + \mathbb{E}^{\mathbb{P}^*}[\sum_{i=0}^{\infty} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}}]. \end{aligned}$$

Hence,

$$\mathbb{E}^{\mathbb{P}^*}[M_0] = \mathbb{E}^{\mathbb{P}^*}[M_N] + \mathbb{E}^{\mathbb{P}^*}\left[\sum_{i=0}^{\infty} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}}\right].$$

We set $\sum_{i=0}^{\infty} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}} =: \sum_{i=0}^{N-1} \alpha_i^{\min}(\mathbb{P}^*)$. Now, let $B \in \mathcal{F}_\sigma$ and define

$$\begin{aligned} S^B &:= \sigma\mathbb{I}_B + \kappa\mathbb{I}_{B^C}, \\ T^B &:= \tau\mathbb{I}_B + \kappa\mathbb{I}_{B^C}, \end{aligned}$$

where $\kappa := \sup N$. Then S^B and T^B are stopping times and we have by equation (3.3)

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}^*}\left[M_\sigma\mathbb{I}_B + \sum_{i=0}^{\sigma-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_B\right] + \mathbb{E}^{\mathbb{P}^*}\left[M_\kappa\mathbb{I}_{B^c} + \sum_{i=0}^{\kappa-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{B^c}\right] \\ = \; & \mathbb{E}^{\mathbb{P}^*}\left[M_{S^B} + \sum_{i=0}^{S^B-1} \alpha_i^{\min}(\mathbb{P}^*)\right] \\ = \; & \mathbb{E}^{\mathbb{P}^*}[M_0] \\ = \; & \mathbb{E}^{\mathbb{P}^*}\left[M_{T^B} + \sum_{i=0}^{T^B-1} \alpha_i^{\min}(\mathbb{P}^*)\right] \\ = \; & \mathbb{E}^{\mathbb{P}^*}\left[M_\tau\mathbb{I}_B + \sum_{i=0}^{\tau-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_B\right] + \mathbb{E}^{\mathbb{P}^*}\left[M_\kappa\mathbb{I}_{B^c} + \sum_{i=0}^{\kappa-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{B^c}\right], \end{aligned}$$

and hence

$$\mathbb{E}^{\mathbb{P}^*}[M_\sigma\mathbb{I}_B] = \mathbb{E}^{\mathbb{P}^*}\left[(M_\tau + \sum_{i=\sigma}^{\tau-1} \alpha_i^{\min}(\mathbb{P}^*))\mathbb{I}_B\right].$$

Since this holds true for all $B \in \mathcal{F}_\sigma$, we have

$$\mathbb{E}^{\mathbb{P}^*}[M_\sigma|\mathcal{F}_\sigma] = \mathbb{E}^{\mathbb{P}^*}[M_\tau + \sum_{i=\sigma}^{\tau-1} \alpha_i^{\min}(\mathbb{P}^*)|\mathcal{F}_\sigma],$$

i.e.

$$M_\sigma = \mathbb{E}^{\mathbb{P}^*}\left[M_\tau + \sum_{i=\sigma+1}^{\tau-1} \alpha_i^{\min}(\mathbb{P}^*)\middle|\mathcal{F}_\sigma\right] + \alpha_\sigma^{\min}(\mathbb{P}^*)\mathbb{I}_{\{\tau>\sigma\}}.$$

Summing up, we have shown for $\tau > \sigma$[15]

$$\begin{aligned} M_\sigma &\geq \mathbb{E}^{\mathbb{P}^*}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{P}^*) \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right) \end{aligned}$$

for a variational martingale M; for $\tau = \sigma$

$$\begin{aligned} M_\sigma &= M_\tau = \mathbb{E}^{\mathbb{P}^*}[M_\tau|\mathcal{F}_\sigma] \\ &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right) \end{aligned}$$

as $\alpha_\sigma^{\min}$ is grounded and $M_\tau \in \mathcal{F}_\sigma$. Hence, for $\tau \geq \sigma$

$$\begin{aligned} M_\sigma &\geq \mathbb{E}^{\mathbb{P}^*}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{P}^*)\mathbb{I}_{\{\tau>\sigma\}} \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right). \end{aligned}$$

For $(S_t)_{t\in\mathbb{N}}$ being a variational supermartingale, the conjecture then follows from the Doob decomposition, Proposition 3.3.8, and the above results for variational martingales:

$$\begin{aligned} &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_\tau - S_\sigma|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right) \\ = &\underbrace{\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right) - M_\sigma}_{\leq 0} + \underbrace{A_\sigma - A_\tau}_{\leq 0} \\ \leq &\ 0. \end{aligned}$$

Hence

$$S_\sigma \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right).$$

□

[15] As usual the empty sum is assumed to equal zero.

For the proofs of our economic results, we just need:

Corollary 3.3.10 (from Propsition 3.3.9). *Let $(S_t)_{t\in\mathbb{N}}$ be a variational supermartingale with respect to time-consistent minimal penalty $(\alpha_t^{\min})_{t\in\mathbb{N}}$. Then we have for every stopping time τ*

$$S_{\tau\wedge t} \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{\tau\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right).$$

Proof. From the first part of the proof of Proposition 3.3.9, we have

$$\begin{aligned} S_{\tau\wedge t} &\geq \mathbb{E}^{\mathbb{P}^*}[S_{\tau\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{\tau>t\}} \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{\tau\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\mathbb{I}_{\{\tau>t\}}\right) \\ &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{\tau\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right). \end{aligned}$$

The last equation follows from $(\alpha_t^{\min})_{t\in\mathbb{N}}$ assumed to be grounded: In case $\tau \leq t$ we have

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = S_\tau + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}) = S_\tau.$$

□

3.4 Main Results

We are now enabled to state and prove the main results of this article. These directly generalize the results in [Riedel, 09] to dynamic variational preferences. In the first subsection, we state the solution of the optimal stopping problem for finite time-horizons and a minimax-theorem similar to [Schied, 07] or [Riedel, 09]. The second subsection is devoted to the solution of the infinite time-horizon problem and an approximation result. The proofs directly follow the lines in [Riedel, 09]

3.4.1 Finite Horizon

Let $T < \infty$. The following result extends the fundamental Propositions VI-1-2 and VI-1-3 in [Neveu, 75] to dynamic variational preferences.

Recall the agent's problem as given by the value function in equation (3.6): At time t, find a stopping rule τ solving

$$V_t := \operatorname*{ess\,sup}_{T \geq \tau \geq t} \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right),$$

where $(\alpha_t^{\min})_{t \leq T}$ is a time-consistent dynamic minimal penalty function.

Theorem 3.4.1 (Solution to the Finite Problem)**.** (a) *The variational Snell envelope $(U_t)_{t \leq T}$ defined in equation (3.7) by*

$$U_t := \max \left\{ X_t, \operatorname*{ess\,inf}_{\mathcal{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1} | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right) \right\} \quad \textit{for } t < T$$

and $U_T = X_T$ is the smallest variational supermartingale with respect to $(\alpha_t^{\min})_{t \leq T}$ that dominates $(X_t)_{t \leq T}$.

(b) *We have $U_t = V_t$ for all $t \leq T$, i.e. the variational Snell envelope, equation (3.7), equals the problem's value function, equation (3.6).*

(c) *$\tau^* := \inf\{t \geq 0 | U_t = X_t\}$ from equation (3.8) is an optimal stopping time, i.e. solves the optimal stopping problem under dynamic variational preferences stated in Remark 3.5. Moreover, it is the smallest optimal stopping time.*

Proof. ad (a): By definition we have $U_t \geq X_t$ for all $t \leq T$ and

$$U_t \geq \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1} | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right)$$

for all $t \leq T-1$. Hence, $(U_t)_{t \leq T}$ is a variational supermartingale with respect to $(\alpha_t^{\min})_{t \leq T}$ exceeding the payoff process $(X_t)_{t \leq T}$. Let $(Z_t)_{t \leq T}$ be another such variational supermartingale with respect to $(\alpha_t^{\min})_{t \leq T}$. We show by (backward) induction that $(Z_t)_{t \leq T} \geq (U_t)_{t \leq T}$: By definition $Z_T \geq X_T = U_T$. Assuming $Z_{t+1} \geq U_{t+1}$, we achieve

$$\begin{aligned} Z_t &\geq \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[\underbrace{Z_{t+1}}_{\geq U_{t+1}} | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right) \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1} | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right). \end{aligned}$$

Thus, as by assumption $Z_t \geq X_t$, we have hence shown (a):

$$Z_t \geq \max\left\{X_t, \operatorname*{ess\,inf}_{Q\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\} = U_t.$$

ad (b): We first show "$\geq$": By Proposition 3.3.9, we have for the variational supermartingale $(U_t)_{t\leq T} \geq (X_t)_{t\leq T}$ and all $t \leq \tau \leq T$

$$\begin{aligned} U_t &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right). \end{aligned}$$

Hence, we have

$$U_t \geq \operatorname*{ess\,sup}_{t\leq\tau\leq T} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = V_t.$$

To show "$\leq$", we define the stopping rule

$$\tau_t^* := \inf\{s \geq t : U_s = X_s\}.$$

Now, fix $t \leq T$. If we can show the stopped variational supermartingale $(U_{s\wedge\tau_t^*})_{t\leq s\leq T}$ to be a variational martingale with respect to $(\alpha_s^{\min})_{t\leq s\leq T}$, we are done: Indeed, in that case, we have since $\tau_t^* \geq t$

$$\begin{aligned} U_t &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{\tau_t^*}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_t^*}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &\leq \operatorname*{ess\,sup}_{t\leq\tau\leq T} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = V_t. \end{aligned}$$

Summing up, we then have $U_t = V_t$ for all $t \leq T$.

Hence, it leaves to show the variational martingale property of the stopped variational Snell envelope $(U_{s\wedge\tau_t^*})_{t\leq s\leq T}$: Let $t \leq s \leq T$.
(i) Whenever $\tau_t^* \leq s$, we have $U_{(s+1)\wedge\tau_t^*} = U_{\tau_t^*} = U_{s\wedge\tau_t^*}$ and hence

$$\begin{aligned} &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{(s+1)\wedge\tau_t^*}|\mathcal{F}_s] + \alpha_s^{\min}(\mathbb{Q})\right) \\ =\ &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{s\wedge\tau_t^*}|\mathcal{F}_s] + \alpha_s^{\min}(\mathbb{Q})\right) \\ =\ &U_{s\wedge\tau_t^*} + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_s^{\min}(\mathbb{Q}) = U_{s\wedge\tau_t^*}. \end{aligned}$$

(ii) For $\tau_t^* > s$, we have (by (a) and the definition of τ_t^*) $U_s > X_s$ and hence

$$\begin{aligned} U_{s\wedge\tau_t^*} &= U_s = \max\left\{X_s, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{s+1}|\mathcal{F}_s] + \alpha_s^{\min}(\mathbb{Q})\right)\right\} \\ &= \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[\underbrace{U_{s+1}}_{=U_{(s+1)\wedge\tau_t^*}}|\mathcal{F}_s] + \alpha_s^{\min}(\mathbb{Q})\right). \end{aligned}$$

(i) and (ii) show the stopped variational martingale property.

ad (c): Let $t = 0$. Then

$$\tau^* := \tau_0^* := \inf\{s \geq 0 : U_s = X_s\}$$

and $(U_{s\wedge\tau^*})_{s\leq T}$ is a variational martingale with respect to $(\alpha_s^{\min})_{s\leq T}$ as already shown. We now obtain

$$\begin{aligned} & \operatorname*{ess\,sup}_{0\leq\tau\leq T} \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau] + \alpha_0^{\min}(\mathbb{Q})\right) \\ = \; & V_0 = U_0 && \text{by (b)} \\ = \; & \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{\tau^*}|\mathcal{F}_0] + \alpha_0^{\min}(\mathbb{Q})\right) && \text{by Proposition 3.3.9} \\ = \; & \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau^*}] + \alpha_0^{\min}(\mathbb{Q})\right) && \text{by definition of } \tau^*. \end{aligned}$$

Hence, τ^* is optimal and the proof of (c) is completed. Morover, any stopping time such that $\mathbb{P}_0[\tau^{**} < \tau^*] > 0$ cannot be optimal since in that case by definition of τ^* and part (b)

$$V_0 > \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau^{**}}] + \alpha_0^{\min}(\mathbb{Q})\right).$$

□

Since coherent risk is just a special case of convex risk, the example in [Riedel, 09], Appendix D, shows that time-consistency is a necessary condition for the above theorem to hold.

We now state a minimax-theorem. Technically, this allows us to interchange the "inf" and "sup" in the formulation of the problem, i.e., intuitively, it does not matter if nature chooses a worst case distribution first and then the agent maximizes, or vice versa. In the not necessarily time-consistent case, this result is achieved in [Schied, 07]. The proof therein makes use of convex conjugates. However, the result in [Schied, 07] does not constitute a constructive device for the calculation of solutions. Here, the result takes the form

$$\operatorname*{ess\,sup}_{T \geq \tau \geq t} \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right) = \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(U_t^{\mathbb{Q}} + \alpha_t^{\min}(\mathbb{Q}) \right),$$

where $U_t^{\mathbb{Q}} := \operatorname{ess\,sup}_{T \geq \tau \geq t} \mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_t]$ is the value of the expected-utility optimal stopping problem with subjective prior $\mathbb{Q}$.

Remark 3.4.2. *As we immediately see, this means that*

$$V_t = U_t = \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(U_t^{\mathbb{Q}} + \alpha_t^{\min}(\mathbb{Q}) \right).$$

Hence, we do not *have the elegant result as in [Riedel, 09] that the variational Snell envelope $(U_t)_t$ is the lower envelope of the individual Snell envelopes $(U_t^{\mathbb{Q}})$ as the penalty is not necessarily zero.*

Of course, upon revelation, payoffs are equal:

$$\operatorname*{ess\,inf}_{\mathcal{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}) \right) = X_\tau + \operatorname*{ess\,inf}_{\mathcal{Q} \in \mathcal{M}} \alpha_\tau^{\min}(\mathbb{Q}) = X_\tau.$$

Remark 3.4.3. *Again, we can show the set*

$$\left\{ U_t^{\mathbb{Q}} + \alpha_t^{\min}(\mathbb{Q}) \quad : \quad \mathbb{Q} \in \mathcal{M} \right\}$$

to be downward directed.

Indeed: Making use of Lemma 3.3 *in [Föllmer & Penner, 06], the proof works as Lemma 3.3.3 when setting* $A := \{U_t^{\mathbb{Q}_1} + \alpha_t^{\min}(\mathbb{Q}_1) < U_t^{\mathbb{Q}_2} + \alpha_t^{\min}(\mathbb{Q}_2)\} \in \mathcal{F}_t$.

Theorem 3.4.4 (Minimax-Theorem)**.** *For every t, we have the following identity:*

$$\begin{aligned} &\operatorname*{ess\,sup}_{T \geq \tau \geq t} \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right) \\ = &\operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\operatorname*{ess\,sup}_{T \geq \tau \geq t} \mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right). \end{aligned}$$

Proof. "$\leq$": This inequality is inter alia shown in [Rockafellar, 70] for general minimax-problems.

"$\geq$": By virtue of Proposition 3.3.6 there exists a $\mathbb{Q}^*$ such that we have the following chain of inequalities:

$$\begin{aligned}
& \operatorname*{ess\,sup}_{T \geq \tau \geq t} \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right) \\
= & \operatorname*{ess\,sup}_{T \geq \tau \geq t} \left(\mathbb{E}^{\mathbb{Q}^*}[X_\tau | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*) \mathbb{I}_{\{\tau > t\}} \right) \\
\geq & \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \operatorname*{ess\,sup}_{T \geq \tau \geq t} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \mathbb{I}_{\{\tau > t\}} \right) \\
= & \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \operatorname*{ess\,sup}_{T \geq \tau \geq t} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau | \mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right)
\end{aligned}$$

since $\alpha_t^{\min}$ is grounded, i.e. on $\{\tau = t\}$, we have

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

□

Remark 3.4.5. *Set $\mathbb{Q}^*$ the worst-case distribution in case of time-consistent dynamic variational preferences and $\mathbb{Q}^{**}$ the worst-case distribution for time-consistent multiple priors in $\mathcal{Q}$ assuming $\mathcal{M} = \mathcal{Q}$, i.e. the sets of distributions with finite penalty coincide. Let $(V_t)_t$ denote the value function of the optimal stopping problem under dynamic variational preferences and $(V_t^{\mathbb{Q}})_t$ the value of the optimal stopping problem with subjective prior $\mathbb{Q}$ for an expected utility maximizer. We then have*

$$\begin{aligned}
V_t &= \operatorname*{ess\,sup}_{T \geq \tau \geq t} \underbrace{\left(\mathbb{E}^{\mathbb{Q}^*}[X_\tau | \mathcal{F}_t] + \underbrace{\alpha_t^{\min}(\mathbb{Q}^*)}_{\geq 0} \right)}_{\geq \mathbb{E}^{\mathbb{Q}^*}[X_\tau | \mathcal{F}_t]} \\
&\geq V_t^{\mathbb{Q}^*}.
\end{aligned}$$

Hence, in addition to Remark 3.4.2, this inequality makes explicit that the uncertainty averse agent under variational preferences does not behave as an expected utility maximizer with respect to the worst case measure as it is the case in [Riedel, 09] under multiple priors. This fact is also elaborated in

Proposition 3.5.16 below. In particular, the optimal stopping time τ^* from the ambiguous problem, does not *coincide with the smallest optimal stopping time from the expected utility optimal stopping problem under the worst-case subjective distribution $\mathbb{Q}^*$.*

Furthermore, we see

$$V_t \geq V_t^{\mathbb{Q}^{**}}.$$

Hence, compared to the multiple priors approach, expected reward from variational preferences is at least as high due to non-triviality of the penalty if we assume $\mathcal{Q} = \mathcal{M}$. In other words, sophistication of $\alpha^{\min}$ increases minimal expected utility. Intuitively: The agent has more information on the likelihood of distributions available under variational preferences than under multiple priors and hence values the problem more. Stated in other terms more important to application in risk management: Convex risk measures assess risk in a more liberal manner than coherent risk measures given the sets of considered distributions coincide.

3.4.2 Infinite Horizon

Let $T = \infty$. Since the variational Snell envelope is only defined for $T < \infty$, the appropriate theorem for the infinite time-horizon case shows the value function to satisfy the Bellman principle.

Theorem 3.4.6 (Infinite Problem & Approximation)**.** (a) *The value process $(V_t)_{t\in\mathbb{N}}$ as defined in equation (3.6) is the smallest variational supermartingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ that dominates the payoff process $(X_t)_{t\in\mathbb{N}}$.*
(b) *The value process $(V_t)_{t\in\mathbb{N}}$ satisfies the Bellman principle, i.e.*

$$V_t = \max\left\{X_t, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\} \quad \text{for all } t \geq 0.$$

(c) *$\tau^* := \inf\{t \geq 0 | V_t = X_t\}$ is the smallest optimal stopping time.*
(d) *Let $(U_t^T)_{t\leq T}$ denote the variational Snell envelope with respect to $(\alpha_t^{\min})_{t\leq T}$ for the optimal stopping problem of $(X_t)_{t\leq T}$ truncated to finite horizon $T <$*

∞. Let $(V_t)_{t\in\mathbb{N}}$ denote the value process of the infinite problem as given in Theorem 3.4.6. Then we have $\lim_{T\to\infty} U_t^T = V_t$ for all $t \geq 0$.

Proof. ad (b): "$\geq$": By Lemma 3.4.7 below, there exists a sequence $(\tau_k)_k$ of stopping times, such that

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_k}|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q})\right) \nearrow_k V_{t+1}.$$

Hence, making use of time-consistency[16] and continuity from below, we have

$$\begin{aligned}
& \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
= \quad & \lim_{k\to\infty} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}\left[\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{P}}[X_{\tau_k}|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P})\right)\middle| \mathcal{F}_t\right]\right. \\
& \qquad \left. + \alpha_t^{\min}(\mathbb{Q})\right) \\
= \quad & \lim_{k\to\infty} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_k}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \leq V_t.
\end{aligned}$$

Furthermore, by definition of $(V_t)_{t\in\mathbb{N}}$, we have $V_t \geq X_t$ and hence

$$V_t \geq \max\left\{X_t, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\}$$

for all $t \geq 0$.

"$\leq$": Given τ, t and set $\sigma := \max\{\tau, t+1\}$. Then, by conditional cash

[16] Here, we use time-consistency directly in terms of preferences, i.e., for τ, t, we have $\pi_t(X_\tau) = \pi_t(\pi_{t+1}(X_\tau))$, or, more elaborately,

$$\begin{aligned}
& \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
= \quad & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}\left[\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q})\right)\middle| \mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q})\right).
\end{aligned}$$

invariance,

$$\begin{aligned}
& \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
= \; & X_t \mathbb{I}_{\{\tau\le t\}} + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\sigma|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \mathbb{I}_{\{\tau\ge t+1\}} \\
= \; & X_t \mathbb{I}_{\{\tau\le t\}} \\
& + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \Bigg(\mathbb{E}^{\mathbb{Q}}\left[\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{P}}[X_\sigma|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q})\right)\middle| \mathcal{F}_t\right] \\
& \qquad\qquad + \alpha_t^{\min}(\mathbb{Q})\Bigg) \mathbb{I}_{\{\tau\ge t+1\}} \\
\le \; & \max\left\{X_t, \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\},
\end{aligned}$$

as $\operatorname{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{P}}[X_\sigma|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q})\right) \le V_{t+1}$.

This shows "$\le$" since the above inequality holds for all $\tau \ge t$ and hence for the $\operatorname{ess\,sup}_{\tau\ge t}$. Hence (b) is achieved.

ad (a): By (b) we have for all t

$$V_t \ge \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \quad \text{and} \quad V_t \ge X_t.$$

Hence, $(V_t)_{t\in\mathbb{N}}$ is a variational supermartingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ and $V_t \ge X_t$. Let $(W_t)_{t\in\mathbb{N}}$ be another variational supermartingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ exceeding $(X_t)_{t\in\mathbb{N}}$. By Proposition 3.3.9 we have for all $\tau \ge t \in \mathbb{N}$

$$\begin{aligned}
W_t &\ge \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[W_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
&\ge \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right),
\end{aligned}$$

as $W_\tau \ge X_\tau$ and, hence,

$$W_t \ge \operatorname*{ess\,sup}_{\tau\ge t} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = V_t.$$

This shows (a).

ad (c): As in the proof of Theorem 3.4.1, we can show $(V_{s\wedge\tau^*})_{s\in\mathbb{N}}$ to be a variational martingale. By our continuity assumption, we hence achieve

$$\begin{aligned}
& \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{\tau^*}|\mathcal{F}_0] + \alpha_0^{\min}(\mathbb{Q})\right) \\
= \; & \lim_{s\to\infty} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{s\wedge\tau^*}|\mathcal{F}_0] + \alpha_0^{\min}(\mathbb{Q})\right) = V_0,
\end{aligned}$$

which shows the assertion in (c).

ad (d): Since $(X_t)_{t\in\mathbb{N}}$ is assumed to be bounded, $(U_t^T)_{t\leq T}$ is bounded, too. Furthermore, enlarging the set of stopping times when considering the process up to $T+1$ instead of T, we have $U_t^T \leq U_t^{T+1}$. Hence, the limit $U_t^\infty := \lim_{T\to\infty} U_t^T$ is well-defined for all t. We thus have by continuity from below

$$\begin{aligned} U_t^\infty &= \lim_{T\to\infty} \overbrace{\max\left\{X_t, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}^T|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\}}^{U_t^T} \\ &= \max\left\{X_t, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}^\infty|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\}. \end{aligned}$$

In particular, $(U_t^\infty)_{t\in\mathbb{N}}$ is a variational supermartingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ exceeding $(X_t)_{t\in\mathbb{N}}$. We now show $(V_t)_{t\in\mathbb{N}} = (U_t^\infty)_{t\in\mathbb{N}}$, where $(V_t)_{t\in\mathbb{N}}$ is the infinite horizon problem's value function: By (a) and $(U_t)_{t\in\mathbb{N}}$ being a variational supermartingale exceeding $(X_t)_{t\in\mathbb{N}}$, we have $(U_t^\infty)_{t\in\mathbb{N}} \geq (V_t)_{t\in\mathbb{N}}$. From the finite horizon problem, we have for all T and t

$$\begin{aligned} U_t^T &= \operatorname*{ess\,sup}_{t\leq\tau\leq T} \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &\leq \operatorname*{ess\,sup}_{t\leq\tau\leq\infty} \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &= V_t. \end{aligned}$$

Hence for all t it holds

$$U_t^\infty = \lim_{T\to\infty} U_t^T \leq V_t.$$

This shows $(V_t)_{t\in\mathbb{N}} = (U_t^\infty)_{t\in\mathbb{N}}$ and completes the proof. □

The last part of the foregoing theorem is particularly valuable for achieving constructive solutions for infinite models in terms of limiting solutions of truncated ones.

Lemma 3.4.7. *Let $(\alpha_t^{\min})_{t\in\mathbb{N}}$ be a time-consistent dynamic minimal penalty. For $t\in\mathbb{N}$, the set*

$$\left\{\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \middle| \tau \geq t\right\}$$

is upward directed, i.e. for any two stopping times τ_1, τ_2, there exists a stopping time, say, $\tau_3 \geq t$ such that

$$
\begin{aligned}
& \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_3}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
= \ & \max\left\{\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right); \right. \\
& \left. \qquad \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\}.
\end{aligned}
$$

Proof. Set

$$
\begin{aligned}
A \ := \ & \left\{\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right. \\
& \left. \qquad > \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\}
\end{aligned}
$$

and define the stopping time

$$
\tau_3 := \tau_1 \mathbb{I}_A + \tau_2 \mathbb{I}_{A^C}.
$$

Note that $A \in \mathcal{F}_t$.

"$\geq$": By Lemma 3.3.4, there exists $\mathbb{Q}_3 \in \mathcal{M}$ such that

$$
\begin{aligned}
& \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_3}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
= \ & \mathbb{E}^{\mathbb{Q}_3}[X_{\tau_3}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_3>t\}} \\
= \ & \mathbb{E}^{\mathbb{Q}_3}[X_{\tau_1}|\mathcal{F}_t]\mathbb{I}_A + \mathbb{E}^{\mathbb{Q}_3}[X_{\tau_2}|\mathcal{F}_t]\mathbb{I}_{A^C} \\
& + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_3>t\}\cap A} + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_3>t\}\cap A^c} \\
= \ & \left(\mathbb{E}^{\mathbb{Q}_3}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_1>t\}}\right)\mathbb{I}_A \\
& + \left(\mathbb{E}^{\mathbb{Q}_3}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_2>t\}}\right)\mathbb{I}_{A^C} \\
\geq \ & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\}\mathbb{I}_A \\
& + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\}\mathbb{I}_{A^C} \\
= \ & \max\left\{\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right); \right. \\
& \left. \qquad \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\},
\end{aligned}
$$

where the indicator function again drops as $\alpha_t^{\min}$ is assumed to be grounded. The last equality follows from the definition of A.

"$\leq$": Since

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}}[X_{\tau_3}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\\ =\ &\left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\mathbb{I}_A+\left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\mathbb{I}_{A^C},\end{aligned}$$

we have

$$\begin{aligned}&\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_3}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\\ =\ &\left[\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\right]\mathbb{I}_A\\ &+\left[\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\right]\mathbb{I}_{A^C}\\ \leq\ &\max\Bigg\{\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right);\\ &\qquad\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\Bigg\},\end{aligned}$$

as $A\cap A^C=\emptyset$ and each factor in front of the indicator function is of course smaller than or equal to the maximum of both factors. □

3.5 Examples

In this section, we consider optimal stopping problems for prominent examples of dynamic variational preferences. First, we consider stopping with dynamic multiplier preferences or, equivalently, dynamic entropic risk measures. Secondly, we apply our theory to a generalized version of average value at risk particularly paying attention to time-consistency issues.

In [Riedel, 09], several examples of optimal stopping problems with multiple priors are considered; in particular for monotone payoff processes, as e.g. American calls or puts. For those, a worst-case distribution is achieved by virtue of stochastic dominance. Then the optimal stopping rule for multiple priors is the optimal stopping rule for the expected utility problem under this worst-case distribution aligning with the intuition of an uncertainty averse agent with multiple priors behaving as an expected utility maximizer under the worst-case distribution.

However, simplicity of these examples is owed to the penalty being trivial for multiple priors. As the penalty is not trivial in case of variational preferences, we might have a trade off between stochastic dominance on the payoff process and the penalty that might increase as nature moves towards stochastically dominated distributions of the payoff process. Hence, the worst-case distribution cannot be attained any longer by stochastic dominance for the payoff process even in the monotone case but by stochastic dominance of the entire expression, the sum of expected payoff and penalty. Furthermore, we observe that correlation is introduced even in quite simple contexts.

3.5.1 Dynamic Entropic Risk Measures

As first fundamental example we consider *dynamic entropic risk measures* or, equivalently, *dynamic multiplier preferences.* Its robust representation is intuitive: the agent expects a reference distribution $\mathbb{Q} \in \mathcal{M}$ most likely and distributions further away seem to be more and more unlikely. Hence, nature shall be punished more severely the further "away" the chosen distribution from that specific $\mathbb{Q}$. *Relative entropy* turns out to be the measure of distance in the robust representation. We introduce *multiplier preferences* as in [Maccheroni et al., 06b]. [Cheridito et al, 06] and [Föllmer & Penner, 06] equivalently consider this example as *dynamic entropic risk measures.* Let again $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\leq T}, \mathbb{P}_0)$, $T \in \mathbb{N}\cup\{\infty\}$, be the underlying space and τ denote a stopping time.

Definition 3.5.1. *For $\mathbb{P} \ll \mathbb{Q}$, locally,*[17] *we define the* relative entropy *of* $\mathbb{P}$

[17] By definition of $\mathcal{M}$ this is satisfied for all distributions under consideration.

with respect to $\mathbb{Q}$ *at time* $t \geq 0$ *as*

$$H_t(\mathbb{P}|\mathbb{Q}) := \mathbb{E}^{\mathbb{P}}\left[\ln(Z_t)\right],$$

where $Z_t := \frac{d\mathbb{P}}{d\mathbb{Q}}|_{\mathcal{F}_t}$. *Furthermore, we define the* conditional relative entropy *of* $\mathbb{P}$ *with respect to* $\mathbb{Q}$ *at time* $t \geq 0$ *as*

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

Basic properties of relative entropy are stated in [Csiszar, 75]: $H_t(\mathbb{P}|\mathbb{Q}) = 0$ if and only if $\mathbb{P} = \mathbb{Q}$ on $\mathcal{F}_t$, i.e. $Z_t = 1$, and non-negative else. As we assume the distributions under consideration to be locally equivalent, the indicator function in the last equation vanishes.

We now formally introduce dynamic multiplier preferences:

Definition 3.5.2. *Let* $\theta > 0$. *We say that dynamic variational expected reward* $\pi_t^e(X_\tau)$, $t, \tau \leq T$, *is obtained by* dynamic multiplier preferences *given reference model* $\mathbb{Q}$ *or, equivalently by* dynamic entropic risk measures, *if its robust representation is of the form*[18]

$$\pi_t^e(X_\tau) = \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_t] + \theta\hat{H}_t(\mathbb{P}|\mathbb{Q})\right). \tag{3.10}$$

Remark 3.5.3. *The variational formula for relative entropy implies*

$$\pi_t^e(X_\tau) = -\theta\ln(\mathbb{E}^{\mathbb{Q}}[e^{-\frac{1}{\theta}X_\tau}|\mathcal{F}_t]).$$

Proposition 3.5.4. *Dynamic multiplier preferences with reference distribution* $\mathbb{Q} \in \mathcal{M}$ *are time-consistent: Its robust representation has minimal penalty* $\alpha_t^{\min}(\mathbb{P}) = \theta\hat{H}_t(\mathbb{P}|\mathbb{Q})$ *for* $t \leq T$, $\mathbb{P} \in \mathcal{M}$, *satisfying the no-gain condition. Hence, we have*

$$\begin{aligned}\pi_t^e(X_\tau) &= X_t\mathbb{I}_{\{\tau=t\}} \\ &\quad + \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}}\left(\int \pi_{t+1}^e(X_\tau)d\mu + \theta\hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t))\right)\mathbb{I}_{\{\tau\geq t+1\}},\end{aligned}$$

[18]This is the generalized version of the respective definition in [Maccheroni et al., 06b]. By conditional cash invariance, for $\tau \leq t$ both sides of the equation equal X_τ.

where we set $\hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t)) := \mathbb{E}^\mu[\ln(\frac{d\mu}{d\mathbb{Q}(\cdot|\mathcal{F}_t)|_{\mathcal{F}_{t+1}}})]$ *which, by abuse of notation, we write as* $\mathbb{E}^\mu[\ln(\frac{d\mu}{d\mathbb{Q}(\cdot|\mathcal{F}_t)}\Big|_{\mathcal{F}_{t+1}})]$, $\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}$.

Proof. The specific form of the penalty is shown in [Föllmer & Penner, 06], Lemma 6.2, in terms of dynamic entropic risk measures: Robust representation of these are equal to those of multiplier preferences up to a minus sign. Time-consistency is shown in [Föllmer & Penner, 06], p.92.

We now show the specific form of π_t^e: By Corollary 3.2.16, we have to show $\gamma_t(\mu) = \theta \hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t))$. For $\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}$ we recall $\gamma_t(\mu) := \operatorname{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P})$. As $\alpha_t^{\min}$ only depends on the conditional distributions given $\mathcal{F}_t$, we may write $\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}) := \alpha_t^{\min}(\mathbb{Q} \otimes_t \mu \otimes_{t+1} \mathbb{P})$ $\forall \mathbb{Q} \in \mathcal{M}$. Hence,

$$\begin{aligned}\frac{1}{\theta}\gamma_t(\mu) &= \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q} \otimes_t \mu \otimes_{t+1} \mathbb{P}) \\ &= \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P}}\left[\ln\left(\frac{\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}|_{\mathcal{F}_T}}{\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}|_{\mathcal{F}_t}}\right)\middle|\mathcal{F}_t\right].\end{aligned}$$

First, note that we have by $d\mu = d(\mathbb{Q} \otimes_t \mu)(\cdot|\mathcal{F}_t)$

$$\mathbb{E}^\mu\left[\ln\left(\frac{d\mu}{d\mathbb{Q}(\cdot|\mathcal{F}_t)}\bigg|_{\mathcal{F}_{t+1}}\right)\right] = \mathbb{E}^{\mathbb{Q}\otimes_t\mu}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu)}{d\mathbb{Q}}\bigg|_{\mathcal{F}_{t+1}}\right)\middle|\mathcal{F}_t\right].$$

As the integrand is $\mathcal{F}_{t+1}$-measurable and $\frac{d(\mathbb{Q}\otimes_t\mu)}{d\mathbb{Q}}\Big|_{\mathcal{F}_{t+1}} = \frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}\Big|_{\mathcal{F}_{t+1}}$, the following equation holds for all $\mathbb{P} \in \mathcal{M}$:

$$\begin{aligned}&\mathbb{E}^\mu\left[\ln\left(\frac{d\mu}{d\mathbb{Q}(\cdot|\mathcal{F}_t)}\bigg|_{\mathcal{F}_{t+1}}\right)\right] \\ = \;&\mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P}}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}\bigg|_{\mathcal{F}_{t+1}}\right)\middle|\mathcal{F}_t\right].\end{aligned}$$

Hence, it leaves to show for all $\mathbb{R} \in \mathcal{M}$ that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{R}}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{R})}{d\mathbb{Q}}\bigg|_{\mathcal{F}_{t+1}}\right)\bigg|\mathcal{F}_t\right] \\
= \;& \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P}}\left[\ln\left(\frac{\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}|_{\mathcal{F}_T}}{\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}|_{\mathcal{F}_t}}\right)\bigg|\mathcal{F}_t\right] \\
= \;& \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P}}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}\bigg|_{\mathcal{F}_T}\right)\bigg|\mathcal{F}_t\right],
\end{aligned}$$

where the last equation follows as $\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}|_{\mathcal{F}_t} = 1$.

We know from the properties of the entropy, that $\hat{H}_t(\mathbb{P}|\mathbb{Q}) \geq 0$ and $= 0$ if and only if $\mathbb{P} = \mathbb{Q}$ on $\mathcal{F}_t$. In the same way, we have that

$$\mathbb{Q} \in \arg\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P}}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}\bigg|_{\mathcal{F}_T}\right)\bigg|\mathcal{F}_t\right].$$

More precisely,

$$\begin{aligned}
&\arg\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P}}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}\bigg|_{\mathcal{F}_T}\right)\bigg|\mathcal{F}_t\right] \\
= \;& \{\mathbb{V} \in \mathcal{M} | \mathbb{V} = \mathbb{R}\otimes_t\mu\otimes_{t+1}\mathbb{Q} \quad \text{for some } \mathbb{R} \in \mathcal{M}\}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P}}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{P})}{d\mathbb{Q}}\bigg|_{\mathcal{F}_T}\right)\bigg|\mathcal{F}_t\right] \\
= \;& \mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{Q}}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{Q})}{d\mathbb{Q}}\bigg|_{\mathcal{F}_T}\right)\bigg|\mathcal{F}_t\right] \\
= \;& \mathbb{E}^{\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{Q}}\left[\ln\left(\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{Q})}{d\mathbb{Q}}\bigg|_{\mathcal{F}_{t+1}}\right)\bigg|\mathcal{F}_t\right],
\end{aligned}$$

where the second equality follows since $q_t := \frac{d\mathbb{Q}}{d\mathbb{Q}}|_{\mathcal{F}_t} = 1$ for all $t \leq T$ and hence $\frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{Q})}{d\mathbb{Q}}\Big|_{\mathcal{F}_{t+1}} = \frac{d(\mathbb{Q}\otimes_t\mu\otimes_{t+1}\mathbb{Q})}{d\mathbb{Q}}\Big|_{\mathcal{F}_\eta}$ for all $\eta \geq t+1$. This completes the proof. $\square$

For the value function, we thus have

$$
\begin{aligned}
V_t &= \operatorname*{ess\,sup}_{t\le\tau\le T} \pi_t^e(X_\tau) \\
&= \operatorname*{ess\,sup}_{t\le\tau\le T} \Big\{ X_t \mathbb{I}_{\{\tau=t\}} \\
&\quad + \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int \pi_{t+1}^e(X_\tau)d\mu + \theta \hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t)) \right) \mathbb{I}_{\{\tau\ge t+1\}} \Big\} \\
&= \max \Big\{ X_t; \\
&\quad \operatorname*{ess\,sup}_{t+1\le\tau\le T} \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int \pi_{t+1}^e(X_\tau)d\mu + \theta \hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t)) \right) \Big\} \\
&= \max \Big\{ X_t; \\
&\quad \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int \operatorname*{ess\,sup}_{t+1\le\tau\le T} \pi_{t+1}^e(X_\tau)d\mu + \theta \hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t)) \right) \Big\} \\
&= \max \Big\{ X_t; \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right) \Big\}
\end{aligned}
$$

again showing the Bellman principle to hold but having applied our minimax theorem. As we want to achieve explicit solutions, we further confine ourselves:

Assumption 3.5.5. *Let the underlying probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\le T}, \mathbb{P}_0)$ *be given as the independent product of the time-t state space,* $(S, \mathcal{S}, \nu_0)$*,* $S \subset \mathbb{R}$*. Then* $\mathbb{P}_0 = \otimes_{t=1}^T \nu_o$ *and* $\mathcal{F}_s$ *is generated by the projection mappings* $\epsilon_t : \Omega \mapsto S$*,* $t \le s$*. In particular, the* $\epsilon_t s$ *are i.i.d. with* ν_0 *under* $\mathbb{P}_0$*.*

As in [Riedel, 09], we confine ourselves to the set

$$
\mathcal{M}^{[a,b]} := \left\{ \mathbb{P}^\beta \approx \mathbb{P}_0 \quad : \quad \frac{d\mathbb{P}^\beta}{d\mathbb{P}_0}\bigg|_{\mathcal{F}_t} = D_t^\beta \quad \forall t,\ (\beta_t)_t \subset [a,b],\ \textit{predictable} \right\},
$$

$D_t^\beta := \exp(\sum_{s=1}^t \beta_s \epsilon_s - \sum_{s=1}^t L(\beta_s))$ *for some predictable process* $(\beta_t)_{t\le T} \subset [a,b] \subset \mathbb{R}$ *and* $L(\beta_t) := \ln \int_S e^{\beta_t x} \nu_0(dx)$*.*

Remark 3.5.6. *As we have now constrained the set of possible probability distributions, we note that we are not in context of general dynamic entropic risk measures any longer.*

Notation 3.5.7. *The reference distribution of the entropic penalty write as* $\mathbb{Q} := \mathbb{P}^{\beta^1}$*, i.e.* $(\beta^1_t)_{t\leq T}$ *denotes the process defining the penalty's reference distribution. Note that* $\mathbb{Q}$ *is in general not equal to* $\mathbb{P}_0$*. Other distributions in* $\mathcal{M}$ *write as* $\mathbb{P} := \mathbb{P}^{\beta^2}$*.*

Then

$$\frac{d\mathbb{P}}{d\mathbb{Q}}\bigg|_{\mathcal{F}_t} = \frac{D_t^{\beta^2}}{D_t^{\beta^1}}\frac{d\mathbb{P}_0}{d\mathbb{P}_0}\bigg|_{\mathcal{F}_t} = \exp\left(\sum_{s=1}^{t}(\beta^2_s - \beta^1_s)\epsilon_s - \sum_{s=1}^{t}[L(\beta^2_s) - L(\beta^1_s)]\right).$$

and the entropic penalty with reference distribution $\mathbb{Q}$ is given by

$$\begin{aligned}\alpha_t^{\min}(\mathbb{P}) &= \theta \hat{H}_t(\mathbb{P}|\mathbb{Q}) \\ &= \theta\mathbb{E}^{\mathbb{P}}\left[\sum_{s=t+1}^{T}(\beta^2_s - \beta^1_s)\epsilon_s - \sum_{s=t+1}^{T}[L(\beta^2_s) - L(\beta^1_s)]\bigg|\mathcal{F}_t\right].\end{aligned}$$

We write $\mathbb{E}^\beta := \mathbb{E}^{\mathbb{P}^\beta}$ and $\hat{H}_t(\beta^2|\beta^1) := \hat{H}_t(\mathbb{P}^{\beta^2}|\mathbb{P}^{\beta^1})$ as well as $\alpha_t^{\min}(\beta^2)$. Note, in case $\mathbb{Q} = \mathbb{P}_0$, we have $(\beta^1_t)_{t\leq T} = 0$ and hence for $\mathbb{P} = \mathbb{P}^{\beta^2}$: $\alpha_t^{\min}(\mathbb{P}) = \theta\mathbb{E}^{\mathbb{P}}\left[\sum_{s=t+1}^{T}\beta^2_s\epsilon_s - \sum_{s=t+1}^{T}L(\beta^2_s)\Big|\mathcal{F}_t\right]$.

To make the value function $(V_t)_{t\leq T}$ more explicit, note for $\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}$ given by previsible $(\beta^2_t)_{t\leq T}$ and penalty's reference distribution $\mathbb{Q} \in \mathcal{M}$ by previsible $(\beta^1_t)_{t\leq T}$, we have

$$\begin{aligned}\hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t)) &= \mathbb{E}^\mu\left[\ln\left(\frac{d\mu}{d\mathbb{Q}(\cdot|\mathcal{F}_t)|_{\mathcal{F}_{t+1}}}\right)\right] \\ &= \mathbb{E}^{\beta^2_{t+1}}\left[(\beta^2_{t+1} - \beta^1_{t+1})\epsilon_{t+1} - (L(\beta^2_{t+1}) - L(\beta^1_{t+1}))\right].\end{aligned}$$

Hence, as above the value is given by

$$\begin{aligned}V_t &= \operatorname*{ess\,sup}_{t\leq\tau\leq T}\operatorname*{ess\,inf}_{\beta^2\subset[a,b]}\left(\mathbb{E}^{\beta^2}[X_\tau|\mathcal{F}_t] + \theta\hat{H}_t(\beta^2|\beta^1)\right) \qquad (3.11) \\ &= \operatorname*{ess\,sup}_{t\leq\tau\leq T}\operatorname*{ess\,inf}_{\beta^2\subset[a,b]}\mathbb{E}^{\beta^2}\left[X_\tau + \theta\left(\sum_{s=t+1}^{T}(\beta^2_s - \beta^1_s)\epsilon_s\right.\right. \\ &\qquad \left.\left. - \sum_{s=t+1}^{T}[L(\beta^2_s) - L(\beta^1_s)]\right)\bigg|\mathcal{F}_t\right] \\ &= \max\Big\{X_t;\end{aligned}$$

$$
\begin{aligned}
&\operatorname*{ess\,sup}_{t+1\le\tau\le T} \operatorname*{ess\,inf}_{\beta^2_{t+1}\in[a,b]} \mathbb{E}^{\beta^2_{t+1}} \Big[\pi_{t+1}(X_\tau) + \theta\left((\beta^2_{t+1}-\beta^1_{t+1})\epsilon_{t+1}\right.\\
&\qquad\qquad \left.-(L(\beta^2_{t+1})-L(\beta^1_{t+1}))\right)\Big]\Big\}\\
= &\max\Big\{X_t \quad ; \quad \operatorname*{ess\,inf}_{\beta^2_{t+1}\in[a,b]} \mathbb{E}^{\beta^2_{t+1}} \Big[V_{t+1} + \theta\left((\beta^2_{t+1}-\beta^1_{t+1})\epsilon_{t+1}\right.\\
&\qquad\qquad \left.-(L(\beta^2_{t+1})-L(\beta^1_{t+1}))\right)\Big]\Big\},
\end{aligned}
$$

where the last equality follows from the Minimax result. In particular, we see that the value of the problem – and hence the worst case distribution – depends on the reference distribution $\mathbb{Q} = \mathbb{P}^{\beta^1}$ of the penalty. In case $T < \infty$, the same recursion has to hold for the Snell envelope $(U_t)_{t\le N}$ by Theorem 3.4.1:

$$
\begin{aligned}
U_t &= \max\{X_t; \pi_t(U_{t+1})\}\\
&= \max\left\{X_t; \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int \pi_{t+1}(U_{t+1})d\mu + \theta H_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t))\right)\right\}\\
&= \max\left\{X_t; \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int U_{t+1}d\mu + \theta H_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t))\right)\right\}\\
&= \max\Big\{X_t; \operatorname*{ess\,inf}_{\beta^2_{t+1}\in[a,b]} \mathbb{E}^{\beta^2_{t+1}} \Big[U_{t+1} + \theta\left((\beta^2_{t+1}-\beta^1_{t+1})\epsilon_{t+1}\right.\\
&\qquad\qquad \left.-(L(\beta^2_{t+1})-L(\beta^1_{t+1}))\right)\Big]\Big\}.
\end{aligned}
$$

To further solve problems under entropic risk, we have to make specific properties of the payoff process explicit. We constraint ourselves to monotone problems:

Assumption 3.5.8. *$X_t := f(t,\epsilon_t)$, where f is a bounded measurable function that is strictly monotone in the state variable ϵ_t.*

For monotone payoff processes in the ambiguous, i.e. multiple priors, case it is shown in [Riedel, 09] that U_t is increasing in ϵ_t. However, having a look at the proof therein (Appendix F), we see that this crucially depends on ϵ_t being independent of $\mathcal{F}_{t-1}$ (cf. equation (12) in [Riedel, 09]) as the process

$(\beta_t^2)_t$ yielding the worst case distribution under multiple priors is constant, and the worst case distribution being the one that is stochastically dominated for the payoff process (Lemma 13). We will see that these arguments do not have to hold in case of variational preferences. Furthermore, in [Riedel, 09]'s multiple priors case, the calculation of a worst case measure is done by virtue of stochastic dominance on the payoff process. It is intuitive that this cannot work as elegant under variational preferences: The penalty is not trivial, i.e. not zero on some set of priors and infinite else. In particular, in the entropic case, the worst-case measure depends on the reference distribution $\mathbb{Q}$: there might be a trade off between stochastic dominance on $(X_t)_t$ and the penalty: The penalty increases the further nature moves away from $\mathbb{Q}$ and in direction of a distribution minimizing the expectation of the payoff process.

To gain insights, we have a look at a special case for the reference distribution of the penalty:

Example 3.5.9. *Let f be increasing and the reference distribution be $\mathbb{Q} = \mathbb{P}^a$, the distribution given by $\beta_t^1 = a$ for all $t \leq T$. We encounter for the first term in the value function, $\mathbb{E}^{\beta^2}[f(\tau, \epsilon_\tau)|\mathcal{F}_t]$: $\mathbb{P}^a$ is stochastically dominated, i.e. minimizes that term on $\mathcal{M}^{[a,b]}$. $\mathbb{P}^a$ also minimizes the penalty: $\hat{H}_t(\beta^2|a) := \hat{H}_t(\mathbb{P}^{\beta^2}|\mathbb{P}^a)$ is increasing in β^2 on $[a, b]$, $\hat{H}_t \geq 0$ and zero if and only if $\mathbb{P}^{\beta^2} = \mathbb{P}^a$. Hence we have equivalence of the problem under dynamic multiplier preferences and the optimality problem under the worst case distribution $\mathbb{P}^a$ as in Theorem 5 in [Riedel, 09].*

Proposition 3.5.10. *Let f be increasing, $T < \infty$, and τ^a denote the optimal stopping time for the classical optimal stopping problem of $(X_t)_{t\leq T}$ under subjective distribution $\mathbb{P}^a$, i.e. τ^a solves $\max_{0\leq\tau\leq T} \mathbb{E}^a[X_\tau]$. Let $\mathbb{Q} = \mathbb{P}^a$ be the reference measure for the penalty, i.e. $\beta_t^1 = a$, $t \leq T$, in equation (3.11). Then, τ^a is the solution to the robust problem with dynamic multiplier preferences $(\pi_t^e)_{t\leq T}$ as given in equation (3.11).*

Proof. Intuitively, in Appendix F in [Riedel, 09], it is shown that $\mathbb{P}^a$ is the worst case distribution for the first term in the value function (3.11). As

$\hat{H}_t(a|a) = 0 \leq \hat{H}_t(\beta^2|a)$ for all β^2, $\mathbb{P}^a$ also minimizes the penalty and hence is the worst case distribution in the multiplier case when $\mathbb{Q} = \mathbb{P}^a$.

Formally: For all increasing bounded measurable functions $h : \Omega \to \mathbb{R}$ and all $t \geq 1$, we have by Lemma 13 in [Riedel, 09]

$$\begin{aligned} \mathbb{E}^a[h(\epsilon_t)|\mathcal{F}_{t-1}] &= \operatorname*{ess\,inf}_{\beta^2\in[a,b]} \mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}] \\ &= \operatorname*{ess\,inf}_{\beta^2[a,b]} \mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}] + \underbrace{\min_{\beta^2\in[a,b]} \theta \hat{H}_{t-1}(\beta^2|a)}_{=\hat{H}_t(a|a)=0} \\ &= \operatorname*{ess\,inf}_{\beta^2\in[a,b]} \left(\mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}] + \theta \hat{H}_{t-1}(\beta^2|a)\right), \end{aligned}$$

where the last equation follows as the joint minimizer of both summands is $\mathbb{P}^a$. Given this result, we can mimic the proof of Theorem 5 in [Riedel, 09]: Let $(U_t)_{t\leq T}$ denote the variational Snell envelope of the problem under multiplier preferences and $(U_t^a)_{t\leq T}$ the classical Snell envelope with respect to subjective prior $\mathbb{P}^a$. For $t = T$, we have $U_T = X_T = f(T, \epsilon_T) = U_T^a$ and hence increasing in ϵ_T. As by induction hypothesis U_{t+1} is an increasing function of ϵ_{t+1}, say $U_{t+1} = u(\epsilon_{t+1})$ for some bounded measurable increasing u, we have for all $t < T$

$$\begin{aligned} U_t &= \max\left\{ f(t,\epsilon_t), \quad \operatorname*{ess\,inf}_{\beta^2\in\mathcal{M}^{[a,b]}} \left(\mathbb{E}^{\beta^2}[U_{t+1}|\mathcal{F}_t] + \theta \hat{H}_t(\beta^2|a)\right)\right\} \\ &= \max\left\{ f(t,\epsilon_t), \quad \mathbb{E}^a[U_{t+1}|\mathcal{F}_t] + \underbrace{\theta \hat{H}_t(a|a)}_{=0} \right\} \\ &= \max\left\{ f(t,\epsilon_t), \quad \mathbb{E}^a[U_{t+1}|\mathcal{F}_t]\right\} =: U_t^a. \end{aligned}$$

This shows the assertion by equality of the recursion formulas: $(U_t)_{t\leq T} = (U_t^a)_{t\leq T}$ and hence the optimal stopping times coincide. □

Remark 3.5.11. *The foregoing proof particularly shows that U_t is increasing in ϵ_t in case $\mathbb{Q} = \mathbb{P}^a$: ϵ_{t+1} is independent of $\mathcal{F}_t$ under $\mathbb{P}^a$ and hence*

$$\begin{aligned} U_t &= \max\{f(t,\epsilon_t), \mathbb{E}^a[u(\epsilon_{t+1})|\mathcal{F}_t]\} \\ &= \max\{f(t,\epsilon_t), \mathbb{E}^a[u(\epsilon_{t+1})]\}. \end{aligned}$$

The argument in the foregoing proof for the case $\mathbb{Q} = \mathbb{P}^a$ is that $\mathbb{P}^a$ minimizes $\mathbb{E}^{\mathbb{P}}[f(t, \epsilon_t)]$ as well as $\hat{H}_t(\mathbb{P}|a)$. Of course, this does not hold true if the reference measure $\mathbb{Q} = \mathbb{P}^{\beta^1}$ is such that β^1_t is not identical a. Then, we have a trade off between a decrease in the first term, $\mathbb{E}^{\mathbb{P}}[f(t, \epsilon_t)]$, which is independent of $\mathbb{P}^{\beta^1}$, and an increase of the penalty in the second term, $\hat{H}_t(\mathbb{P}|\beta^1)$, the further nature deviates from the reference distribution $\mathbb{P}^{\beta^1}$ "downwards" to the distribution $\mathbb{P}^a$. More elaborately, considering a distribution $\mathbb{P}^{\beta^2}$ with $\beta^2_t \in [a, \beta^1_t]$, $t \leq T$: When nature moves towards $\mathbb{P}^a$, decreasing the first term, the second term increases; when nature moves towards the reference distribuiton $\mathbb{P}^{\beta^1}$, minimizing the second term, the first term increases. However, moving from $\mathbb{P}^{\beta^1}$ in direction of the upper extremal distribution $\mathbb{P}^b$, both terms increase:

Proposition 3.5.12. *Let $\mathbb{Q} = \mathbb{P}^{\beta^1} \in \mathcal{M}^{[a,b]}$ be the reference distribution of the entropic penalty, and f be increasing. Then, the worst case distribution $\mathbb{P}^{\bar{\beta}^2}$ satisfies $\bar{\beta}^2_t \in [a, \beta^1_t]$.*

Proof. For h as above, we have

$$\begin{aligned} & \operatorname*{ess\,inf}_{\beta \in [a,b]} \left\{ \mathbb{E}^{\beta}[h(\epsilon_t)|\mathcal{F}_{t-1}] + \hat{H}_{t-1}(\beta|\beta^1) \right\} \\ \leq \quad & \mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}] + \hat{H}_{t-1}(\beta^2|\beta^1) \end{aligned}$$

for all $\beta^2_t \in [\beta^1_t, b]$ for all t as $\hat{H}_{t-1}(\beta^1|\beta^1) = 0$ and ≥ 0 else and furthermore $\mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}]$ is increasing in β^2 as seen in the proof of Lemma 13 in [Riedel, 09]. As $\hat{H}_t(\cdot|\beta^1)$ is strictly increasing on $[\beta^1_t, b]$, we have strict inequality on $]\beta^1_t, b]$. □

We see, that the approaches e.g. in [Karatzas & Zamfirescu, 08], with nature maximizing over the set of priors, are easier to handle in this context as there is no tradeoff.

Example 3.5.13. *The second extreme case for monotone increasing problems to be considered is the penalty's reference distribution set to $\mathbb{Q} = \mathbb{P}^b$:*

Here, the smaller $(\beta_t^2)_t$ *is chosen and hence the smaller the first term, the more increases the penalty as nature deviates further from the reference distribution. In particular, we see that the worst case distribution depends on the specific form of* f*, not just on* f *being increasing: Due to tradeoff, it depends on the slope of* f *at a particular state of the world. This has severe consequences for the complexity of calculations: Let us for example take the case of an American call as considered in [Riedel, 09]. As long as it is in the money, the slope of* f *is one, whereas it is zero when out of the money. I.e., when out of the money, nature cannot just apply a distribution low enough to likely staying out of the money but also has to take care of it being close enough to* $\mathbb{Q}$ *not to increase the penalty too much. In this sense, the penalty comes relatively more severely into account when the call is out of the money and, hence, the one step ahead worst case distribution depends on the current state:*

Remark 3.5.14. *In case of variational preferences, correlation is already introduced for the call that has independent rewards under multiple priors as shown in [Riedel, 09].*

In general, we see that an increase in penalty by deviating further from $\mathbb{P}^{\beta^1}$ to $\mathbb{P}^a$ is less severe the steeper f, i.e. the tradeoff effect is in favor of minimizing the first part of the value function, the expectation. In extreme cases we might even still have $\mathbb{P}^a$ to be the worst case distribution if f is "steep enough", i.e. the increase in penalty might be outweighed by the decrease in expected f, and $\mathbb{P}^{\beta^1}$ "is not too far away" from $\mathbb{P}^a$. To sum up:

Proposition 3.5.15. *As we have already seen, the worst case distribution depends on the reference distribution* $\mathbb{Q}$ *of the penalty, i.e. on* $(\beta_t^1)_{t\leq T}$*. Furthermore, as we have argued, it is a function of the current state of the world and the specific form of the function* f *at that state and particularly of the whole history.*

It is hence immediate that not even a constant reference process $(\beta_t^1)_{t\leq T}$ induces a constant worst case $(\bar{\beta}_t^2)_{t\leq T}$. This insight can be seen in the following calculations: Let $U_t = h(\epsilon_1, \dots, \epsilon_t)$, bounded and $\mathcal{F}_t$-measurable. Then, the right hand side of the Snell envelope becomes

$$\begin{aligned} & \mathbb{E}^{\beta_t^2}[h(\epsilon_1, \dots, \epsilon_t)|\mathcal{F}_{t-1}] + \theta \hat{H}_t(\beta_1^2|\mathbb{P}^{\beta^1}(\cdot|\mathcal{F}_{t-1})|_{\mathcal{F}_t}) \\ = \;& \mathbb{E}^{\beta_t^2}[h(\epsilon_1, \dots, \epsilon_t) + \theta\left((\beta_t^2 - \beta_t^1)\epsilon_t - (L(\beta_t^2) - L(\beta_t^1))\right)|\mathcal{F}_{t-1}]. \end{aligned}$$

In order to recursively obtain a worst-case distribution, we have to minimize this expression with respect to $\beta_t^2 \in [a, b]$ and obtain some $\bar{\beta}_t^2 = \bar{\beta}_t^2(\epsilon_1, \dots, \epsilon_{t-1}, \beta_t^1)$. In particular, we can see that the process achieving the worst-case distribution is again previsible, i.e. $\bar{\beta}_t^2$ is $\mathcal{F}_{t-1}$-measurable. Hence, given a specific structure of $(X_t)_{t\leq T}$ and a reference $\mathbb{P}^{\beta^1}$ for the penalty, we receive a worst case measure $\mathbb{P}^{\bar{\beta}^2}$ where $(\bar{\beta}_t^2)_t$ is achieved as above. Having achieved this worst case distribution, we can calculate the optimal stopping time τ^*. However, as in general $\hat{H}_t(\bar{\beta}_t^2|\beta_t^1) \neq 0$, we obtain a *negation* of Theorem 5 in [Riedel, 09] for our approach:

Proposition 3.5.16. *Let $(\bar{\beta}_t^2)_t$ denote the process inducing the worst-case distribution for the monotone problem under dynamic multiplier preferences $(\pi_t^e)_{t\leq T}$. Then,*

$$\begin{aligned} U_t &= \max\left\{X_t; \operatorname*{ess\,inf}_{\beta_{t+1}^2\in[a,b]} \left(\mathbb{E}^{\beta_{t+1}^2}[U_{t+1}|\mathcal{F}_t] + \theta H_{t+1}(\beta_{t+1}^2|\mathbb{P}^{\beta^1}(\cdot|\mathcal{F}_t))\right)\right\} \\ &= \max\left\{X_t; \mathbb{E}^{\bar{\beta}_{t+1}^2}[U_{t+1}|\mathcal{F}_t] + \theta H_{t+1}(\bar{\beta}_{t+1}^2|\mathbb{P}^{\beta^1}(\cdot|\mathcal{F}_t))\right\} \\ &\geq \max\left\{X_t; \mathbb{E}^{\bar{\beta}_{t+1}^2}[U_{t+1}|\mathcal{F}_t]\right\} = U_t^{\bar{\beta}^2}, \end{aligned}$$

where $U_t^{\bar{\beta}^2}$ denotes the classical Snell envelope of the optimal stopping problem under subjective prior given by $\bar{\beta}^2$. In particular, we see that

$$\tau^* = \inf_t\{X_t = U_t\} \geq \inf_t\{X_t = U_t^{\bar{\beta}^2}\} = \tau^{\bar{\beta}^2 *}.$$

As the recursion formulas for the Snell Envelopes and hence the optimal stopping times of the problem under dynamic multiplier preferences and the one for an expected utility maximizer under the respective worst case distribution *differ, we see that the intuition in [Riedel, 09] is not valid anymore:* The agent does not behave as the expected utility maximizer under the worst case distribution.

As a tangible example, we apply the problem of an American put to variational preferences. We assume the agent to consider the market as "emerging", i.e. she considers distributions more favorable under which the value of the underlying is likely to go up. We hence set the reference distribution of the entropic penalty to $\mathbb{P}^b$. We will formally show the following result: As the value of the American put is decreasing in the value of the underlying and the penalty is minimal for $\mathbb{P}^b$, the worst case distribution is given by $\mathbb{P}^b$. Moreover, as $\hat{H}_t(\mathbb{P}^b|\mathbb{P}^b) = 0$ for all t, the agent behaves as expected utility maximizer under the subjective prior $\mathbb{P}^b$. Formally:

Example 3.5.17 (American Options in CRR-Model)**.** *Let the agent assess utility in terms of dynamic multiplier preferences with entropic penalty given by parameter $\theta = 1$ and reference distribution $\mathbb{P}^b$. We consider American options for the Cox-Ross-Rubinstein (CRR) model: Let $\Omega := \{0,1\}^T$, $T < \infty$.*[19] *Let $\epsilon_t : \Omega \to \{0,1\}$, $t \leq T$, be the projection mappings and $\mathbb{P}_0$ such that ϵ_t's are i.i.d. under $\mathbb{P}_0$ with $\mathbb{P}_0[\epsilon_t = 1] = \mathbb{P}_0[\epsilon_t = 0] = \frac{1}{2}$. Let $\mathcal{M}^{[a,b]}$ be given as in Assumption 3.5.5. As in [Riedel, 09], we then have for all $\beta := (\beta_t)_t$ that $\mathbb{P}^\beta[\epsilon_t = 1|\mathcal{F}_{t-1}] \in [\underline{p}; \bar{p}]$, where $\underline{p} := \frac{e^a}{1+e^a}$ and $\bar{p} := \frac{e^b}{1+e^b}$. Let $\mathbb{P}^a$ be again the distribution induced by the constant process with $\beta_t = a$ for all t and equivalently for $\mathbb{P}^b$. Then, under $\mathbb{P}^a$, ϵ_t's are i.i.d. with $\mathbb{P}^a[\epsilon_t] = \underline{p}$ and equivalently for $\mathbb{P}^b$ with $\mathbb{P}^b[\epsilon_t] = \bar{p}$.*

The "ingredients" of the CRR-model are given by a risk-less asset with value process $B_t = (1+r)^t$ for some fixed interest rate $r > -1$ and a risky asset with value process S_t at t such that $S_0 = 1$ and

$$S_{t+1} = S_t \cdot \begin{cases} (1+d) & \text{if } \epsilon_{t+1} = 1, \\ (1+c) & \text{if } \epsilon_{t+1} = 0, \end{cases}$$

[19] The infinite case can be achieved by virtue of Theorem 3.4.6.

where we assume the constants to satisfy $-1 < c < r < d$ for the model not to allow for arbitrage opportunities.

Now, consider an American option with payoff $A(t, S_t)$ from exercising at t. The agent has to solve the problem[20]

$$\operatorname*{ess\,sup}_{\tau} \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}^{[a,b]}} \left\{ \mathbb{E}^{\mathbb{P}}\left[A(\tau, S_\tau)\right] + H_0(\mathbb{P}|\mathbb{P}^b) \right\}.$$

To further elaborate the example: Assume $A^p(t, S_t)$ being an American put and, hence, decreasing in S_t for all t.[21] *Let $(U_t^b)_{t\leq T}$ denote the classical Snell envelope of $A^p(t, S_t)$ under subjective probability $\mathbb{P}^b$, i.e.*

$$\begin{aligned} U_t^b(t, S_t) &= \max\left\{ A^p(t, S_t); \bar{p} U_t^b(t+1, S_t(1+d)) \right. \\ &\qquad \left. +(1-\bar{p})U_t^b(t+1, S_t(1+c)) \right\}. \end{aligned}$$

The following assertion holds: The variational Snell envelope $(U_t)_{t\leq T}$ of the American put problem with dynamic multiplier preferences and reference distribution $\mathbb{P}^b$ satisfies $(U_t)_{t\leq T} = (U_t^b)_{t\leq T}$. In particular, the worst case distribution is given by $\mathbb{P}^b$ and, as the penalty vanishes for this distribution, the optimal stopping time is given by $\tau^* = \inf\{t \geq 0 | A^p(t, S_t) = U_t^b\} = \tau^{b*}$, i.e. the optimal stopping time τ^{b*} of the problem under subjective prior $\mathbb{P}^b$.

The proof of this assertion is immediate by virtue of stochastic dominance: As in Appendix H in [Riedel, 09], we show for the variational Snell envelope $(U_t)_{t\leq T}$ that $U_t = u(t, S_t) = U_t^b$, $t \leq T$, for a function u that is decreasing in the second variable: First, we have $U_T = A^p(T, S_T) = U_T^b$ by definition. For an inductive proof, we write with a slight but intuitively understandable misuse of notation $\hat{H}_t(p_{t+1} \otimes p_{t+2} \otimes \ldots |\mathbb{P}^b)$[22] *for $p_i \in [\underline{p}; \bar{p}]$ and note that $\hat{H}_t(\bar{p} \otimes \bar{p} \otimes \ldots |\mathbb{P}^b) = 0$ and ≥ 0 else, i.e. $\bar{p}$ at any t minimizes the penalty. From the induction hypothesis, we have $u(t+1, S_t(1+d)) \leq u(t+1, S_t(1+c))$.*

[20] [Riedel, 09] achieves a general theory for American options under multiple priors.

[21] Equivalent results hold for an American call with $\mathbb{P}^a$ as reference distribution.

[22] Formally: $\hat{H}_t(p_{t+1} \otimes p_{t+2} \otimes \ldots |\mathbb{P}^b) := \hat{H}_t(\mathbb{P}^\beta|\mathbb{P}^b)$ with $(\beta_t)_{t\leq T}$ such that $\mathbb{P}^\beta[\epsilon_t = 1|\mathcal{F}_{t-1}] = p_t$ for $t \leq T$; well defined as $p_1, \ldots, p_t$ drops by general definition of $\hat{H}_t$.

We hence have

$$
\begin{aligned}
U_t &= \max\Big\{A^p(t,S_t) \quad ; \quad \min_{p_{t+1}\in[\underline{p};\bar{p}]} \{p_{t+1}u(t+1,S_t(1+d)) \\
&\qquad\qquad\qquad\qquad\qquad +(1-p_{t+1})u(t+1,S_t(1+c)) \\
&\qquad\qquad\qquad\qquad\qquad +H_t(p_{t+1}\otimes\bar{p}\otimes\ldots|\mathbb{P}^b)\}\Big\} \\
&= \max\Big\{A^p(t,S_t) \quad ; \quad \bar{p}u(t+1,S_t(1+d)) \\
&\qquad\qquad\qquad +(1-\bar{p})u(t+1,S_t(1+c)) \\
&\qquad\qquad\qquad +\underbrace{H_t(\bar{p}\otimes\bar{p}\otimes\ldots|\mathbb{P}^b)}_{=0}\Big\} \\
&= U_t^b.
\end{aligned}
$$

Thus, we have the equality of the variational Snell envelope and the classical Snell envelope under the worst case measure, i.e. $(U_t)_{t\leq T} = (U_t^b)_{t\leq T}$*, and the coincidence of the respective optimal stopping times, i.e.* $\tau^* = \tau^{b*}$*.*

To conclude: The problem of optimally exercising an American put under dynamic entropic risk with reference distribution $\mathbb{P}^b$ *for the entropic penalty coincides with the problem for the American put for an expected utility maximizer with respect to subjective prior* $\mathbb{P}^b$*.*

In a way, the result in the example is more like a self fulfilling prophecy as the agent assumes the worst-case distribution to be the most likely one. The same holds true for an American call with reference distribution $\mathbb{P}^a$: In that case, the reference distribution is also the worst-case one. However, due to the tradeoff effects, $\mathbb{P}^a$ is not the worst-case distribution for the American call when $\mathbb{P}^b$ is the reference distribution; as $\mathbb{P}^b$ is not worst-case distribution for the American put when $\mathbb{P}^a$ is the reference one.

[Föllmer & Schied, 02] introduce convex risk measures based on expected loss or shortfall risk in a static framework. Entropic risk measures are just a special case when loss is exponential. Carrying over these risk measures to a dynamic framework, a fruitful further application could be achieved as risk measures based on *shortfall risk* have a quite intuitive appeal.

3.5.2 Dynamic Generalized AVaR

In the financial industry *value at risk (VaR)* still is a standard method for risk quantification and risk management. Given a confidence level $\lambda \in]0,1[$, VaR of a risky project X is "commonly" defined as

$$VaR_\lambda(X) := \inf \left\{ l \in \mathbb{R} | \, \mathbb{P}(X + l < 0) < \lambda \right\},$$

i.e. the negative of the upper quantile, a definition that might inter alia be found in [Cheridito & Stadje, 09]. Prominence of VaR might be due to its simplicity in applications and its intuitive appeal. Though widely used, VaR is neither convex nor coherent as it is not sub-additive: Applying VaR, a risk officer runs the danger or accumulating a highly risky portfolio. A standard example is inter alia given in [McNeil et al., 05]. Moreover, VaR does not account for the actual magnitude of losses but just loss events. Being aware of VaR's shortcomings, *average value at risk ($AVaR$)* is introduced taking into account not only loss probabilities in terms of quantiles, as VaR does, but also the amount of possible loss. Nevertheless, $AVaR$ is still intuitive and easily implemented by virtue of

$$AVaR_\lambda(X) := \frac{1}{\lambda} \int_0^\lambda VaR_m(X) dm$$

for some level $\lambda \in]0,1[$. It can be shown that AVaR satisfies the robust representation

$$AVaR_\lambda(X) = \sup_{\mathbb{Q} \in \mathcal{M}} \left\{ \mathbb{E}^{\mathbb{Q}}[-X] - \alpha(\mathbb{Q}) \right\}$$

for

$$\alpha^{\min}(\mathbb{Q}) = \begin{cases} 0 & \text{if } \frac{d\mathbb{Q}}{d\mathbb{P}_0} \leq \frac{1}{\lambda}, \\ \infty & \text{else.} \end{cases}$$

Hence $AVaR$ is a coherent risk measure giving raise for multiple prior preferences as considered in [Riedel, 09]. Elaborate discussions on AVaR and further representations can be found in [McNeil et al., 05]. [Föllmer et al., 09] introduce a generalization of AVaR, called *utility based shortfall risk measure.* [Cheridito & Li, 09] use a convenient representation for $AVaR$ which has an immediate generalization to a convex risk measure, called *generalized* $AVaR$ *(*$gAVaR$*)* here. This convex risk measure gives then raise to a variational preference by multiplying the robust representation with -1.

As shown in [Cheridito & Stadje, 09] as well as [Artzner et al., 07] the natural dynamic extension of $AVaR$, and hence of $gAVaR$, just in terms of conditional expectations is *not* time-consistent, cp. [Artzner et al., 99]'s Definition 5.5. We thus define a time-consistent dynamic version of $gAVaR$, called dyn_gAVaR: In a first approach as in [Cheridito & Stadje, 09] recursively in terms of the definition of time-consistency. Thereafter, recursively in terms of the penalty function as in [Maccheroni et al., 06b] by composing one period ahead penalties directly achieving the robust representation. By Corollary 4.8 in [Cheridito et al, 06], both approaches induce the same time-consistent dynamic convex risk measure or, equivalently, the same time-consistent dynamic variational preference.

Consider again the underlying filtered reference space $(\Omega, (\mathcal{F}_t)_{t\leq T}, \mathbb{P}_0)$. Set $L^i_t := L^i(\Omega, \mathcal{F}_t, \mathbb{P}_0|_{\mathcal{F}_t})$ for $i \in \{0\} \cup \mathbb{N} \cup \{\infty\}$, $t \leq T$. To start with, we first consider the static convex risk measure $gAVaR$ for some end period payoff $X_T \in L^\infty_T$ as in [Cheridito & Li, 09], $T < \infty$. Later this static risk measure will serve as dyn_gAVaR_0 in the definition of the dynamic convex risk measure $(dyn\_gAVaR_t)_{t\leq T}$. We obtain robust representations for $gAVaR$ in terms of a penalty $\alpha^{\min}$, serving as $\alpha^{\min}_0$ in the penalty function $(\alpha^{\min}_t)_{t\leq T}$ for $(dyn\_gAVaR_t)_{t\leq T}$.

Definition 3.5.18. *For* $(\theta, \beta, p) \in]0, \infty[\times]1, \infty[\times[1, \infty[$*, define the risk measure gAVaR for* $X_T \in L^\infty_T$*, called* generalized Average Value at Risk (gAVaR)*, by virtue of*

$$gAVaR^{\beta,p}_\theta(X_T) := \min_{s\in\mathbb{R}} \left\{ \frac{1}{\theta} \left\| (s - X_T)^+ \right\|^\beta_p - s \right\},$$

where $\|\cdot\|_p := (\mathbb{E}^{\mathbb{P}_0|_{\mathcal{F}_T}}[|\cdot|^p])^{\frac{1}{p}}$ *denotes the usual p-norm.*

For ease of notation, we do not explicitly state the parameters but just write $gAVaR$ instead of $gAVaR_\theta^{\beta,p}$ when these are obvious. We have:

Proposition 3.5.19. *(a) For* $(\theta,\beta,p) \in \;]0,1[\times\{1\}\times[1,\infty[$, $gAVaR_\theta^{\beta,p}$ *is a* coherent *risk measure for* $X_T \in L_T^\infty$ *with robust representation in terms of minimal penalty* $\alpha^{\min}$ *by virtue of*

$$\alpha^{\min}(\mathbb{Q}) = \begin{cases} 0 & \text{if } \|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\|_q \leq \frac{1}{\theta}, \\ \infty & \text{else,} \end{cases}$$

for $\mathcal{Q} \in \mathcal{M}$, *where* $q := \frac{p}{p-1}$ *and* $\|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\|_q = \left(\mathbb{E}^{\mathbb{P}_0|_{\mathcal{F}_T}}[|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}|^q]\right)^{\frac{1}{q}}$.
(b) For $\theta \in]0,1[$, $\beta = p = 1$, *we have* $\|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\|_\infty = \operatorname{ess\,sup}|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}|$ *and hence the robust representation becomes*

$$\begin{aligned} gAVaR_\theta^{1,1}(X_T) &= \sup_{\mathbb{Q}\in\mathcal{M}} \left\{ \mathbb{E}^{\mathbb{Q}|_{\mathcal{F}_T}}[-X_T] \,\middle|\, 0 \leq \frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}} \leq \frac{1}{\theta} \right\} \\ &= AVaR_\theta(X_T), \end{aligned}$$

which again shows $AVaR$ *to be a coherent risk measure.*
(c) For $(\theta,\beta,p) \in \;]0,\infty[\times]1,\infty[\times[1,\infty[$, $gAVaR_\theta^{\beta,p}$ *is a* convex *risk measure for* $X_T \in L_T^\infty$ *with minimal penalty* $\alpha^{gAVaR}(\mathbb{Q}) := c\|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\|_q^d$, *where* $q := \frac{p}{p-1}$, $d := \frac{\beta}{\beta-1}$ *and* $c = \theta^{d-1}\beta^{1-d}d^{-1}$. *Hence*

$$gAVaR_\theta^{\beta,p}(X_T) = \sup_{\mathbb{Q}\in\mathcal{M}} \left\{ \mathbb{E}^{\mathbb{Q}|_{\mathcal{F}_T}}[-X_T] - c\left\|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\right\|_q^d \right\}.$$

Proof. cp. [Cheridito & Li, 09]. □

[Cheridito & Stadje, 09] recursively achieve a time-consistent dynamic version of $AVaR$ for end period payoff X_T. Mimicking this approach by virtue of the definition of time-consistency for dynamic convex risk measures, i.e. $\rho_t = \rho_t(-\rho_{t+1})$ or ,equivalently, $\pi_t = \pi_t(\pi_{t+1})$ for dynamic variational preferences,[23] we obtain a time-consistent dynamic version of $gAVaR_\theta^{\beta,p}$.

[23] As we assume T being finite, time-consistency of dynamic risk measures is by Proposition 4.5 in [Cheridito et al, 06] equivalent to “one-step time consistency” as applied in this article.

As in [Cheridito & Stadje, 09], we now define a time-consistent version for the more general risk measure $gAVaR$ in terms of the definition of time-consistency:

Definition 3.5.20. *We recursively define the dynamic convex risk measure called* dynamic generalized average value at risk*, $(dyn\_gAVaR_t)_{t\leq T}$, as follows: Let $X_i \in \mathcal{F}_i$, $i \leq T$, then we set for all t*

$$\begin{aligned} dyn\_gAVaR_t(X_j) &:= -X_j, \\ dyn\_gAVaR_t(X_{t+1}) &:= \operatorname*{ess\,inf}_{s\in L_t^\infty} \left\{ \frac{1}{\theta} \left(\mathbb{E}\left[\left|(s - X_{t+1})^+\right|^p \middle| \mathcal{F}_t \right] \right)^{\frac{\beta}{p}} - s \right\}, \\ dyn\_gAVaR_t(X_z) &:= dyn\_gAVaR_t(-dyn\_gAVaR_{t+1}(X_z)) \end{aligned}$$

for $j \leq t$, $t+1 < z \leq T$.[24]

Remark 3.5.21. *In terms of Definition 3.5.20, for an adapted payoff process $(X_t)_{t\leq T}$ and a stopping time $\tau \leq T$, the term $dyn\_gAVaR_t(X_\tau)$ is well defined for $t \leq T$.*

Remark 3.5.22. *From [Cheridito & Stadje, 09], we see that the natural dynamic generalization*

$$gAVaR_t(X_T) := \operatorname*{ess\,inf}_{s\in L_t^\infty} \left\{ \frac{1}{\theta} \left(\mathbb{E}\left[\left|(s - X_T)^+\right|^p \middle| \mathcal{F}_t \right] \right)^{\frac{\beta}{p}} - s \right\}$$

is not time-consistent. But in these terms our definition becomes

$$dyn\_gAVaR_t(X_z) = gAVaR_t(-dyn\_gAVaR_{t+1}(X_z)).$$

Proposition 3.5.23. *$(dyn\_gAVaR_t)_{t\leq T}$ is a time-consistent dynamic convex risk measure, i.e. satisfies for $t < T$*

$$dyn\_gAVaR_t = dyn\_gAVaR_t(-dyn\_gAVaR_{t+1}).$$

[24] The last term is well-defined as $dyn\_gAVaR_{t+1}(X_z)$ is $\mathcal{F}_{t+1}$-measurable. A special case is of course z = T, in which case we are back in the setting of [Cheridito & Stadje, 09] but for gAVaR instead of AVar.

In our optimal stopping approach time-consistency can be written as: For $(X_t)_{t\leq T}$*, and a stopping time* $\tau \leq T$ *we obtain for* $t < T$

$$\begin{aligned} dyn\_gAVaR_t(X_\tau) &= dyn\_gAVaR_t\left(X_\tau \mathbb{I}_{\{\tau\leq t\}} \right. \\ &\qquad \left. - dyn\_gAVaR_{t+1}(X_\tau)\mathbb{I}_{\{\tau\geq t+1\}}\right) \\ &= -X_\tau \mathbb{I}_{\{\tau\leq t\}} \\ &\quad + dyn\_gAVaR_t(-dyn\_gAVaR_{t+1}(X_\tau)\mathbb{I}_{\{\tau\geq t+1\}}). \end{aligned}$$

Proof. Being a dynamic time-consistent convex risk measure is immediate by Definition 3.5.20 in terms of the static convex risk measure $gAVaR$ as the recursion formula is just the definition of time-consistency.

Our special form of time-consistency follows immediately as we have already seen in the theoretical section. Nevertheless, we make it explicit here: As $\tau \leq T$, X_τ is $\mathcal{F}_T$-measurable, i.e. at time T we know when we have stopped the process. Writing $X_\tau = X_\tau \mathbb{I}_{\{\tau\leq t\}} + X_\tau \mathbb{I}_{\{\tau\geq t+1\}}$ we obtain with conditional cash invariance

$$\begin{aligned} & dyn\_gAVaR_t(-dyn\_gAVaR_{t+1}(X_\tau)) \\ = \; & dyn\_gAVaR_t(-dyn\_gAVaR_{t+1}(X_\tau \mathbb{I}_{\{\tau\leq t\}} + X_\tau \mathbb{I}_{\{\tau\geq t+1\}})) \\ = \; & dyn\_gAVaR_t(-\underbrace{dyn\_gAVaR_{t+1}(X_\tau \mathbb{I}_{\{\tau\leq t\}})}_{=-X_\tau \mathbb{I}_{\{\tau=t\}}} \\ & \qquad - dyn\_gAVaR_{t+1}(X_\tau)\mathbb{I}_{\{\tau\geq t+1\}}) \\ = \; & dyn\_gAVaR_t(X_\tau \mathbb{I}_{\{\tau\leq t\}} - dyn\_gAVaR_{t+1}(X_\tau)\mathbb{I}_{\{\tau\geq t+1\}}) \\ = \; & -X_\tau \mathbb{I}_{\{\tau\leq t\}} + dyn\_gAVaR_t(-dyn\_gAVaR_{t+1}(X_\tau)\mathbb{I}_{\{\tau\geq t+1\}}). \end{aligned}$$

□

By [Föllmer & Penner, 06], Theorem 4.5, $(dyn\_gAVaR_t)_{t\leq T}$ then of course possesses a robust representation in terms of a minimal penalty satisfying the no-gain condition by Proposition 3.2.15.

Definition 3.5.24. *We say that the dynamic variational preference* $(\pi_t^{aR})_{t\leq T}$ *is obtained by dynamic generalized average value at risk* $(dyn\_gAVaR_t)_{t\leq T}$ *if* *it is of the form*

$$\pi_t^{aR} := -dyn\_gAVaR_t.$$

Remark 3.5.25. *By Proposition 3.5.23, $(\pi_t^{aR})_{t\leq T}$ is time-consistent, i.e. for $t < T$, $z \leq T$, we have*

$$\pi_t^{aR}(X_z) = \pi_t^{aR}(\pi_{t+1}^{aR}(X_z)),$$

more elaborately for a stopping time $\tau \leq T$

$$\pi_t^{aR}(X_\tau) = X_\tau \mathbb{I}_{\{\tau\leq t\}} + \pi_t^{aR}\left(\pi_{t+1}^{aR}(X_\tau)\mathbb{I}_{\{\tau\geq t+1\}}\right)$$

which shows time-consistency in terms of Proposition 3.2.16.

As the assertion in Theorem 3.4.1 can be reformulated not to make use of the robust representation of dynamic variational preferences, we can directly apply the variational Snell envelope approach[25] and achieve for $t < T$

$$\begin{aligned} U_t &= \max\left\{X_t; \pi_t^{aR}(U_{t+1})\right\} \\ &= \max\left\{X_t; -dyn\_gAVaR_t(U_{t+1})\right\} \\ &= \max\left\{X_t; \operatorname*{ess\,sup}_{s\in L_t^\infty}\left\{s - \frac{1}{\theta}\left(\mathbb{E}\left[\left|(s-U_{t+1})^+\right|^p \middle| \mathcal{F}_t\right]\right)^{\frac{\beta}{p}}\right\}\right\} \end{aligned}$$

as U_{t+1} is $\mathcal{F}_{t+1}$-measurable. In order to achieve explicit solutions in terms of worst-case distributions as done in the theoretical section, we rather want to have the robust representation of $(dyn\_gAVaR_t)_{t\leq T}$. Hence, we end this section by establishing an alternative way to introduce a time-consistent dynamic version of $gAVaR$ in terms of a robust representation, i.e. we appropriately define a penalty $(\alpha_t^{\text{gAVaR}})_{t\leq T}$: Hereto, we will use the minimal penalty α^{gAVaR} of the static $gAVaR$ as defined in Proposition 3.5.19. We apply the recursive procedure from [Maccheroni et al., 06b] in terms of one

[25] In [Cheridito et al, 06], Section 5.3, optimal stopping problems with general monetary risk measures are considered. In that case, the Snell envelope can only be given in this form as the risk measure does not necessarily possess a robust representation.

period ahead penalties $(\gamma_t)_{t\leq T}$ to achieve a time-consistent dynamic minimal penalty $(\alpha_t^{\text{gAVaR}})_{t\leq T}$. We then show that the dynamic time-consistent variational preferences obtained by virtue of both procedures coincide.

To ease notation, we do not state this example in terms of one-period ahead penalties γ_t but in terms of s-period ahead penalties $\alpha_{t,t+s}^{\min}$, $s \geq 0$, as defined in [Föllmer & Penner, 06], p. 76. s-period ahead penalties constitute a direct generalization of our one-period ahead penalties by virtue of $\gamma_t(\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)) = \alpha_{t,t+1}^{\min}(\mathbb{Q})$. We do not rigorously introduce the theory in terms of these more general s-period ahead penalties: All assertions, in particular the no-gain condition, can be analogously stated in terms of $\alpha_{t,t+1}^{\min}$. The respective results are given in [Föllmer & Penner, 06], Theorem 4.5.

Making use of α^{gAVaR} in Proposition 3.5.19(c), define the s-period ahead penalty at t by

$$\alpha_{t,t+s}^{gAVaR}(\mathbb{Q}) := \alpha^{gAVaR}(\mathbb{Q}|_{\mathcal{F}_{t+s}}(\cdot|\mathcal{F}_t)) = c\left(\mathbb{E}^{\mathbb{P}_0}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}_0}\bigg|_{\mathcal{F}_{t+s}}\right)^q\bigg|\mathcal{F}_t\right]\right)^{\frac{d}{q}}$$

for $s \geq 0$, $t+s \leq T$, $\mathbb{Q} \in \mathcal{M}$, and the parameters as in Proposition 3.5.19(c). Note, that we then have $\alpha_{0,0+T}^{gAVaR}(\mathbb{Q}) = \alpha^{gAVaR}(\mathbb{Q}) = c\left\|\frac{d\mathbb{Q}}{d\mathbb{P}_0}\Big|_{\mathcal{F}_T}\right\|_q^d$. Then, the one period ahead penalty γ_t^{gAVaR} on $\mathcal{M}|_{\mathcal{F}_{t+1}}$ is defined by

$$\gamma_t^{gAVaR}(\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)) := \alpha_{t,t+1}^{gAVaR}(\mathbb{Q}) = c\left(\mathbb{E}^{\mathbb{P}_0}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}_0}\bigg|_{\mathcal{F}_{t+1}}\right)^q\bigg|\mathcal{F}_t\right]\right)^{\frac{d}{q}}.$$

Given this one-step ahead penalty, we recursively define a dynamic penalty $(\alpha_t^{gAVaR})_{t\leq T}$ as in Theorem 2 in [Maccheroni et al., 06b]:

Definition 3.5.26. *Let $F_t \in \mathcal{F}_t$. We define the dynamic penalty $(\alpha_t^{gAVaR})_{t\leq T}$ by virtue of*

$$\begin{aligned}
\alpha_T^{gAVaR}(\mathbb{Q})(\omega) &:= \begin{cases} 0 & \text{if } \mathbb{Q} = \mathbb{I}_{\{\omega\}}, \\ \infty & \text{else} \end{cases} \qquad \text{for } \omega \in \Omega, \\
\alpha_t^{gAVaR}(\mathbb{Q})(F_t) &:= \int \alpha_{t+1}^{gAVaR}(\mathbb{Q}(\cdot|\mathcal{F}_{t+1}))d\mathbb{Q}(\cdot|F_t) + \gamma_t^{gAVaR}(\mathbb{Q}(\cdot|F_t)|_{\mathcal{F}_{t+1}})
\end{aligned}$$

$$\begin{aligned} & & & \textit{if } \mathbb{Q}(F_t) > 0, \\ \alpha_t^{gAVaR}(\mathbb{Q})(F_t) &:=& \infty & \textit{ if } \mathbb{Q}(F_t) = 0, \end{aligned}$$

for $t < T$.[26]

Applying $(\alpha_t^{gAVaR})_{t\leq T}$ *to a robust representation, we define dynamic variational preferences* $(\pi_t^{\alpha^{gAVaR}})_{t\leq T}$ *by*

$$\pi_t^{\alpha^{gAVaR}}(X_T) := \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{ \mathbb{E}^{\mathbb{Q}}[X_T|\mathcal{F}_t] + \alpha_t^{gAVaR}(\mathbb{Q}) \right\}$$

for $X_T \in L_T^\infty$.

Remark 3.5.27. $(\pi_t^{\alpha^{gAVaR}})_{t\leq T}$ *is a time-consistent dynamic variational preference. Indeed: It is a dynamic variational preference by virtue of its definition in terms of a robust representation. Time-consistency of* $(\pi_t^{\alpha^{gAVaR}})_{t\leq T}$ *follows by Proposition 3.2.15 as the penalty* $(\alpha_t^{gAVaR})_{t\leq T}$ *is defined recursively in terms of the no-gain condition.*

We have achieved two distinct time-consistent variational preferences generalizing $gAVaR$: $(\pi_t^{aR})_{t\leq T} = (-dyn\_gAVaR_t)_{t\leq T}$ and $(\pi_t^{\alpha^{gAVaR}})_{t\leq T}$. We now show that these preferences coincide, i.e.

$$(\pi_t^{aR})_{t\leq T} = (\pi_t^{\alpha^{gAVaR}})_{t\leq T},$$

given equality of the respective model parameters not explicitly stated here. By Corollary 4.8 in [Cheridito et al, 06], it suffices to check that $\pi_0^{aR}(X_T) = \pi_0^{\alpha^{gAVaR}}(X_T)$ for $\mathcal{F}_T$-measurable random variables X_T. However, for both π_0^{aR} as well as $\pi_0^{\alpha^{gAVaR}}$ we have a robust representation:

$$\pi_0^{\alpha^{gAVaR}}(X_T) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{ \mathbb{E}^{\mathbb{Q}}\left[X_T\right] + \alpha_0^{gAVaR}(\mathbb{Q}) \right\},$$

[26]Intuitively, $\alpha_T^{\text{gAVaR}}(\mathbb{Q})(\omega)$ is the penalty that only allows for the observed path $(\omega_1, \ldots, \omega_T)$.

and on the other hand we have

$$\begin{aligned} \pi_0^{aR}(X_T) &= -dyn\_gAVaR_0(X_T) \\ &= -\operatorname*{ess\,inf}_{s\in\mathbb{R}} \left\{ \frac{1}{\theta} \left(\mathbb{E}^{\mathbb{P}_0} \left[|s - X_T|^p \right] \right)^{\frac{\beta}{p}} - s \right\} \\ &= -gAVaR(X_T) \\ &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{ \mathbb{E}^{\mathbb{Q}|\mathcal{F}_T} \left[X_T \right] + \alpha^{gAVaR}(\mathbb{Q}) \right\}, \end{aligned}$$

where the second equality follows by Definition 3.5.20 and time-consistency, and the last by Proposition 3.5.19. Hence, it suffices to show equality of the minimal penalties at $t = 0$, i.e. for all $\mathbb{Q} \in \mathcal{M}$, we have to show

$$\alpha^{gAVaR}(\mathbb{Q}) = \alpha_0^{gAVaR}(\mathbb{Q}).$$

Indeed: As we have seen that $\alpha_{0,0+T}^{gAVaR}(\mathbb{Q}) = \alpha^{gAVaR}(\mathbb{Q})$, it leaves to show $\alpha_{0,0+T}^{gAVaR}(\mathbb{Q}) = \alpha_0^{gAVaR}(\mathbb{Q})$. By Theorem 4.5 in [Föllmer & Penner, 06], we have the no-gain condition for s-period ahead penalties reducing to

$$\alpha_0^{gAVaR}(\mathbb{Q}) = \alpha_{0,0+T}^{gAVaR}(\mathbb{Q}) + \mathbb{E}^{\mathbb{Q}} \left[\alpha_T^{gAVaR}(\mathbb{Q}) \middle| \mathcal{F}_0 \right].$$

The right hand side equals $\alpha_{0,0+T}^{gAVaR}(\mathbb{Q})$ as $\mathbb{E}^{\mathbb{Q}} \left[\alpha_T^{gAVaR}(\mathbb{Q}) \middle| \mathcal{F}_0 \right] = 0$ by definition of α_T and the assumption that $\mathbb{Q} \in \mathcal{M}$: Otherwise $\mathbb{E}^{\mathbb{Q}} \left[\alpha_T^{gAVaR}(\mathbb{Q}) \middle| \mathcal{F}_0 \right] = \infty$ contradicting $\mathbb{Q} \in \mathcal{M}$.

Hence, both time-consistent dynamic variational preferences, π^{aR} and $\pi^{\alpha^{gAVaR}}$, coincide and we have

$$\begin{aligned} \pi_t^{aR}(X_\tau) &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{ \mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{gAVaR}(\mathbb{Q}) \right\} \\ &= X_\tau \mathbb{I}_{\{\tau\le t\}} \\ &\quad + \min_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int \pi_{t+1}^{aR}(X_\tau) d\mu + \gamma_t^{gAVaR}(\mu) \right) \mathbb{I}_{\{\tau\ge t+1\}}. \end{aligned}$$

We have the following recursive representation of the Snell envelope of time-consistent dynamic variational preferences induced by dynamic generalized average value at risk, $(dyn\_gAVaR_t)_{t\leq T}$:

$$\begin{aligned} U_t &= \max\left\{X_t; \pi_t^{aR}(U_{t+1})\right\} \\ &= \max\left\{X_t; \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}}\left(\int \pi_{t+1}^{aR}(U_{t+1})d\mu + \gamma_t^{gAVaR}(\mu)\right)\right\} \\ &= \max\left\{X_t; \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}}\left(\int U_{t+1}d\mu + c\left(\mathbb{E}\left[\left|\frac{d\mu}{d\mathbb{P}_0|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)}\right|^q \middle| \mathcal{F}_t\right]\right)^{\frac{d}{q}}\right)\right\}. \end{aligned}$$

This representation enables us, given an explicit structure of $(X_t)_{t\leq T}$, to solve the problem for an optimal stopping time τ^* as in Theorem 3.4.1.

3.6 Conclusions

We have generalized the theory of optimal stopping under multiple priors as set out in [Riedel, 09] to dynamic convex risk measures or, equivalently, dynamic variational preferences introduced in [Maccheroni et al., 06b]. To achieve our results, we have introduced the notion of variational supermartingales as a generalization of the usual notion of supermartingales. For this concept, we have obtained results including a Doob decomposition and optional sampling. These enabled us to generalize the classical optimal stopping approach for an expected utility maximizer in [Neveu, 75] (Section VI.1) in terms of Snell envelopes to the case of dynamic variational preferences by virtue of variational Snell envelopes. We have achieved general optimal stopping times for this problem and have shown that the solution to the infitite horizon problem can be approximated by a sequence of solutions for an approximating sequence of finite horizon problems. A further insight is a minimax theorem similar to a minimax result in [Schied, 07] but making use of time-consistency.

Our results were then applied to prominent examples: dynamic entropic risk measures and dynamic generalized average value at risk. For the latter, we are not aware of any reference having obtained this notion to a dynamic context. We applied static generalized average value at risk to a dynamic set up solving a severe time-consistency issue. We have achieved a recursive representation directly applicable to the optimal stopping approach in terms of variational Snell envelopes.

To conclude, the virtue of the present article is that optimal stopping problems are now solved for convex risk measures. This is important for applications on financial markets: coherent risk measures, as a robust approach reducing model risk, are quite conservative. Convex risk measures are a comprehensive vehicle to more liberally assess risk while still being robust: No specific probabilistic model is assumed but a penalty representing the likelihood of distinct models.

Of course, our approach leaves a realm for further generalizations. It seems possible to achieve the results in this article for general time-consistent (monotone) monetary risk measures, i.e. relaxing the convexity assumption. Of course, in that case, the robust representation in terms of penalty α does not hold anymore. Hence, proofs have to be adjusted accordingly. However, as we have explicitly stated in one of the examples, the variational Snell envelope does not need a robust representation and can hence be generalized to more general risk measures.[27] The next direction in which theory might be generalized is to relax the assumption of the payoff process being bounded.

[27] In [Cheridito et al, 06], Chapter 5.3, the authors introduce a stopping problem for more general dynamic risk measures relaxing the convexity assumption. It is assumed that expected reward $(\pi_t)_{t\leq T}$ is induced by a dynamic time-consistent monetary risk measure, i.e. a dynamic time-consistent monotone translation invariant risk measure. As the convexity assumption is relaxed, $(\pi_t)_{t\leq T}$ does not convey the robust representation crucial for our recursive solution. However, having the agent maximizing over her set of stopping times, the usual Snell envelope approach as set out in [Neveu, 75] is still valid. Hence, [Cheridito et al, 06] achieve equality of the Snell envelope and the value function as well as the smallest optimal stopping time as in Theorem 3.4.1. Moreover, they show the value function to be time-consistent and again a monetary utility function, i.e. the value function again has all properties of expected reward $(\pi_t)_{t\leq T}$. Due to a missing robust representation, the solution is not explicit.

Several of the cited references consider convex risk measures for L^p processes or, as in [Cheridito & Li, 09], risk measures defined on Orlicz spaces.

Besides these theoretical considerations, further examples and concrete applications might be elaborated. As mentioned in the text, the theory should be applied to dynamic risk measures based on expected shortfall as a generalization of dynamic entropic risk measures or dynamic variational preferences. These can inter alia be found in [Föllmer et al., 09].

At last, the problem might be considered in a time-continuous setting. Several approaches to convex risk measures in a time-continuous framework are available: In [Bion-Nadal, 08], dynamic convex risk measures are achieved by virtue of BMO martingales. A special case of this approach is given in [Rosazza Gianin, 04] and [Rosazza Gianin, 06] via BSDE resulting in g-expectations as introduced in [Peng, 97].

Chapter 4

Learning for Convex Risk Measures with Increasing Information

4.1 Introduction

Reaching decisions concerning risky projects in a dynamic system, an agent faces new information consecutively influencing her assessment of risk instantaneously.

In this article, we answer the question how anticipation of risk evolves over time when an agent gathers information. We show that, in the limit, all uncertainty is revealed but risk remains if the agent perceives risk in terms of time-consistent dynamic convex risk measures and, hence, generalize the famous Blackwell-Dubins Theorem to convex risk measures. We then relax the time-consistency assumption and show the result to still be valid. Hereto, a fundamental assumption is existence of a reference distribution that fixes impossible and sure events by virtue of equivalence of distributions under consideration.

Coherent risk measures were introduced by virtue of an axiomatic ansatz 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, 09]. 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. 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 system: [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] applies 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. 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.

A parametric learning model in an uncertain environment for dynamic coherent risk measures or, equivalently, dynamic multiple priors as introduced in [Epstein & Schneider, 03], is elaborated in [Epstein & Schneider, 07]. The main virtue of this article is to introduce learning based on experience to convex risk measures models. First, we try to introduce learning in a constructive approach: we design a minimal penalty function and plug it into the robust representation: Since the penalty might be seen as some inverse likelihood of a specific prior distribution, we first apply a quite simple and intuitive learning mechanism to the penalty. We calculate the likelihood of a distribution given past experience and use this as updated penalty. The intuition behind this approach is quite simple: observing good events, distributions of a payoff process that are "stochastically more dominated", i.e. put more weight on bad events, become more unlikely, i.e. have a higher penalty. However, besides its intuitive appeal, it turns out that this procedure does not result in a penalty function as it is backwards oriented and a penalty function, by definition, incorporates probability distributions of the future movement of the payoff process. In a second, more sophisticated approach, we model a penalty incorporating projections of "past" likelihoods on future distributions. Here, we make use of the conditional relative entropy as penalty function: we achieve a proper penalty that penalizes distributions according to "distance" from the "most likely" distribution serving as reference distribution. However, the convex risk measure in terms of this penalty turns out not to be time-consistent in general as shown by a counterexample. In [Epstein & Schneider, 07], time-consistency is not an issue as multiplicity of priors is not introduced in terms of multiple equally likely distributions of the payoff process as e.g. in [Riedel, 09] or [Maccheroni et al., 06a], but in terms of multiple distributions on the parameter space.

Our further approach is not constructive but takes the robust representation of a risk measure in terms of minimal penalty for granted. As the 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. We have hence achieved *learning as an intrinsic property* of dynamic convex risk measures.

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.

[Schnyder, 02] discusses H.P. Minsky's theory of financial instability, a huge portion of which is caused by herding on financial markets. Besides, herding is usually one of the major objections towards Basel II. Our result however shows that, in the long run, there is hardly any chance to circumvent herding behavior.

The article is considered in a parametric setting. However, the second part can be restated in a non parametric setting. It is structured as follows: The next section formally introduces the underlying probabilistic model. Section 3 elaborately discusses robust representation of dynamic (time-consistent) convex risk measures. Constructive approaches to learning in terms of dynamic minimal penalty as well as their shortcomings are stated in Section

4. Section 5 generalizes the Blackwell-Dubins Theorem to conditional expectations. The following two sections then apply this result to coherent and convex risk measures first in the time-consistent case and then in the case without time-consistency. Section 8 states examples. Then we conclude.

4.2 Model

For our model we start with a discrete time set $t \in \{0, ..., T\}$ where T is an infinite time horizon. We will now construct an underlying filtered reference space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P}^{\theta_0})$ and define risky projects X:

We fix $(S, \mathcal{A})$ as a measurable space where S describes the possible states of the world at a fixed point in time t and define Ω to be all possible states of the world, formally the set of sequences of elements of S. For this let $S_t = S$ for all $t \in \{0, ..., T\}$ and then define $\Omega := \bigotimes_{t=0}^{T} S_t$. 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 $\pi_1, ..., \pi_t$. Additionally define all sequences up to time t by $\Omega^t := \bigotimes_{s=0}^{t} S_s$. Denote generic elements on these spaces by $s_t \in S_t$, $s \in \Omega$, $s^t \in \Omega^t$ and $a_t \in \mathcal{A}$.

Let Θ be a set of parameters where every $\theta \in \Theta$ uniquely defines a distribution $\mathbb{P}^\theta$ on $(\Omega, \mathcal{F})$ with filtration $(\mathcal{F}_t)_t$ and fix $\mathbb{P}^{\theta_0}$ as a reference distribution which can be seen as the true distribution of the states. For all $\theta \in \Theta$, $\mathbb{P}^\theta$ is assumed to be equivalent to $\mathbb{P}^{\theta_0}$. Let $\mathcal{M}^e(\mathbb{P}^{\theta_0})$ denote the set of all distributions on $(\Omega, \mathcal{F})$ equivalent to $\mathbb{P}^{\theta_0}$. Assume that all these can be achieved by parameters $\theta \in \Theta$, i.e. $\mathcal{M}^e(\mathbb{P}^{\theta_0}) = \{\mathbb{P}^\theta | \theta \in \Theta\}$. For $\mathbb{P}^\theta \in \mathcal{M}^e(\mathbb{P}^{\theta_0})$ let $\mathbb{P}^\theta(\cdot|\mathcal{F}_t)$ denote the distribution conditional on $\mathcal{F}_t$. Due to our assumption to only consider distributions equivalent to $\mathbb{P}^{\theta_0}$, the reference distribution merely fixes the null-sets of the model, i.e. distinct agents at least agree on impossible and sure events. 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 by Epstein and Marinacci in [Epstein & Marinacci, 07]. They related it to an axiom on preferences first postulated by Kreps in [Kreps, 79]. He claimed that if an agent 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 $\operatorname{ess\,sup}|X| = \kappa > 0$. Having constructed the filtered reference space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}^{\theta_0})$ as above, the sets of almost surely bounded $\mathcal{F}$-measurable and $\mathcal{F}_t$-measurable random variables are denoted by $L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P}^{\theta_0})$ and $L^\infty_t := L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}^{\theta_0})$, respectively. All equations have to be understood $\mathbb{P}^{\theta_0}$-almost surely.

Remark 4.2.1. *As we will see in course of the article, the parametric setting is only needed in the first part on the constructive approach to learning. All statements in the second part, the generalization of the Blackwell-Dubins theorem, can be posed in terms of an arbitrary underlying filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}_0)$ with distributions in $\mathcal{M}^e(\mathbb{P}_0)$, where $\mathbb{P}_0$ denotes the reference distribution, i.e. in a non-parametric setting. Moreover, for these results, we do not need the particular structure of Ω in terms of a product of marginal spaces S_t. We however follow the parametric approach throughout to obtain a unified appearance.*

4.3 Dynamic Convex Risk Measures

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 described 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 4.3.1 (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*

$$\rho_t(X) = \underset{\mathbb{Q}\in\mathcal{M}^e(\mathbb{P}^{\theta_0})}{\text{ess sup}} \left(\mathbb{E}^{\mathbb{Q}}\left[-X|\,\mathcal{F}_t\right] - \alpha_t(\mathbb{Q})\right),$$

where $(\alpha_t)_t$ *denotes the dynamic penalty function, i.e. a family of mappings* $\alpha_t : \mathcal{M}^e(\mathbb{P}^{\theta_0}) \to L_t^\infty$*,* $\alpha_t(\mathbb{Q}) \in \mathbb{R}_+ \cup \infty$*, 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.*

$$u_t(X) := \underset{\mathbb{Q}\in\mathcal{M}^e(\mathbb{P}^{\theta_0})}{\text{ess inf}} \left(\mathbb{E}^{\mathbb{Q}}\left[X|\,\mathcal{F}_t\right] + \alpha_t(\mathbb{Q})\right).$$

Remark 4.3.2. (a) *By Theorem* 4.5 *in [Föllmer & Penner, 06], the above robust representation in terms of* $\mathcal{M}^e(\mathbb{P}^{\theta_0})$ *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 equivalently for dynamic variational preferences. However, we have chosen to concentrate on dynamic convex risk measures here.*

Assumption 4.3.3. *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* $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}^{\theta_0})$

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

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. Throughout, we assume a representation in terms of the minimal penalty $(\alpha_t^{\min})_t$.

Remark 4.3.4 (Equivalent Notation)**.** *In our parametric set-up, a distribution $\mathbb{P}^\theta$ of the process is uniquely defined by a parameter $\theta \in \Theta$. Hence, we write*

$$\rho_t(X) = \operatorname*{ess\,sup}_{\theta\in\Theta} \left(\mathbb{E}^{\mathbb{P}^\theta}\left[-X|\,\mathcal{F}_t\right] - \alpha_t^{\min}(\theta)\right).$$

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

Remark 4.3.5 (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*

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

for $\tilde{\mathcal{Q}}$ the set of prior distributions induced by all θ in some set $\tilde{\Theta} \subset \Theta$. Throughout, $\tilde{\mathcal{Q}}$ is assumed to be convex and weakly compact or, equivalently, $\tilde{\Theta}$ is assumed to be such.

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

Definition 4.3.6 (Time-Consistency)**.** *A dynamic convex risk measure $(\rho_t)_t$ is called* time-consistent *if, for all $t, s \in \mathbb{N}$, it holds*

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

or, equivalently, $u_t = u_t(u_{t+s})$.

Remark 4.3.7. *For the special approach here, [Cheridito et al, 06] show that it suffices to consider $s = 1$ in the above definition.*

Remark 4.3.8. *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 $(\alpha_t^{\min})_t$ 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 4.3.9 (Relative Conditional Entropy)**.** *For $\mathbb{P} \ll \mathbb{Q}$, we define the* relative entropy *of $\mathbb{P}$ with respect to $\mathbb{Q}$ at time $t \geq 0$ as*

$$H_t(\mathbb{P}|\mathbb{Q}) := \mathbb{E}^{\mathbb{P}}\left[\log Z_t\right],$$

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 $\mathbb{Q}$. Furthermore, we define the conditional relative entropy *of $\mathbb{P}$ with respect to $\mathbb{Q}$ at time $t \geq 0$ as*

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

Definition 4.3.10 (Entropic Risk Measures)**.** *Given reference model $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$. Let $\delta > 0$. 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 $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}^{\theta_0})$ if it is of the form*

$$\rho_t^e(X) = \operatorname*{ess\,sup}_{\mathbb{P}\in\mathcal{M}^e(\mathbb{P}^{\theta_0})} \left(\mathbb{E}^{\mathbb{P}}[-X|\mathcal{F}_t] - \delta\hat{H}_t(\mathbb{P}|\mathbb{Q})\right). \tag{4.1}$$

Equivalently, dynamic multiplier preferences *$(u_t^e)_t$ are defined by virtue of*

$$u_t^e(X) = \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}^e(\mathbb{P}^{\theta_0})} \left(\mathbb{E}^{\mathbb{P}}[X|\mathcal{F}_t] + \delta\hat{H}_t(\mathbb{P}|\mathbb{Q})\right). \tag{4.2}$$

Remark 4.3.11. *The variational formula for relative entropy implies*

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

Intuitively, 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. Again, we can write $(\rho_t^e)_t$ by virtue of

$$\rho_t^e(X) = \operatorname*{ess\,sup}_{\theta \in \Theta} \left(\mathbb{E}^{\mathbb{P}^\theta}[-X|\mathcal{F}_t] - \delta \hat{H}_t(\theta|\eta) \right),$$

where $\mathbb{P}^\eta$ defines the reference model.

4.4 A Constructive Approach to Learning

In this section, we try to explicitly develop a learning mechanism by virtue of penalty functions that are then used for the robust representation of dynamic convex risk measures. We will encounter, that this is not an eligible approach to model learning as it is still not clear how to explicitly form a penalty. In a later section, we will just take the robust representation as given and pose the question what can be said about learning when distinct properties of the penalty are assumed.

4.4.1 The Intuition of Learning via Penalties

In a first, intuitive approach, we explicitly introduce a learning mechanism to the penalty $(\alpha_t)_t$ in terms of a likelihood function. The fundamental idea is that the penalty might be viewed as a measure for the likelihood of a distribution. In the extreme case of coherent risk, this means

- $\alpha_t(\theta) = \infty$: $\mathbb{P}^\theta$ is not possible,
- $\alpha_t(\theta) = 0$: $\mathbb{P}^\theta$ is among the most likely.

In general, the larger α_t, the less likely the respective distribution. Stated in other terms, $(\alpha_t)_t$ is a measure for uncertainty aversion: given two penalties $(\alpha_t^1)_t$ and $(\alpha_t^2)_t$, the a.s. larger one corresponds to the less uncerteinty averse agent. In the entropic case, $\alpha_t(\theta) = H_t(\mathbb{P}^\theta|\mathbb{P}^{\bar{\theta}})$, the conditional relative entropy of $\mathbb{P}^\theta$ with respect to $\mathbb{P}^{\bar{\theta}}$ at time t, the agent considers $\mathbb{P}^{\bar{\theta}}$ most likely as $H_t(\mathbb{P}^{\bar{\theta}}|\mathbb{P}^{\bar{\theta}}) = 0$ and distributions "further away" as more and more unlikely.

In the coherent case characterized by a trivial penalty, learning means to alternate the sets $\tilde{\mathcal{Q}}_t := \{\mathbb{P} \in \tilde{\mathcal{Q}} \mid \mathbb{P}(\cdot|\mathcal{F}_t) \in \tilde{\mathcal{Q}}(\cdot|\mathcal{F}_t)\}$, $t = 0, ..., T$ of conditional priors on which the penalty as value zero: when more information is available and hence, more might be known about the distribution that rules the world, $\tilde{\mathcal{Q}}_t \supset \tilde{\mathcal{Q}_{t+1}}$, i.e. penalty is increasing in t. For some cut off value β, an intuitive approach would be in terms of some likelihood function l:

$$\alpha_t(\theta) = \begin{cases} 0 & \text{if } l(\mathbb{P}^\theta|\theta, \mathcal{F}_t) \geq \beta, \\ \infty & \text{else.} \end{cases}$$

As a direct generalization to convex risk measures, one might consider the log-likelihood $-\log(l(\mathbb{Q}|\theta, \mathcal{F}_t))$ as penalty. It will turn out that this approach is not eligible since a penalty defined in terms of likelihood functions is not feasible. Hence, we come up with a distinct ansatz in which penalty is given by relative conditional entropy. We then achieve a dynamic convex risk measure but run into trouble regarding time-consistency. A model defined as above serves as a measure theoretic fundament of H.P. Minsky's theory of financial instability: A sequence of "good" events causes the penalty to be smaller for distributions that stochastically dominate for the payoff under consideration. Upon observing favorable events, the agent thinks that nature has become kinder. This might help to understand underestimation of risk leading to bubbles and financial instability in times of growth and financial success.

4.4.2 Special Case: Explicit Learning for Coherent Risk

[Epstein & Schneider, 07] introduce learning for coherent risk in terms of likelihood ratio tests. As we will see later, they do not consider the sets of priors $(\mathcal{Q}_t)_t$ as for example in [Riedel, 04] but the process $\mathcal{P}_t(\mathcal{F}_t)$ of one-step ahead conditional beliefs, formally introduced below, as these immediately represent the learning process. Moreover, [Epstein & Schneider, 07] distinguish between information that can be learned and information that cannot: Information that can be learned is incorporated in a the set of priors not being singleton, information that cannot be learned is incorporated in the set of likelihood functions not being singleton.

Formally, let the state space be given by $S^T := \otimes_{t=1}^T S_t$, $S_t = S$, Θ as in the general model. The space of parameters will be slightly modified, i.e. every $\theta \in \Theta$ uniquely characterizes a distribution on S and not on Ω; however, this modification is restricted to the current subsection. Let $\mathcal{Q}_0 \subset \mathcal{M}(\Theta)$ be the set of priors on Θ and $\mathcal{L}$ the set of likelihoods, i.e. every $l \in \mathcal{L}$ satisfies $l(\cdot|\theta) \in \mathcal{M}(S)$ and $l(s_t|\cdot)$ is $\mathcal{F}_t$-measurable for $s_t \in S_t$. Set $s^t = (s_1, \dots, s_t)$, $s_i \in S_i$. Every $\mu_0 \in \mathcal{Q}_0$ together with a family of likelihoods $(l_1, l_2, \dots) \in \mathcal{L}^\infty$ induces a prior $\mathbb{P} \in \mathcal{M}^e(\mathbb{P}_0)$ of the payoff process or, equivalently, the process $(p_t)_t$ of one-step-ahead conditionals

$$p_t(\cdot|s^t) = \int_\Theta l(\cdot|\theta) d\mu_t(\theta|s^t) \quad \in \mathcal{M}(S_{t+1}),$$

where μ_t is derived from μ_0 as described below and $\mu_t(\cdot|s^t) \in \mathcal{Q}_t(s^t)$, the set of posterior beliefs on Θ given history s^t. Hence, multiplicity of beliefs is described by

$$\begin{aligned} \mathcal{P}_t(s^t) &= \left\{ p_t(\cdot|s^t) = \int_\Theta l(\cdot|\theta) d\mu_t(\theta) \;\middle|\; \mu_t \in \mathcal{Q}_t^\alpha(s^t), l \in \mathcal{L} \right\} \\ &:= \int_\Theta \mathcal{L}(\cdot|\theta) d\mathcal{Q}_t^\alpha(\theta). \end{aligned}$$

To complete the model, it leaves to show how $(\mu_0; l_1, \dots)$ induce μ_t or, equivalently, how $\mathcal{Q}_t(s^t)$ is obtained. For $(\mu_0; l_1, \dots)$, the posteriors are obtained by Bayesian updating:

$$\begin{aligned} & d\mu_t(\cdot, s^t, \mu_0, l^t) \\ = \; & \frac{l_t(s_t|\cdot)}{\int_\Theta l_t(s_t|\tilde{\theta} d\mu_{t-1}(\tilde{\theta}, s^{t-1}, \mu_0, l^{t-1})} d\mu_{t-1}(\cdot, s^{t-1}, \mu_0, l^{t-1}). \end{aligned}$$

Then, the posteriors are achieved by virtue of a likelihood ratio test in terms of the unconditional data density:

$$\mathcal{Q}_t^\alpha(s^t) := \left\{ \mu_t(s^t, \mu_0, l^t) \middle| \mu_0 \in \mathcal{Q}_0, l^t \in \mathcal{L}^t, \int \prod_{j=1}^t l_j(s_j|\theta) d\mu_0(\theta) \right.$$
$$\left. \geq \beta \max_{\bar{\mu}_0 \in \mathcal{Q}_0, \bar{l}^t \in \mathcal{L}^t} \int \prod_{j=1}^t \bar{l}_j(s_j|\theta) d\bar{\mu}_0(\theta) \right\}$$

for some bound $\beta \in \mathbb{R}^+$.

Remark 4.4.1. *Conceptually, there is a huge difference between the approach in [Epstein & Schneider, 07] and [Gilboa & Schmeidler, 89]: In the latter, the term "multiple priors" means multiple distributions of the payoff stream, all being equally likely, in the former, it means multiple distributions of the parameter, i.e. multiple distributions on the distributions of the payoff stream. Hence, [Epstein & Schneider, 07] is a generalization of [Gilboa & Schmeidler, 89] as the latter framework is achieved with $\mathcal{Q}_0 = \{\mu_0\}$ with μ_0 the uniform distribution on some subset of Θ. In that case we have a trivial α and hence a coherent risk measure. Intuitively, a uniform distribution on a subset of Θ corresponds to the agent believing all distributions in that subset being equally likely and the others impossible.*

Nevertheless, fruitful insights from [Epstein & Schneider, 07] can be gained for our approach in particular the incorporation of a likelihood ratio test. We go a step closer to [Gilboa & Schmeidler, 89] and introduce a single distribution on Θ inducing a unique penalty for a dynamic convex risk measure.

4.4.3 A First, Particularly Intuitive Approach: Simplistic Learning

As stated above, multiple prior preferences mean the agent has a uniform distribution on a subset of Θ: She is sure about which parameters are possible and which not, but has no tendency towards their likeliness. In a way, this corresponds to a non-informative weighting or a trivial penalty function α_0. We act on this non-informative approach and assume the following penalty at time zero: Let $\tilde{\Theta} \subset \Theta$. The penalty corresponding to this distribution is given by:

$$\alpha_t(\theta) = \begin{cases} 0 & \text{if } \theta \in \tilde{\Theta}, \\ \infty & \text{else.} \end{cases}$$

Hence, initially the convex risk measure is actually coherent:

$$\rho_0(X) := \operatorname*{ess\,sup}_{\theta \in \Theta} \left\{ \mathbb{E}^{\mathbb{P}^\theta}[-X] - \alpha_0(\theta) \right\} = \operatorname*{ess\,sup}_{\theta \in \tilde{\Theta}} \mathbb{E}^{\mathbb{P}^\theta}[-X].$$

We now come up with a simple learning mechanism directly defining the dynamic penalty function $(\alpha_t)_t$ in terms of likelihoods. At $t = 0$, we have already characterized the penalty. Furthermore, we set

$$\alpha_1(\theta) := -\ln\left(\frac{l(s_1|\theta)}{\sup_{\bar{\theta}} l(s_1|\bar{\theta})}\right) = -\ln\left(\frac{\mathbb{Q}^\theta(s_1)}{\sup_{\bar{\theta}} \mathbb{Q}^{\bar{\theta}}(s_1)}\right),$$

where $s_1 = s^1$ and

$$\alpha_2(\theta) = -\ln\left(\frac{l(s^2|\theta)}{\sup_{\bar{\theta}} l(s^2|\bar{\theta})}\right) = -\ln\left(\frac{\mathbb{Q}^\theta(s_1)\mathbb{Q}^\theta(s_2|\theta, s^1)}{\gamma_2}\right),$$

where $\gamma_2 := \sup_{\theta \in \Theta} \mathbb{Q}^\theta(s_1)\mathbb{Q}^\theta(s_2|\theta, s^1)$.

Definition 4.4.2. *We say that the penalty $(\alpha_t)_t$ in the robust representation of the convex dynamic risk measure $(\rho_t)_t$ is achieved by* simplistic learning*, if it is of the form:*

$$\alpha_t(\theta) := -\ln\left(\frac{\prod_{i=1}^t \mathbb{Q}^\theta(s_i|\theta, s^{i-1})}{\gamma_t}\right),$$

where $\gamma_t := \sup_{\theta \in \Theta} \prod_{i=1}^t \mathbb{Q}^\theta(s_i|\theta, s^{i-1})$.

Remark 4.4.3 (On improperness of simplistic learning)**.** *$(\alpha_t)_t$ achieved by simplistic leaning is* not *a feasible penalty function.*

Proof. A penalty at t shall include the conditional distributions from t onwards as seen in the definition. In our likelihood approach α_t only depends on distributions up to time t, i.e. already realized entities of the density process. □

4.4.4 A Second, More Sophisticated Approach: Entropic Learning

We now incorporate the likelihood function in the relative entropy in order to achieve a risk measure based on the well known and elegant entropic risk measures.

Here, 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 $\theta^t := (\theta_1, \dots, \theta_t)$ analogous to s^t.

In the foregoing section, we have seen the major problem to be that our "penalty" was only contingent on the past evolution of the density process. There is however a whole bunch of possibilities to estimate the future by use of past information. A prominent route is by virtue of maximum likelihood estimator.

Definition 4.4.4 (Experience Based Learning)**.** (a) *Given likelihood l. Being at time t, learning is said to be* naive *if the estimator $\hat{\theta}_t$ for θ_t is achieved solely by taking into account maximum likelihood for the observation s_t at time t.*
(b) *Learning is called* intermediate *or* experience based at level m, *if $\hat{\theta}_t$ is the maximum likelihood estimator of the last m observations $(s_{t-m}, \dots, s_t)$*

$$\mathrm{MLE}_{-m} \in \arg\max_{\theta_t \in \Theta} l(s_{t-m}, \dots, s_t | \theta_t, \hat{\theta}^{t-1}, s^{t-m-1}).$$

(c) *Learning is said to be of* maximum likelihood *type, if, at any t, $\hat{\theta}_t$ is the maximum likelihood estimator of the whole history.*

Note that the naive estimator is just the intermediate one at level zero. Furthermore, note that our definition of experience based maximum likelihood. In the next definition, we characterize how learning results in a distribution for the payoff.

Definition 4.4.5 (Learning Distributions)**.** *Being at time t, having obtained $\hat{\theta}_t$ and the foregoing estimators $(\hat{\theta}_i)_{i<t}$, the reference family $\hat{\theta}$ of parameters is achieved by*

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

Having seen how agents learn about the best fitting distribution, we now formally introduce entropic learning for wich dynamic entropic risk measures in Definition 4.3.10 serve as a vehicle: We choose the best fitting distribution as reference distribution in the conditional relative entropy.

The agent's variational utility incorporating learning is in our setup given by a convex risk measure with an entropic penalty function:

Definition 4.4.6 (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 $\delta > 0$ and $\hat{\theta} = (\hat{\theta}_t)_t$ achieved as in Definition 4.4.5, $\eta = (\eta_t)_t \in \Theta$. The resulting convex risk measure $(\hat{\rho}_t)_t$ incorporating this very penalty function is then called experience based entropic risk*.*

Remark 4.4.7. *$(\hat{\alpha}_t^\theta)_t$ is well defined as penalty; this is inter alia shown in [Föllmer & Schied, 04]. Due to our construction, the penalty now incorporates conditional distributions of future movements.*

Remark 4.4.8. *When the parameter is also the realization of an entity in the density process, e.g. in a tree (cp. the example below), relative entropy can directly be written as*

$$\hat{\alpha}_t(\theta) = \mathbb{E}^{\mathbb{P}^\theta}\left[\ln \left(\frac{d\mathbb{P}^\theta}{d\mathbb{P}^{\theta_0}} \Big/ \frac{d\mathbb{P}^{\hat{\theta}}}{d\mathbb{P}^{\theta_0}} \right) \Bigg| \mathcal{F}_t \right].$$

Remark 4.4.9. *Naive entropic learning reflects the tendency of the agent to forget (or ignore) about the distant past and just assume the present to be the* *best estimator of the underlying model. This learning mechanism is then of course particularly adjuvant in explaining a bubble as it is harder to see that the financial system moves away from the fundamentals.*

Despite [Epstein & Schneider, 07] we do not consider multiplicity of likelihoods here. Hence, we do not incorporate information that cannot be learned upon in our model. Though real world applications with several true parameters, e.g. in incomplete financial markets with a multiplicity of equivalent martingale measures, would be modeled in terms of multiple likelihoods. However, our main result in this section on "time-*in*consistency" of experience based entropic risk would not change when extending the model to multiple likelihoods.

Proposition 4.4.10. *The model is well defined, i.e. for every t, $\hat{\rho}_t$ is a conditional convex risk measure.*

Proof. As can easily be seen, the model satisfies the axioms of convex risk measures: $\hat{\rho}_t : L^\infty \to L_t^\infty$ and

- $\hat{\rho}_t$ is monotone, i.e. $\hat{\rho}_t(X) \leq \hat{\rho}_t(Y)$ for $X \geq Y$ a.s.
- $\hat{\rho}_t$ is cash-invariant, i.e. $\hat{\rho}_t(X + m) = \hat{\rho}_t(X) - m \; \forall m \in L_t$, $X \in L_T$
- $\hat{\rho}_t$ is convex as a function on L_T

□

As inter alia shown [Föllmer & Penner, 06], Proposition 4.4, dynamic entropic risk measures are time-consistent when the reference distribution is not learned but fixed at the beginning. However, now that the reference distribution is also stochastic, we achieve:

Proposition 4.4.11. *Experience based entropic risk is in general not time-consistent.*

Proof. As proof we construct the following counterexample showing an experience based entropic risk measure which is not time-consistent. □

Example 4.4.12 (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 $l(\cdot \mid \theta)$. For the same reason we will also use the extreme case of naive updating which means our reference measure will merely depend on the last observed event in our tree. 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*

$$\hat{\rho}_t(X) = \hat{\rho}_t(-\hat{\rho}_{t+s}(X)) \qquad \textit{for all } t \in [0, T] \textit{ and } s \in \mathbb{N}$$

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

$$\begin{aligned} &\hat{\rho}_2(X)(\text{up}, \text{up}) \\ &= \operatorname*{ess\,sup}_{p \in [a,b]} \mathbb{E}\left[-X \mid \mathcal{F}_2\right](\text{up}, \text{up}) - \mathbb{E}\left[\ln\left(\frac{\theta_2}{\theta_2^*}\right) \mid \mathcal{F}_2\right](\text{up}, \text{up}) \\ &= \sup_{p \in [a,b]} \left(-3p - 1 + p - p \ln \frac{p}{b} - (1-p) \ln\left(\frac{1-p}{1-b}\right)\right) \\ &= \ln\left(be^{-3} + (1-b)e^{-1}\right), \end{aligned}$$

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

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

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:*

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

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

Here one can nicely observe the extremeness of the naive learning approach. Even though the decision maker in these two calculations is located at the same vertex in the tree he has very different beliefs about the probability of going up or down which causes strong shifts in his risk conception.

In the case of going first down then up he clearly believes up will be more probable in the next step. This is visible in his choice of reference measure $\mathbb{P}^{\theta^*}$ *in the penalty function which he sets* b *for going up and* $1-b$ *for going down.*

In contrast to this the decision maker who has observed up and then down will put more weight on the probability of going down in the next step and therefore sets his reference measure a *for up and* $1-a$ *for down.*

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*

$$\begin{aligned}\hat{\rho}_1(X)(\mathrm{up}) &= \hat{\rho}_1(-\hat{\rho}_2(X))(\mathrm{up}) = \ln \mathbb{E}^{\mathbb{P}^{\theta^*}}[\exp(\hat{\rho}_2(X)) \mid \mathcal{F}_1](\mathrm{up}) \\ &= \ln\left(b\left(be^{-3}+(1-b)e^{-1}\right)+(1-b)\left(ae^{-1}+(1-a)e^{1}\right)\right) \\ &= \ln\left(b2e^{-3}+(a+b)(1-b)e^{-1}+(1-a)e^{1}\right).\end{aligned}$$

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:

$$\begin{aligned}\hat{\rho}_1(X)(up) &= \operatorname*{ess\,sup}_{p,q\in[a,b]} \mathbb{E}^{p,q}\left[-X \mid \mathcal{F}_1\right](up) - \mathbb{E}^{p,q}\left[\ln\left(\frac{\theta_1\theta_2}{\theta_1^*\theta_2^*}\right) \mid \mathcal{F}_1\right](up) \\ &= \ln\left(b2e^{-3}+2b(1-b)e^{-1}+(1-b)2e^{1}\right).\end{aligned}$$

This clearly is 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.

To emphasize the reason for these inconsistencies set $Z_t := \frac{d\mathbb{P}^{\theta_1}}{d\mathbb{P}^{\theta_2}}\Big|_{\mathcal{F}_t}$*, where* $\mathbb{P}^{\theta_i}$ *is the reference distribution the agent obtains at time i when looking at past realizations and then maximizing the respective likelihood function. Then for instance for $t = 1$ and $\omega = up$ we obtain:*

$$\begin{aligned}&\hat{\rho}_1(-\hat{\rho}_2(X-\ln\frac{Z_T}{Z_2}))(up) \\ &= \ln\left[\mathbb{E}^{\mathbb{P}^{\theta_1}}\left[\exp\left(\rho_2\left(X_3-\ln\frac{Z_3}{Z_2}\right)\right)\right] \mid \mathcal{F}_1\right](up) \\ &= \ln\left[b\mathbb{E}^{\mathbb{P}^{\theta_2}}\left[e^{-X}\frac{Z}{Z_2} \mid \mathcal{F}_2\right](up,up)\right. \\ &\qquad \left.+(1-b)\mathbb{E}^{\mathbb{P}^{\theta_2}}\left[e^{-X}\frac{Z}{Z_2} \mid \mathcal{F}_2\right](up,down)\right] \\ &= \ln\left[b\left(be^{-3}\frac{bbb}{bbb}\frac{bb}{bb}+(1-b)e^{-1}\frac{bb(1-b)}{bb(1-b)}\frac{bb}{bb}\right)\right.\end{aligned}$$

$$+ (1-b)\left(ae^{-1}\frac{b(1-b)b}{b(1-b)a}\frac{b(1-b)}{b(1-b)}\right.$$
$$\left.\left. +(1-a)e^{1}\frac{b(1-b)(1-b)}{b(1-b)(1-a)}\frac{b(1-b)}{b(1-b)}\right)\right]$$
$$= \ln\left[b2e^{-3} + 2b(1-b)e^{-1} + (1-b)2e^{1}\right] = \rho_1(X)(up),$$

which, if $\frac{Z_T}{Z_i} \neq 1$ *(generally true), clearly contradicts time-consistency.*

In this special case for example the measure $\mathbb{P}^{\theta_1}$ *corresponds to the measure assigning the probability* b *to up in every time period, whereas* $\mathbb{P}^{\theta_2}$ *is the measure assigning* b *to up in the first 2 time periods and* a *in the last. That is why e.g.* $Z_3(up, down, up) = \frac{b(1-b)b}{b(1-b)a}$ *and* $\frac{Z_3}{Z_2}(up, down, up) = \frac{b}{a}$.

4.4.5 Lack of Time Consistency

As we have seen in the foregoing paragraph our definition of experience based entropic risk does not result in a time-consistent dynamic convex risk measure. This insight is somewhat disappointing as time consistency is a prosperous vehicle to solve tangible problems. On the other hand, [Schied, 07] shows that a meaningful theory of convex risk can even be achieved in a not generally time-consistent setting.

We have to pose the following question: Does there exist any learning model for the reference distribution such that dynamic entropic risk becomes time-consistent?

Remark 4.4.13. *The major issue that might come into mind is the independence of the reference distribution of future histories. As we will see, this is basically the reason for the general impossibility result below. Furthermore, the worst-case distribution chosen by nature is heavily dependent on the reference distribution. As the latter one may change in a broad variety of manners, there is no good reason to expect nature to choose in a time-consistent way.*

Next, we pose the most general definition of learning in entropic set-ups.

Definition 4.4.14. *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, i.e. not deterministically fixed from scratch. We call the resulting dynamic convex risk measure* $(\tilde{\rho}^g_t)_t$ *defined by virtue of* $\tilde{\alpha}^g_t := \hat{H}_t(\cdot|(\tilde{\theta}_t)_t)$ *in the robust representation* general experience based entropic risk.

We see that our definition of experience based entropic risk satisfies the above definition as in that context learning takes place in terms of maximum likelihood.

Using this general definition of learning, we can show an impossibility result for time-consistency of general experience based entropic risk.

Proposition 4.4.15. *General experience based entropic risk* $(\tilde{\rho}^g_t)_t$ *is in general not time-consistent.*

Proof. Let $\tilde{\theta} = (\tilde{\theta}_1, \ldots)$ 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 $Z_{t+1} := \frac{d\mathbb{Q}^{{}^t\tilde{\theta}}}{d\mathbb{Q}^{{}^{t+1}\tilde{\theta}}}\Big|_{\mathcal{F}_{t+1}}$. Then, we have

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

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. □

The foregoing result immediately implies our main intuition for experience based entropic risk not being time-consistent though quite puzzling as entropic risk measures are broadly used as standard example for time-consistent convex risk.

Remark 4.4.16 (Main Intuition)**.** *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. Hence, an experience based entropic risk measure is actually a family of dynamic entropic risk measures and our definition of time-consistency is not even applicable.*

4.4.6 A Retrospective – In Between

In this section, we have stated a constructive approach to learning for convex risk measures. We have encountered several problems in doing that:

- In our first intuitive approach, we ran into problems in defining a penalty function not entirely contingent on the past evolution of the density process.
- In our second one, we ran into time-consistency problems.

In a way, in the next section, we put the cart before the horse: We just take the robust representation in terms of minimal penalty of time-consistent dynamic convex risk measures as given and ask ourselves what can be said about "learning" in that respect. We will show an equivalent to the fundamental Blackwell-Dubins Theorem for convex risk measures. As will be seen, this result will be equivalently satisfied whenever the true parameter is eventually learned upon as defined in the subsequent subsection. Our result states some kind of herding behavior as every market participant will eventually perceive risk in the same manner.

4.4.7 Learning for a given Time-Consistent Convex Risk Measure

We now want to encounter, whether we actually have to construct a learning mechanism or if learning is not already incorporated in some sense in the concept of a time-consistent convex risk measure.

Remark 4.4.17. *We have stated that the time-consistency problem encountered so far in learning models is due to the fact that penalties are not just random variables but random itself, i.e. also the functional form depends on the observations. This assumption in general contradicts time-consistency as we actually may achieve distinct risk measures at a particular point in time. However, the basis for learning is already incorporated in convex risk as the domain of penalty consists of bayesian updated distributions of the process.*

Let us hence assume a true underlying parameter $\theta_0 \in \Theta$ and the agent evaluates risk in terms of robust representation of time-consistent dynamic convex risk $(\rho_t)_t$ with *minimal* penalty $(\alpha_t^{\min})_t$. We then state the following definition:

Definition 4.4.18. *We say that θ_0 is* eventually learned upon *if*

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X|\mathcal{F}_t]\right| \to 0 \quad \mathbb{P}^{\theta_0} - a.e.$$

for $t \to \infty$.

Proposition 4.4.19. *The above definition is satisfied if and only if*

$$\lim_{t\to\infty}\left|\rho_t(X) - \int_{S_{t+1}} -\rho_{t+1}(X)\mathbb{P}^{\theta_0}(ds_{t+1}|\mathcal{F}_t)\right| = 0 \quad \mathbb{P}^{\theta_0} - a.e.$$

Proof. cp. [Klibanoff et al., 09], Proposition 5. □

In the time-consistent case, the following assertion is equivalent to Definition 4.4.18:

Proposition 4.4.20. *Given a time-consistent dynamic convex risk measure $(\rho_t)_t$, then θ_0 is eventually learned upon if and only if*

$$\alpha_t^{\min}(\theta) \overset{t\to\infty}{\longrightarrow} 0 \quad \mathbb{P}^{\theta_0} - a.e$$

for all θ such that $\alpha_0^{\min}(\theta) < 0$.

Proof. As $(\rho_t)_t$ is assumed to be time-consistent, it holds for all t

$$\rho_t = \rho_t(-\rho_{t+1})$$

or, more elaborately, for all X

$$\begin{aligned} & \rho_t(X) \\ = \; & \sup_{\theta\in\Theta} \left\{ \mathbb{E}^{\mathbb{P}^\theta}\left[-X \middle| \mathcal{F}_t \right] - \alpha_t^{\min}(\theta) \right\} \\ = \; & \sup_{\theta\in\Theta} \left\{ \mathbb{E}^{\mathbb{P}^\theta}\left[-\rho_{t+1}(X) \middle| \mathcal{F}_t \right] - \alpha_t^{\min}(\theta) \right\}. \end{aligned}$$

As further for all X

$$\begin{aligned} & \int_{S_{t+1}} -\rho_{t+1}(X)\mathbb{P}^{\theta_0}(ds_{t+1}|\mathcal{F}_t) \\ = \; & \mathbb{E}^{\mathbb{P}^{\theta_0}}\left[-\rho_{t+1}(X) \middle| \mathcal{F}_t \right] \\ = \; & \sup_{\theta\in\Theta} \left\{ \mathbb{E}^{\mathbb{P}^\theta}\left[-\rho_{t+1}(X) \middle| \mathcal{F}_t \right] - \bar{\alpha}_t^{\min}(\theta) \right\}, \end{aligned}$$

where $(\bar{\alpha}_t^{\min})_t$ is defined as

$$\bar{\alpha}_t^{\min}(\theta) := \begin{cases} 0 & \text{if } \theta = \theta_0 \\ \infty & \text{else,} \end{cases}$$

the proof follows readily: $\alpha_t^{\min}(\theta) \overset{t\to\infty}{\longrightarrow} \bar{\alpha}_t^{\min}(\theta)$ by Proposition 4.4.19. Theorem 5.4.(4) in [Föllmer & Penner, 06] then shows equivalence to a vanishing limit given time-consistency. □

In the subsequent sections, we show the notion of being eventually learned upon to be satisfied by convex risk measures in case of time-consistency and under less stringent assumptions in terms of Blackwell & Dubins.

4.5 Adaption of Blackwell-Dubins Theorem

As a cornerstone for our main result on convergence of 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. As set out in the model section, we assume existence of a reference distribution $\mathbb{P}^{\theta_0}$, $\theta \in \Theta$, as in [Blackwell & Dubins, 62]. This reference has to be interpersonally being agreed upon.

Proposition 4.5.1. *Let $\mathbb{P}^{\theta}$ be absolutely continuous with respect to $\mathbb{P}^{\theta_0}$ for some $\theta \in \Theta$,*[1] *X as in the definition of the model, then*

$$\left|\mathbb{E}^{\mathbb{P}^{\theta}}[X \ |\mathcal{F}_t] - \mathbb{E}^{\mathbb{P}^{\theta_0}}[X \ |\mathcal{F}_t]\right| \to 0 \quad \mathbb{P}^{\theta_0}\text{-almost surely for } t \to \infty.$$

Proof. For improving readability denote $\mathbb{P}^{\theta_0}$ by $\mathbb{P}$ and $\mathbb{P}^{\theta}$ by $\mathbb{Q}$.

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

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

and hence

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

which converges to zero $\mathbb{P}$-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. □

Remark 4.5.2. *As we see in the proof, the parametric setting is not needed. The assertion can be shown in the same fashion in a non-parametric setting. The same holds true for subsequent results.*

[1] Note that we have assumed all distributions induced by parameters $\theta \in \Theta$ to be equivalent. In particular, all those are absolutely continuous with respect to each other and this assumption is no restriction within our setup. Also note that the respective θ does not have to be θ_0.

4.6 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.

4.6.1 Time-Consistent Coherent Risk

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

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

with $\tilde{\Theta} \subset \Theta$ assumed to be a convex and compact set of parameters inducing a weakly compact and convex set of priors $\tilde{\mathcal{Q}} \subset \mathcal{M}^e(\mathbb{P}^{\theta_0})$.

Proposition 4.6.1. *For every essentially bounded $\mathcal{F}$-measurable random variable X and time-consistent coherent risk measure $(\rho_t)_t$ we have*

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X \ |\mathcal{F}_t]\right| \to 0 \quad \mathbb{P}^{\theta_0}\text{-almost surely for } t \to \infty.$$

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

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X \ |\mathcal{F}_t]\right| = \left|\mathbb{E}^{\mathbb{P}^{\theta^*}}[-X \ |\mathcal{F}_t] - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X \ |\mathcal{F}_t]\right|$$

and this converges to zero as t increases and $\mathbb{P}^{\theta^*} \sim \mathbb{P}^{\theta_0}$ by Proposition 4.5.1. □

Remark 4.6.2. *Note that we have not assumed $\theta_0 \in \tilde{\Theta}$.*

Remark 4.6.3. *The assumption that $\tilde{\Theta}$ is weakly compact is a very crucial assumption, 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 4.6.4.*

Proposition 4.6.4. *Weak compactness of the set* $\{\mathbb{P}^\theta | \theta \in \tilde{\Theta}\}$ *of priors is a necessary condition for our result in Proposition 4.6.1 to hold.*

Proof. For the proof, see the counterexample in Section 4.8.2. □

4.6.2 Time-Consistent Convex Risk

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

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

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

Assumption 4.6.5. *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 4.6.6. *In the coherent case, continuity from below is equivalent to weak compactness of the set* $\{\mathbb{P}^\theta | (\alpha_t(\theta))_t = 0\} = \{\mathbb{P}^\theta | \theta \in \tilde{\Theta}\}$ *of priors as inter alia shown in [Riedel, 09].*

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

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X \; |\mathcal{F}_t]\right| \to 0 \quad \mathbb{P}^{\theta_0}\text{-almost surely for } t \to \infty$$

if there exists $\theta \in \Theta$ such that $\alpha_t^{\min}(\theta) \to 0$ $\mathbb{P}^{\theta_0}$-almost surely and $\alpha_0^{\min}(\theta) < \infty$.

Remark 4.6.8 (On the Assumption)**.** *By the main assumption in Proposition 4.6.7 there ought to be some θ 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 4.6.9. *By Theorem* 5.4 *in [Föllmer & Penner, 06] due to time-consistency the assumption $\alpha_t^{\min}(\theta) \to 0$ $\mathbb{P}^{\theta_0}$-almost surely for some $\theta \in \Theta$ is equivalent to $\alpha_t^{\min}(\theta) \to 0$ $\mathbb{P}^{\theta_0}$-almost surely for all $\theta \in \Theta$ with $\alpha_0(\theta) < \infty$.*

Proof of the proposition. By our assumptions on $(\rho_t)_t$ there exists $\theta^* \in \Theta$ such that the assertion becomes

$$\left|\mathbb{E}^{\mathbb{P}^{\theta^*}}[-X|\mathcal{F}_t] - \alpha_t^{\min}(\theta^*) - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X|\mathcal{F}_t]\right| \to 0 \quad \mathbb{P}^{\theta_0}\text{-a.s.}$$

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

$$\left|\alpha_t^{\min}(\theta^*)\right| \to 0 \quad \mathbb{P}^{\theta_0}\text{-a.s.}$$

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

$$\left|\alpha_t^{\min}(\theta)\right| \to 0 \quad \mathbb{P}^{\theta_0}\text{-a.s.}$$

for some $\theta \in \Theta$ such that $\alpha_0(\theta) < \infty$ as assumed to hold in the assertion. □

Corollary 4.6.10. *By Proposition 4.4.20 under the conditions of Proposition 4.6.7, θ_0 is eventually learned upon.*

Again, note that we have not assumed θ_0 such that $\alpha_0(\theta_0) < \infty$.

Corollary 4.6.11. *Every dynamic time-consistent convex risk measure $(\rho_t)_t$ satisfying the assumptions of the Proposition 4.6.7 is asymptotically precise as in the sense of [Föllmer & Penner, 06], i.e. $\rho_t(X) \to \rho_\infty(X) = -X$, and vice versa. In particular, this holds for the coherent case as $t \to \infty$.*

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 $\alpha_t^{\min}(\theta_0) \to 0$. Then the assertion is shown by Proposition 5.11 in [Föllmer & Penner, 06]. □

Remark 4.6.12. *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, $\lim_{t\to\infty}(E^{\mathbb{P}^{\theta_0}}[-X\ |\mathcal{F}_t])$ 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 4.6.13. *$(\rho_t)_t$ being continuous from below is a necessary condition for the result in Theorem 4.6.7 to hold.*

Proof. In Proposition 4.6.4 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 4.6.14. *In Proposition 4.6.7, if there does not exist θ such that $\alpha_t^{\min}(\theta) \to 0$ but $\alpha_t^{\min}(\theta^*) \leq c \in \mathbb{R}_+$ for all $t \geq n_0$ for some $n_0 \in \mathbb{N}$ then there is at least an upper bound on the remaining uncertainty:*

$$|\rho_t(X) - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X|\mathcal{F}_t]| \leq c$$

as $t \to \infty$.

4.7 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 parameter θ_0.

4.7.1 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 4.8.3.

Definition 4.7.1. (a) *Given a dynamic convex risk measure $(\rho_t)_t$, continuous from below but not necessarily time-consistent, we call a distribution $\mathbb{P}^{\theta_t^*}$* instantaneous worst case distribution at t *if it satisfies*[2]

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

(b) *We say* learning takes place *if there exists a $\theta \in \Theta$, $\mathbb{P}^{\theta} \sim \mathbb{P}^{\theta_0}$, such that the instantaneous worst case measures $\mathbb{P}^{\theta_t^*} \to \mathbb{P}^{\theta}$ weakly for $t \to \infty$. In the coherent case we need $\theta \in \tilde{\Theta}$ as the penalty is infinite otherwise.*

In this very definition, we see however, that the agent does not have to learn the true underlying parameter θ_0. In this sense, nature might mislead her to a wrong parameter.

[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.

We can now relax the time-consistency assumption in the main result of this article. Note that time-consistency is a special case of Definition 4.7.1 given continuity from below as in that case the sequence of instantaneous worst case parameters is constant. Hence, we achieve the more general result:

Proposition 4.7.2. *Let $(\rho_t)_t$ be a not necessarily time-consistent dynamic coherent risk measure for which learning takes place. Then*

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X \,|\mathcal{F}_t]\right| \to 0 \quad \mathbb{P}^{\theta_0}\text{-almost surely for } t \to \infty.$$

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

$$\alpha_t^{\min}(\theta) = \begin{cases} 0 & \text{if } \mathbb{P}^\theta(\cdot|\mathcal{F}_t) \in \tilde{\mathcal{Q}}(\cdot|\mathcal{F}_t), \\ \infty & \text{else} \end{cases}$$

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

First, note that in case $\alpha_t^{\min}(\theta) \to \infty$ for all $\theta \in \tilde{\Theta}$[3], our convergence result cannot hold, as $\lim_{t\to\infty} \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X|\mathcal{F}_t]$ exists and is finite by assumption.

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

Let us now turn to the proof itself: As $\tilde{\mathcal{Q}}$ is assumed to be weakly compact and non-empty, i.e. there exists a distribution that has penalty zero, we achieve an instantaneous worst case distribution at each time step, i.e. at any t, there exists $\theta_t^* \in \Theta$ s.t.

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

Of course, due to "time-*in*consistency", we might have $\theta_i^* \neq \theta_j^*$ for $i \neq j$.

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

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

$$\mathbb{E}^{\mathbb{P}^{\theta_n^*}}[-X|\mathcal{F}_t] \to \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X|\mathcal{F}_\infty] \qquad \text{for } n,t \to \infty.$$

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:

$$\mathbb{E}^{\mathbb{P}^{\theta_n^*}}[-X|\mathcal{F}_t] = \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X\frac{d\mathbb{P}^{\theta_n^*}}{d\mathbb{P}^{\theta_0}}\Big|_{\mathcal{F}_n}|\mathcal{F}_t].$$

Define the following sequence of random variables $Y_n := -X\frac{d\mathbb{P}^{\theta_n^*}}{d\mathbb{P}^{\theta_0}}\Big|_{\mathcal{F}_n}$. These have finite expectation and thanks to our assumption that learning takes place and the original Blackwell-Dubins result we have

$$\mathbb{P}^{\theta_0}[\lim_{n\to\infty} Y_n = -X] = \mathbb{P}^{\theta_0}[-X\frac{d\mathbb{P}^{\theta_\infty^*}}{d\mathbb{P}^{\theta_0}}\Big|_{\mathcal{F}_\infty} = -X] = 1.$$

Then, by Lemma 4.7.4, the assertion follows. □

Remark 4.7.3. *Again, note that we have not assumed* $\theta_0 \in \tilde{\Theta}$.

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

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

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

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

A positive answer is given in the following lemma.

[4] By our assumptions we know:

- $\mathbb{E}^{\mathbb{P}^{\theta_n^*}}[-X|\mathcal{F}_t] \to \mathbb{E}^{\mathbb{P}^\theta}[-X|\mathcal{F}_t]$ for $n \to \infty$ as $\theta_n^* \to \theta$ by Portemonteau's Theorem.
- $\mathbb{E}^{\mathbb{P}^{\theta_n^*}}[-X|\mathcal{F}_t] \to \mathbb{E}^{\mathbb{P}^{\theta_n^*}}[-X|\mathcal{F}_\infty]$ for $t \to \infty$ by Proposition 4.5.1.

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

In the time-consistent case, where $\theta_i^* = \theta_j^*$ for all i, j, this is immediate by Proposition 4.5.1.

Lemma 4.7.4. *Fix θ. Let $(Y_n)_n$ be a sequence of $\mathcal{F}$-measurable random variables such that $\mathbb{E}^{\mathbb{P}^\theta}[\sup_n |Y_n|] < \infty$. Assume $Y_n \to_{n\to\infty} Y$ almost surely for some $\mathcal{F}$-measurable random variable Y. Then, it holds*[5]

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

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

$$\mathbb{E}^{\mathbb{P}^\theta}\left[Y_n \middle| \mathcal{F}_t\right] \leq \mathbb{E}^{\mathbb{P}^\theta}\left[G_k \middle| \mathcal{F}_t\right] \tag{4.3}$$

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

$$\begin{aligned} z &:= \lim_{j\to\infty} \sup_{n,t\geq j} \mathbb{E}^{\mathbb{P}^\theta}\left[Y_n \middle| \mathcal{F}_t\right] \\ &\overset{(4.3)}{\leq} \lim_{j\to\infty} \sup_{t\geq j} \mathbb{E}^{\mathbb{P}^\theta}\left[G_k \middle| \mathcal{F}_t\right] \\ &= \lim_{t\to\infty} \mathbb{E}^{\mathbb{P}^\theta}\left[G_k \middle| \mathcal{F}_t\right] \\ &\overset{\text{Doob}}{=} \mathbb{E}^{\mathbb{P}^\theta}\left[G_k \middle| \mathcal{F}\right] \end{aligned}$$

and

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

In the same token,

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

which completes the proof since

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

□

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

Remark 4.7.5 (On Blackwell-Dubins Type Learning). *Blackwell-Dubins applies for learning models but does not necessarily result in time-consistency as this notion is now motivated as a special case of our notion of* θ_0 *to be eventually learned upon.*

We have built a bridge between the first and the second part of this article: in the first part we have achieved dynamic convex risk measures by virtue of learning that did not turn out to be time-consistent. Hence, we have shown, that our result even holds for those models, e.g. entropic learning.

Remark 4.7.6. *Note, that the above new version of the fundamental result particularly holds for time-consistent dynamic coherent risk measures as then such a limiting* θ *as in the Definition 4.7.1(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].*

4.7.2 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 $\theta \in \Theta$ such that the instantaneous worst case $\theta_t^* \to \theta$ as $t \to \infty$. Furthermore, we have to assume $\alpha_t^{\min}(\theta_t^*) \to 0$ as $n \to \infty$:[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 θ_0.

Proposition 4.7.7. *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*

$$\left|\rho_t(X) - \mathbb{E}^{\mathbb{P}^{\theta_0}}[-X \,|\mathcal{F}_t]\right| \to 0 \quad \mathbb{P}^{\theta_0}\text{-almost surely for } t \to \infty$$

[6] Note, again, we do not have to assume $\alpha_t^{\min}(\theta_0) \to 0$.

if learning takes place for an instantaneous worst case sequence $(\theta_t^*)_t$ *toward some* $\theta \in \Theta$ *and we have*

$$\alpha_t^{\min}(\theta_t^*) \to 0.$$

Proof. Applying the procedure used in the proof of Proposition 4.7.2 to the proof of Proposition 4.6.7 shows the assertion. □

4.8 Examples

In this section, we first consider dynamic entropic risk measures as a prominent economic example of time-consistent dynamic convex risk measures. In the second part we state a counterexample serving as proof for Proposition 4.6.4 and 4.6.13. Lastly, we consider a dynamic risk measure that is not time-consistent.

4.8.1 Entropic Risk

Here, we will have a look at time-consistent dynamic entropic risk measure $(\rho_t^{\mathrm{e}})_t$. Recall its Definition 4.3.10 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

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

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

The fundamental Blackwell-Dubins Theorem immediately shows that

$$\left|\mathbb{P}^\theta(\cdot|\mathcal{F}_t) - \mathbb{P}^\eta(\cdot|\mathcal{F}_t)\right| \to 0$$

for every θ, η. Hence, we have that $\frac{Z_T}{Z_t} \to 1$ $\mathbb{P}^{\theta_0}$-a.s. for $t \to \infty$and hence

$$\alpha_t^{\min}(\theta) \to 0$$

showing Proposition 4.6.7 to hold. This is an alternative way to show the last assertion in Theorem 6.3 in [Föllmer & Penner, 06] directly.

4.8.2 Counterexample

To show necessity of continuity from below in Proposition 4.6.7 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 $\mathbb{P}^{\theta_0}$ 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 $\alpha_t^{\min}(\theta_0) \to 0$ $\mathbb{P}^{\theta_0}$-a.s. of the following form:

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

That this sequence from all properties assumed in Proposition 4.6.7 is only missing continuity from below (here equivalent to weak compactness 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 $\mathbb{P}^{\theta_0}[A] > 0$ and $\mathbb{P}^{\theta_0}[A^c \cap J_{t,k}] \neq 0$ for all t and k. For this set, it holds

$$\lim_{t\to\infty}\left|\rho_t(\mathbb{1}_A)-\mathbb{E}^{\mathbb{P}^{\theta_0}}[-\mathbb{1}_A\,|\mathcal{F}_t]\right|=\lim_{t\to\infty}\left|0+\mathbb{P}^{\theta_0}[A\,|\mathcal{F}_t]\right|=\mathbb{P}^{\theta_0}[A]>0$$

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

$$A:=\left(\bigcup_{t=1}^{\infty}\bigcup_{k=1}^{2^t-1}U_{\epsilon_t}(k2^{-t})\right)^c,$$

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

4.8.3 A Non Time-Consistent Example

Here, we consider the entropic learning model introduced in Definition 4.4.6 explicitly in terms of $\Omega=\otimes_t S_t$. Let $\mathbb{P}^\theta$ denote the distribution induced by $\theta=(\theta_t)_t$, θ_t inducing a marginal distribution in $\mathcal{M}(S_t)$. Though the model looks quite similar to dynamic entropic risk measures, we briefly recall it: Let the robust representation of a dynamic convex risk measure $(\hat\rho_t)_t$ be given by virtue of the penalty

$$\hat\alpha_t^{\min}(\theta):=\delta\hat H_t(\mathbb{P}^\theta|\mathbb{P}^{\hat\theta}),$$

$\delta>0$ and $\hat\theta=(\hat\theta_t)_t$ be achieved as in Definition 4.4.5: for $t\in\mathbb{N}$, $\hat\theta_t$ is the maximum likelihood estimator of the foregoing observations and $\hat\theta_i:=\hat\theta_t$ for $i>t$. Restricting ourselves to the iid case, we know that we achieve $\hat\theta_t\to\bar\theta_0$, $\mathbb{P}^{\theta_0}$-a.ss, where $\theta_0=(\bar\theta_0)_t$ for some $\bar\theta_0$ inducing a marginal distribution in $\mathcal{M}(S_t)$. By definition, $(\hat\rho_t)_t$ is a dynamic convex risk measure. As shown in Proposition 4.4.15, $(\hat\rho_t)_t$ is not time-consistent. By standard results on conditional entropic risk measures, $(\hat\rho_t)_t$ is continuous from below.

Furthermore, Proposition 4.7.7 is applicable and hence, our generalization of Blackwell-Dubins' theorem holds for experience based entropic risk.

Indeed: By definition of the penalty and our considerations in Section 4.8.1, $\hat{\alpha}_t^{\min}(\theta) \to 0$ as $t \to \infty$ for all $\theta \in \Theta$. Secondly, as the maximum likelihood estimator is asymptotically stable, i.e. $\hat{\theta}_t \to \bar{\theta}_0$, the conditional reference distributions $\mathbb{P}^{\hat{\theta}}(\cdot|\mathcal{F}_t)$ converge. Thus, the worst case instantaneous distributions $\mathbb{P}^{\theta_t^*}$ converge as in Definition 4.7.1 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]

4.9 Conclusions

The major contribution of our results is to carry over the famous Blackwell-Dubins theorem from probability distributions to convex risk measures. It is particularly striking that the results still hold when time-consistency is not posed as an assumption.

Hereto, the present article is twofold: In the first part, we show that explicitly constructing dynamic convex risk measures by virtue of a penalty emerging from a learning mechanism and inserted in the robust representation of convex risk measures leads to time-consistency problems. In the second part, we have then assumed a time-consistent dynamic convex risk measure for granted and asked the question of limit behavior; more elaborately its convergence to the expected value under the true underlying distribution.

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

[7] The notation is quite misleading at this point: the worst case instantaneous distributions $\mathbb{P}^{\theta_t^*} \in \mathcal{M}^e(\mathbb{P}^{\theta_0})$ as in Definition 4.7.1 is a distribution on $(\Omega, \mathcal{F})$ as θ_t^* is an element of Θ and not a "marginal" parameter as the above θ_ts.

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 result to hold. In particular, we have obtained a more general existence result for the limiting risk measure ρ_∞ than in [Föllmer & Penner, 06]. Our generalization of the Blackwell-Dubins theorem was shown to be equivalent to the notion of the parameter being eventually learned upon and the notion of asymptotic precision in [Föllmer & Penner, 06] in the time-consistent case.

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, 09]. Hence, the main question turns out to be if the result still holds assuming local instead of global equivalence as done here.

Chapter 5

Closing Remarks

Within the three essays of this thesis we have tackled several problems arising in case of dynamic coherent as well as convex risk measures or, equivalently, dynamic variational preferences. Each essay is elaborately given in one chapter and finalized by a conclusion stating achievements of that essay's results as well as limitations and ideas for further research. Nevertheless, we briefly summarize our results here at the very end:

First, we have generalized the Best-Choice or Secretary problem to the case of an ambiguous number of applicants. For this problem we have achieved a result on the number of stopping islands generalizing the main theorem in [Presman & Sonin, 72]. In order to achieve this, we have encountered several problems in directly generalizing the risky to the ambiguous problem and hence have built a model in terms of assessments.

Thereafter, we have build a general theory for optimal stopping of payoff processes in context of time-consistent dynamic variational preferences. In order to achieve our results on optimal stopping times by virtue of so called variational Snell envelopes extending [Riedel, 09], we have introduced the notion of variational supermartingales and have built an accompanying martingale theory. We have applied our insights to dynamic entropic risk measures and average value at risk.

In the third article, we have considered dynamic convex risk measures when information is gathered in course of time. We have generalized the fundamental Blackwell-Dubins theorem from [Blackwell & Dubins, 62] to not necessarily time-consistent dynamic convex risk measures and have thus shown their convergence to conditional expected values with respect to the true underlying distribution: Intuitively the result shows that uncertainty vanishes but risk endures.

Bibliography

[Abdel-Hamid et al., 82] Abdel-Hamid, A. R.; Bather, J. A. & Trustrum, G. B.: *The Secretary Problem with an Unknown Number of Candidates*, Journal of Applied Probability, Vol.19, No.3, pp.619-630, 1982.

[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.

[Artzner et al., 07] Artzner, Philippe; Delbaen, Freddy; Eber, Jean-Marc; Heath, David & Ku, Hyejin: *Coherent Multiperiod Risk Adjusted Values and Bellman's Principle*, Annals of Operations Research, Vol.152, No1, pp.5-22, 2007.

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

[Barrieu & El Karoui, 04] Barrieu, Pauline & El Karoui, Nicole: *Optimal Derivatives Design under Dynamic Risk Measures*, Contemporary Mathematics, Vol.351, pp.13-26, 2004.

[Barrieu & El Karoui, 05] Barrieu, Pauline & El Karoui, Nicole: *Inf-Convolution of Risk Measures and Optimal Risk Transfer*, Finance and Stochastics, Vol.9, pp.269-298, 2005.

[Bion-Nadal, 06] Bion-Nadal, Jocelyne: *Dynamic Risk Measuring: Discrete Time in a Context of Uncertainty, and Continuous Time on a Probability Space*, CMAP Preprint No.596, 2006.

[Bion-Nadal, 06b] Bion-Nadal, Jocelyne: *Time Consistent Dynamic Risk Processes, Cadlag Modification*, eprint arXiv:math/0607212, 2006.

[Bion-Nadal, 08] Bion-Nadal, Jocelyne: *Dynamic Risk Measures: Time Consistency and Risk Measures from BMO Martingales*, Finance and Stochastics, Vol.12, No.2, pp.219-244, 2008.

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

[Bruss, 84] Bruss, F. Thomas: *A Unified Approach to a Class of Best Choice Problems with an Unknown Number of Options*, Annals of Probability, Vol.12, No.3, pp.882-889, 1984.

[Bruss & Samuels, 87] Bruss, F. Thomas & Samuels, Stephen M.: *A Unified Approach to a Class of Optimal Selection Problems with an Unknown Number of Options*, Annals of Probability, Vol.15, No.2, pp.824-8830, 1987.

[Bruss & Rogers, 91] Bruss, F. Thomas & Rogers, L. Chris G.: *Embedding Optimal Selection Problems in a Poisson Process*, Stochastic Processes and their Applications, Vol.38.2, pp.267-278, 1991.

[Bruss, 00] Bruss, F. Thomas: *Sum the Odds to One and Stop*, Annals of Probability, Vol.28, No.3, pp.1384-1391, 2000.

[Carr et al., 01] Carr, Peter; Geman, Helyette & Madan, Dilip B.: *Pricing and Hedging in Incomplete Markets*, Journal of Financial Economics, Vol.62, pp.131-167, 2001.

[Chateauneuf, 1991] Chateauneuf, Alain: *On the Use of Capacities in Modeling Uncertainty Aversion and Risk Aversion*, Journal of Mathematical Economics, Vol.20, pp.343-369, 1991.

[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.

[Chen & Epstein, 02] Chen, Zengjing & Epstein, Larry: *Ambiguity, Risk, and Asset Returns in Continuous Time*, Econometrica, Vol.70, No.4, pp.1403-1443, 2002.

[Cheridito et al, 04] Cheridito, Patrick; Delbaen Freddy & Kupper, Michael: *Coherent and Convex Monetary Risk Measures for Bounded Cadlag-Processes*, Stochastic Processes and Their Applications, Vol.112, No.1, pp.1-22, 2004.

[Cheridito et al, 05] Cheridito, Patrick; Delbaen Freddy & Kupper, Michael: *Coherent and Convex Monetary Risk Measures for Unbounded Cadlag-Processes*, Finance and Stochastics, Vol.9, pp.369-387, 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.

[Cheridito & Li, 09] Cheridito, Patrick & Tianhui, Li: *Risk Measures on Orlicz Hearts*, Mathematical Finance, Vol.19, No.2, pp.189-214, 2009.

[Cheridito & Stadje, 09] Cheridito, Patrick & Stadje, Mitja: *Time-inconsistency of VaR and time-consistent alternatives*, Finance Research Letters, Vol.6, No.1, pp.40-46, 2009.

[Chow & Robbins, 63] Chow, Yuan S. & Robbins, Herbert: *On Optimal Stopping Rules*, Zeitschrift für Wahrscheinlichkeitstheorie, Vol.2, pp.33-49, 1963.

[Chow, 67] Chow, Yuan S.: *On the Expected Value of a Stopped Submartingale*, The Annals of Mathematical Statistics, Vol.38, No.2, pp.608-609, 1967.

[Chudjakow & Riedel, 09] Chudjakow, Tatjana & Riedel, Frank: *The Best Choice Problem under Ambiguity*, IMW Working Paper, 2009.

[Cont, 06] Cont, Rama: *Model Uncertainty and its Implication on the Pricing of Derivative Instruments*, Mathematical Finance, Vol.16, No.3, pp.519-547, 2006.

[Cozman & Chrisman, 97] Cozman, Fabio & Chrisman, Lonnie: *Learning Convex Sets of Probability from Data*, Working Paper, CMU RI TR 97 25, 1997.

[Csiszar, 75] Csiszar, I: *I-Divergence Geometry of Probability Distributions an Minimization Problems*, The Annals of Probability, Vol.3, No.1, pp.146-158, 1975.

[Cvitanic & Karatzas, 99] Cvitanic, Jaksa & Karatzas, Ioannis: *On Dynamic Measures of Risk*, Finance and Stochastics, Vol.3, No.4, pp.451-482, 1999.

[DeGroot, 68] DeGroot, Morris H.: *Some Problems of Optimal Stopping*, Journal of the Royal Statistical Society, Series B (Methodological), Vol.30, No.1, pp.108-122, 1968.

[Delbaen, 02] Delbaen, Freddy: *Coherent Risk Measures on General Probability Spaces*, Advances in Finance and Stochastics, pp.1-37, Springer-Verlag, Berlin, 2002.

[Delbaen et al., 02] Delbaen, Freddy; Grandidts, Peter; Rheinländer, Thorsten; Samperi, Dominick; Schweizer, Martin & Stricker, Christoph: *Exponential Hedging and Entropic Penalties*, Mathematical Finance, Vol.12, No.2, pp.99-123, 2002.

[Detlefsen & Scandolo, 05] Detlefsen, Kai & Scandolo, Giacomo: *Conditional and Dynamic Convex Risk Measures*, Finance and Stochastics, Vol.9, No.4, pp.539-561, 2005.

[Engen & Seim, 87] Engen, Steinar & Seim, Eva: *Uniformly Best Invariant Stopping Rules*, Journal of Applied Probability, Vol.24, No.11, pp.77-87, 1987.

[Epstein & Marinacci, 07] 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, No.1, pp.1-13, 2003.

[Epstein & Schneider, 03b] Epstein, Larry & Schneider, Martin: *IID: independently and indistinguishably distributed*, Journal of Economic Theory, Vol.113, No.1, pp.32-50, 2003.

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

[Epstein & Schneider, 08] Epstein, Larry & Schneider, Martin: *Ambiguity, Information Quality and Asset Pricing*, Journal of Finance, Vol.63, No.1, pp.197-228, 2008.

[Ferguson, 89] Ferguson, Thomas S.: *Who Solved the Secretary Problem?*, Statistical Science, Vol.4, No.3, pp.282-296, 1989.

[Föllmer & Schied, 02] Föllmer, Hans & Schied, Alexander: *Convex Measures of Risk and Trading Constraints*, Finance and Stochastics, Vol.6, No.4, pp.429-447, 2002.

[Föllmer & Schied, 02b] Föllmer, Hans & Schied, Alexander: *Robust Preferences and Convex Measures of Risk*, Advances in Finance and Stochastics, pp.39-56, Springer-Verlag, Berlin. 2002.

[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, No.1, 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.

[Freeman, 83] Freeman, P. R.: *The Secretary Problem and Its Extensions: A Review*, International Statistical Review, Vol.51, No.2, pp.189-206, 1983.

[Frittelli & Rosazza Gianin, 02] Frittelli, Marco & Roszza Gianin, Emanuela: *Putting Order in Risk Measures*, Journal of Banking and Finance, Vol.26, No.7, pp.1473-1468, 2002.

[Frittelli & Rosazza Gianin, 04] Frittelli, Marco & Roszza Gianin, Emanuela: *Dynamic Convex Risk Measures*, in *Risk Measures for the 21st Century*, G. Szegö ed., pp.227-248,Wiley & Sons, 2004.

[Frittelli & Rosazza Gianin, 05] Frittelli, Marco & Roszza Gianin, Emanuela: *Law Invariant Convex Risk Measures*, Advances in

Mathematical Economics, Vol.7, pp.42-53, Springer-Verlag, Berlin, 2005.

[Gianini-Pettitt, 79] Gianini-Pettitt, Jacqueline: *Optimal Selection Based on Relative Ranks with a Random Number of Individuals*, Advances in Applied Probability, Vol.11, No.4, pp.720-736, 1979.

[Gilbert & Mosteller, 66] Gilbert, John P. & Mosteller Frederick: *Recognizing the Maximum of a Sequence*, Journal of the American Statistical Association, Vol.61, No. 313, pp.35-73, 1966.

[Gilboa, 87] Gilboa, Itzhak: *Expected Utility with Purely Subjective Non-additive Probabilities*, Journal of Mathematical Economics, Vol.16, No.1, pp.65-88, 1987.

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

[Gnedin, 94] Gnedin, Alexander V.: *A Solution to the Game of Googol*, The Annals of Probability, Vol.22, No.3, pp.1588-1585, 1994.

[Gnedin, 96] Gnedin, Alexander V.: *On the Full-Information Best-Choice Problem*, Journal of Applied Probability, Vol.33, No.3, pp.678-687, 1996.

[Gnedin & Krengel, 95] Gnedin, Alexander V. & Krengel, Ulrich: *A Stochastic Game of Optimal Stopping and Order Selection*, The Annals of Probability, Vol.5, No.1, pp.310-321, 1995.

[Hill & Krengel, 91] Hill, Theodore P. & Krengel, Ulrich: *Minimax-Optimal Stop Rules and Distributions in Secretary Problems*, The Annals of Probability, Vol.19, No.1, pp.342-353, 1991.

[Hill & Kennedy, 92] Hill, Theodore P. & Kennedy, D.P.: *Sharp Inequalities for Optimal Stopping with Rewards based on Ranks*, The Annals of Probability, Vol.2, No.2, pp.503-517, 1992.

[Irle, 80] Irle, Albrecht: *On the Best Choice Problem with Random Population Size*, Zeitschrift für Operations Research, Vol.24, No.5, pp.177-190, 1980.

[Jouini et al., 04] Jouini, Elyes; Meddeb, Moncef & Touzi, Nizar: *Vector Valued Coherent Risk Measures*, Finance and Stochastics, Vol.8, No.4, pp.531-552, 2004.

[Kalai & Lehrer, 94] Kalai, Ehud & Lehrer, Ehud: *Weak and Strong Merging of Opinions*, Journal of Mathematical Economics, Vol. 23, No.1, pp.73-86, 1994.

[Karatzas & Zamfirescu, 06] Karatzas, Ioannis & Zamfirescu, Ingrid-Mona: *Martingale Approach to Stochastic Control with Discretionary Stopping*, Applied Mathematics & Optimization, Vol.53, No.2, pp.163-184, 2006.

[Karatzas & Zamfirescu, 08] Karatzas, Ioannis & Zamfirescu, Ingrid-Mona: *Martingale Approach to Stochastic Differential Games of Control and Stopping*, Annals of Probability, Vol.36, No.4, pp.1495-1527, 2008.

[Klibanoff et al., 09] Klibanoff, Peter; Marinacci, Massimo & Mukerji,Sujoy: *Recursive Smooth Ambiguity Preferences*, Journal of Economic Theory, forthcoming.

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

[Lindley, 61] Lindley, Dennis Victor: *Dynamic Programming and Decision Theory*, Applied Statistics, Vol.10, No.1, pp.39-51, 1961.

[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, No.1, pp. 4-44, 2006.

[Marinacci, 99] Marinacci, Massimo: *Limit Laws for Non-additive Probabilities and Their Frequentist Interpretation*, Journal of Economic Theory, Vol.84, No.2, pp.145-195, 1999.

[Marinacci, 02] Marinacci, Massimo: *Learning from Ambiguous Urns*, Statistical Papers, Vol.43, No.1, pp. 145-151, 2002.

[McNeil et al., 05] McNeil, Alexander; Frey, Rüdiger & Embrechts, Paul: *Quantitative risk management : concepts, techniques, and tools*, Princeton University Press, Princeton, 2005.

[Neveu, 75] Neveu, Jaques: *Discrete-Parameter Martingales*, North-Holland, Oxford, 1975.

[Peng, 97] Peng, S.: *Backward SDE and related g-expectations* in: El Karoui, Nicole & Mazliak, L (eds.).: *Backward Stochastic Differential Equations*, Pitman Research Notes Mathematical Series, Vol.364, pp.141-159, 1997.

[Petrucelli, 83] Petrucelli, Joseph D.: *On the Best-Choice Problem when the Number of Observations is Random*, Journal of Applied Probability, Vol.20, No.1, pp.165-171, 1983.

[Presman & Sonin, 72] Presman, Ernst L. & Sonin, Isaac M.: *The Best Choice Problem for a Random Number of Objects*, Theory of Probability and its Applications, Vol.17, No.4, pp.657-668, 1972.

[Rasmussen, 75] Rasmussen, Will T.: *A Generalized Choice Problem*, Journal of Optimization Theory and Applications, Vol.15, No.3, pp.311-325, 1975.

[Rasmussen & Robbins, 75] Rasmussen, Will T. & Robbins, Herbert: *The Candidate Problem with Unknown Population Size*, Journal of Applied Probability, Vol.12, No.4, pp.692-701, 1975.

[Riedel, 04] Riedel, Frank: *Dynamic Coherent Risk Measures*, Stochastic Processes and their Applications, Vol.112, No.2, pp.185-200, 2004.

[Riedel, 09] Riedel, Frank: *Optimal Stopping with Multiple Priors*, Econometrica, forthcoming.

[Rockafellar, 70] Rockafellar, R. Tyrell: *Convex Analysis*, Princeton University Press, Princeton, 1970.

[Roorda et al., 05] Roorda, Berend; Schumacher, J.M. & Engwerda, Jacob: *Coherent Acceptability Measures in Multiperiod Modesl*, Mathematical Finance, Vol.15, No.4, pp.589-612, 2005.

[Roorda & Schumacher, 07] Roorda, Berend & Schumacher, Hans: *Time Consistency Conditions for Acceptability Measures, with an Application to Tail Value at Risk*, Insurance: Mathematics and Economics, Vol.40, No.2, pp.209-230, 2007.

[Rosazza Gianin, 04] Roszza Gianin, Emanuela: *Some Examples of Risk Measures via g-Expectations*, Working Paper No. 41, University of Milano, 2004.

[Rosazza Gianin, 06] Roszza Gianin, Emanuela: *Risk Measures via g-Expectations*, Insurance: Mathematics and Economics, Vol.39, No.1, pp.19-34, 2006.

[Sakaguchi, 76] Sakaguchi, Masamichi: *Optimal Stopping Problems for Randomly Arriving Offers*, Math. Japonica, Vol.21, No.9, pp.201-217, 1976.

[Samuel-Cahn, 96] Samuel-Cahn, Ester: *Optimal Stopping With Random Horizon With Application to the Full-Information Best-Choice Problem With Random Freeze*, Journal of the American Statistical Association, Vol.91, No.433, pp.357-364, 1996.

[Samuel-Cahn, 95] Samuel-Cahn, Ester: *The Best-Choice Secretary Problem with Random Freeze on Jobs*, Stochastic Processes and Applications, Vol.55, No.2, pp.315-327, 1995.

[Sardelis & Valahas, 99] Sardelis, Dimitris A. & Valahas Theodoros M.: *Decision Making: A Golden Rule*, The American Mathematical Monthly, Vol.106, No.3, pp.215-226, 1999.

[Schied, 06] Schied, Alexander: *Risk Measures and Robust Optimization Problems*, Stochastic Models, Vol.22, No.4, pp.753-831, 2006.

[Schied, 07] Schied, Alexander: *Optimal Investment for Risk- and Ambiguity-Averse Preferences: A Duality Approach*, Finance and Stochastics, Vol.11, No.1, pp.107-129, 2007.

[Schmeidler, 86] Schmeidler, David: *Integral Representation without Additivity*, Proceedings of the American Mathematical Society, Vol.97, No.2, pp.255-261, 1986.

[Schmeidler, 89] Schmeidler, David: *Subjective Probability and Expected Utility without Additivity*, Econometrica, Vol.57, No.3, pp.571-587, 1989.

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

[Skulj, 01] Skulj, Damjan: *Non-additive Probability*, Faculty of Social Sciences, University of Ljubljana, Slovenia, 2001.

[Shepp, 69] Shepp, Larry A.: *Explicit Solutions to some Problems of Optimal Stopping*, The Annals of Mathematical Statistics, Vol.40, No.3, pp.993-1010, 1969.

[Siegmund, 67] Siegmund, David O.: *Some Problems in the Theory of Optimal Stopping Rules*, The Annals of Mathematical Statistics, Vol.38, No.6, pp.1627-1640, 1967.

[Stewart, 78] Stewart, Theodor J.: *Optimal Selection from a Random Sequence with Learning of the Underlying Distribution*, Journal of the American Statistical Association, Vol.73, No.364, pp.775-780, 1978.

[Stewart, 81] Stewart, Theodor J.: *The Secretary Problem with an Unknown Number of Options*, Operations Research, Vol.29, No.1, pp.130-145, 1981.

[Szegö, 04] Szegö, Giorgio (ed.): *Risk Measures for the 21st Century*, The Wiley Finance Series, 2004.

[Walley & Fine, 82] Walley, Peter & Fine, Terrence L.: *Towards a Frequentist Theory of Upper and Lower Probability*, The Annals of Statistics, Vol.10, No.3, pp.741-761, 1982.

[Weber, 06] Weber, Stefan: *Distribution-Invariant Risk Measures, Information, and Dynamic Consistency*, Mathematical Finance, Vol.16, No.2, pp.419-441, 2006.

[Yaari, 87] Yaari, Menahem E.: *The Dual Theory of Choice under Risk*, Econometrica, Vol.55, No.1, pp.95-115, 1987.

[Yahav, 66] Yahav, Joseph A.: On Optimal Stopping, The Annals of Mathematical Statistics, Vol.37, No.1, pp.30-35, 1966.