On Knightian Uncertainty Models: Optimal Behavior in Presence of Model Uncertainty

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors

der Wirtschaftswissenschaften

durch die

Fakultät für Wirtschaftswissenschaften

der Universität Bielefeld

vorgelegt von

Tatjana Chudjakow

aus Belousowka, Kasachstan

Bielefeld, September 2010

Dekan:	Prof. Dr. Fred G. Becker
Erstreferent:	Prof. Dr. Frank Riedel
Zweitreferent:	Prof. Dr. Moritz Kassmann

to my family

Acknowledgments

This work would not have been possible without the support of several people who contributed with ideas, conversations, guidance and care.

First and foremost, I appreciate the permanent encouragement and support of my advisor Professor Dr. Frank Riedel. His motivation, guidance and excitement for economic theory have been of enormous benefit to me. Not only did he encourage me to widen my horizons and to explore my own ideas — at the same time he also prevented me from getting lost by giving valuable comments and suggestions. Furthermore, Professor Riedel made possible a research stay at Boston University which was an enriching and inspiring experience both from an academic and personal perspective. In addition, I am indebted to Professor Dr. Moritz Kassmann for his effort taken in surveying this thesis.

A huge contribution to this work was made by Jörg who coauthored the second paper. I am grateful for having shared this experience with him.

I acknowledge financial grants from the German Academic Exchange Service (DAAD) and the German Research Foundation (DFG) through the EBIM program.

I am sincerely grateful for institutional support by the IMW and thank all IMW members for being great colleagues making my time at IMW enjoyable and pleasant. In this respect, I would like to specially mention Monika, Simon, Michael, Jörg and Jan-Philip for being not only colleagues but also valuable friends.

Finally special thanks go to my family for supporting me unconditionally throughout my studies.

Contents

Chapter 1

General Introduction

Almost all economic activities and real life decisions take place under uncertainty. The future payoff often depends on an uncertain state of the world realized after the decision has been made. First models on decision making in presence of uncertainty proposed by Von Neumann and Morgenstern (1953) and Savage (1954) assume that there exists a unique probability over possible states of the world. An act is to be preferred to an alternative one if the expected utility of the first one exceeds the expected utility of the second evaluated under a certain probability measure. The question arises however on how to find the probability in question. This seems to be an easy task while gambling in the casino, tossing a coin or in similar lab like situations. Here, the exact probability is known and the situation bears only risk. In more complex situations the right pick is not that easy anymore. Assigning a probability to the event that Germany wins the next world cup includes accounting for contingencies that are far beyond subject's understanding and is a task impossible to be accomplished. While a casino bet is risky, Germany winning the World Cup is uncertain.

The distinction between risk and uncertainty was first discussed by Knight

(1921) and is now referred to as ambiguity or *Knightian uncertainty* [1].

1.1 Knightian Uncertainty

In his seminal book Knight (1921) suggests that there exist random outcomes which cannot be represented by numerical probabilities, i.e. he establishes a clear distinction between *measurable uncertainty* he calls risk and *unmeasurable uncertainty*. Although in later version of his famous book he concludes that modeling real world decisions with economic models is hardly possible his ideas gave rise to a strand of literature on modeling *unknown unknowns*. Ellsberg (1961) constructs an urn experiment in which agents make choices that systematically violate the axioms of expected utility. In the experiment in which probabilities of outcomes are not given completely, participants prefer to bet on outcomes with known probability and systematically refuse to bet on outcomes where probabilities of winning are not given. Behaving this way they make choices that make it impossible to attach probabilities to outcomes in the way prescribed by Savage (1954). This observation suggests that agents not only care about relative likelihood of obtaining a payoff but also about ambiguity attached to it. The work of Ellsberg (1961) first challenges the prevailing subjective expected utility (SEU) paradigm by disentangling aversion to lack of knowledge of future outcomes from the aversion to lack of knowledge of its odds. Several attempts has been made to rationalize choices made by Ellsberg agents. Gilboa and Schmeidler (1989) set up axioms for preferences which lead to Maxmin Expected Utility or the Multiple Priors Model by weakening the Sure Thing principle of Anscombe and Aumann (1963) and add an axiom that captures Uncertainty Aversion to the model. These axioms lead to a representation in which a single prior used in classical models to compute the expected utility needs to be replaced by a whole set

[1] In the following we use terms ambiguity and uncertainty interchangeably meaning the same.

of possible distributions.

Epstein and Schneider (2003b) modify the Gilboa and Schmeidler (1989) axioms to extend the model to a dynamical setting. To do so they modify the preference relation to be time and state dependent and add the dynamic consistency axiom to the preferences. Loosely speaking, dynamic consistency requires that decisions made at a node are still optimal from the perspective of later nodes. The resulting representation is the so called multiple prior recursive utility. As in the static setting the expectation is taken with respect to a set of priors. However, not every set of measures can arise in the representation but only sets satisfying a certain regularity condition. The restriction on the set of priors steams from the dynamic consistency axiom imposed on the preferences and will be discussed later.

Heuristically, the use of a set of measures instead of one captures the inability of the decision maker to assign concrete probabilities. She expresses her aversion to this uncertainty by putting more weight on states unfavorable to her and thus considering the minimal expected utility over all possible probabilities. As a result the ambiguity averse preferences assign utilities in very conservative and cautious manner by only considering the minimal expected utility of a future payoff.

Several authors work on different approaches to rationalize ambiguity averse behavior. Maccheroni, Marinacci, and Rustichini (2006) employ a multiple prior approach but account for plausibility of different models by penalizing models that the agents considers as unlikely, Schmeidler (1989) uses a non-additive measure notion to capture ambiguity, Bewley (2002) relaxes completeness maintaining independence, Klibanoff, Marinacci, and Mukerji (2005) choose a smooth ambiguity approach by disentangling ambiguity from the attitude towards it.

While previous works are all preference based there also have been attempts to formalize robust behavior in other economic disciplines. Hansen and Sargent (2001) introduce the notion of robust control in macroeconomics, another strand in the literature which is mathematically equivalent to the multiple prior utility sketched above is the coherent risk measures approach introduced by Artzner, Delbaen, Eber, and Heath (1999).

By providing a solid decision theoretical ground works of above mentioned authors gave rise to a strand of literature analyzing the consequences of these preferences for predictions of classical economic models. The second generation of literature on ambiguity aversion discussed real world problems under uncertainty.

1.2 Ambiguity in Real Life Applications

1.2.1 Modeling Uncertainty

Decision theoretical models discussed above take the uncertainty as given and abstract from its source and the exact structure of the set of possible models. In his experiment Ellsberg generically imposes ambiguity by not disclosing the decomposition of the urn. The situation is not that clear in real world applications since there is information available that allows to draw conclusions about the odds. The existence and the structure of uncertainty is the first challenge to overcome while modeling meaningful multiple prior models.

A frequent objection of critics of multiple prior models is the argument of learning. A rational agent learns from past observations and infer the right model from data using common filtering methods. As the data set grows large by the Law of Large Numbers empirical frequencies converge to the right probability making multiple prior approach obsolete.

There are several reasons why this reasoning does not fully capture the reality. From subjective perspective individual investor might lack knowledge and information needed to infer the right probability model or just not own powerful machines to deal with the data. Knowing that she is not able to interpret the information available to her, she thus prefers to use a set of models. She might also take into account not only her subjective forecasts but also opinions of experts available on the market.

Even if we take the stand of a more sophisticated agent such as a large bank there are still good reasons for ambiguity. The ongoing financial crisis spectacularly demonstrated how excessive confidence in models used by banks can lead to a disaster and revealed a knew source of uncertainty that had been mainly ignored before the crisis – the model uncertainty. It arises due to the impossibility to completely express the reality through abstract models and is greatly amplified by the complexity of modern quantitative models used in finance. In these situations minimax expected utility seems an appropriate choice. Since the value of the portfolio strongly depends on the chosen model more robust decision making procedure is desirable.

By setting up a meaningful multiple prior model one still has to be careful when choosing the size of the set of measures since the degree of uncertainty captured by this set strongly affects predictions of the model. A natural way to choose the set of models is to consider sets of measures that are statistically close and thus hard to distinguish from a series of observations as proposed by Kogan and Wang (2002). One then chooses a confidence level and considers all models that cannot be distinguished at this confidence level. Another way is to consider all models consistent with current market conditions such as bid-ask spread or regulations routines imposed by authorities. The wast majority of models we are considering stick to one of above mentioned techniques while modeling ambiguity. It is noteworthy however, that the way one sets up the multiple prior structure can drastically impact prediction of the model.

In what follows we will give a short overview on multiple prior models of financial markets. The main focus will lie on optimal stopping problems that arise in the context of American options and on asset pricing and allocation under ambiguity.

1.2.2 Ambiguity and Optimal Stopping Problems

Optimal Stopping Problems arise in economics as right timing is often crucial for realizing a profit opportunity. This might be launching a new product before your opponent does, selling a house at a good price or hiring the right person for a given job. The classical solution formulated by Snell (1952) relies heavily on the specification of the probability space and imposes a fixed prior over future payoffs. Riedel (2009) generalizes the method of backward induction in order to account for uncertainty by using a set of measures instead of one. The regularity condition that already appeared in Epstein and Schneider (2003b) also plays a crucial role here. Formulated in a slightly different though equivalent way it ensures that the Law of Iterated Expectations still holds true and thus allows dynamic programming reasoning. Mathematically, the assumption imposed on the set of priors is equivalent to the stability of the set of measures under cutting and pasting them together. Intuitively speaking time-consistency ensures that the decision maker can change his mind in every period about which measure he thinks is the true one or the worst one. Time-consistency guarantees that this measure which might be time- and state-dependent is contained in the set of her possible measures. This implies that as time passes she will never regret his previous decisions since at every point in time he can decide optimally. The concept has also been discussed by Föllmer and Schied (2004), Delbaen (2002). A brief review of different concepts along with a proof of equivalence can be found in Riedel (2009).

Given time-consistency assumption the ideas of Snell (1952) carry over: the value function can be formulated in a similar way through Snell envelope, the optimal stopping times have similar form. This allows to solve much more complicated problems of multiple prior optimal stopping using familiar techniques of dynamic programing. As in the classical version the backward induction techniques provides the optimal stopping strategy and assigns value to the problem. However, the multiple prior structure drastically increases the computational complexity of the model since a set of expectations has to be computed instead of a single one in every step. This complexity can be reduced by exploiting a duality result already appeared in Karatzas and Kou (1998): the multiple prior stopping problem can be reduced to a single prior problem for a specific measure, the so called *worst-case measure*. Several papers use this technique to tackle multiple prior stopping problems.

In an early paper Nishimura and Ozaki (2004) studied the effects of Knightian uncertainty on job search in a discrete model and provided a closed form solution for the optimal reservation wage. Several examples of optimal stopping problems under uncertainty can be found in Riedel (2009). He shows that ambiguity decreases the value of the option and leads to earlier stopping for many classes of payoffs.

Nishimura and Ozaki (2007) and Kort and Trojanowska (2007) analyze the firm's entry and exit decision under uncertainty in continuous time and show that ambiguity decreases the option value but does not lead to earlier stopping in general. A similar point has been made by Miao and Wang (2004) for a slightly more general class of payoffs.
Recently much work has been done on the dynamic models in continuous time. Mathematical grounds to this go back to Peng (1997) who introduced the notion of *g-expectations*. Optimal stopping problems in continuous time then correspond to Backward Stochastic Differential Equation[2].

1.2.3 Ambiguity and Asset Markets

Ambiguity about market conditions or the right market model to use affects the valuation of investors and thus the choices they makes on a market. As a

[2]See Riedel (2010) for further references.

result portfolio allocations in presence of uncertainty structurally differ from allocations predicted by SEU models. This changes the resulting market outcomes and helps explaining market anomalies and failures not predicted by standard SEU models.

While an SEU agent solves a maximization problem when choosing optimal portfolios, an ambiguity averse decision maker solves a maxmin problem maximizing her minimal expected utility.

First analysis of the qualitative changes of investment behavior under ambiguity is due to Dow and Werlang (1992). They show that the valuation of the security depends on the direction of the trade. In a multiple prior model the investor evaluates a long position according to the lowest return of this security and short position using the highest return on it. As a result she only buys the security if the lowest expected return exceeds its price. Similarly, she shorts the security if the highest return is below the market price. If the security price lies within the interval of the highest and the lowest valuation the agent chooses not to trade the asset at all.

A growing body of literature uses this fundamental observation to explain phenomena observed in reality and not captured by classical models. Epstein and Miao (2003) assume that the uncertainty about returns on assets on foreign company is more pronounced than the uncertainty about home country stocks and show that this can explain home-bias. Using a similar argument Boyle and Wang (2009) and Uppal and Wang (2003) explain overinvestment in the own company and underdiversification.

Epstein and Schneider (2003b), Illeditsch (2009) show that ambiguity averse investor overreacts to bad news and underreact to good news again underpinning the pessimistic attitude of the preferences. In extreme case the optimal portfolio exhibits inertia not reacting to the changes of fundamentals at all for a range of parameters.

Moreover, while non-participation is never optimal for an SEU investor it arises naturally if there is ambiguity about expected returns due to the

existence of the no-trade interval. Epstein and Chen (2002) observed non-participation in a dynamical Lucas-style model, Cao, Wang, and Zhang (2005) and Ui (2009) studied endogenous participation in a static model with investors differing in their degree of ambiguity. Mukerji and Tallon (2001) show that no-trade decision can generate incompleteness of the market endogenously.
A large body of literature including many studies mentioned above approaches optimal asset allocation under ambiguity by employing a mean-variance analysis. Maenhout (2004) shows that uncertainty decreases the portfolio weight of risky assets. In a model where ambiguity is generated by different opinions of experts advising the investor Lutgens and Schotman (2010) point out that the portfolio weights are more stable compared to portfolios resulting in a single prior model.

Most of the works discussed above take the prices of securities as given. However, portfolio decisions impact the aggregate demand and supply of the security and thus prices. Among others Epstein and Chen (2002), Cao, Wang, and Zhang (2005), study asset pricing and resulting equity premia on markets with ambiguity. Epstein and Schneider (2003b) point out that ambiguity aversion can explain a part of the premium puzzle since it acts as extra risk aversion, Routledge and Zin (2009) shows that ambiguity generates illiquidity increasing the bid-ask spread. Caballero and Krishnamurthy (2007) point out that an increase in ambiguity about asset returns may induce flights to quality causing distress in the market for the ambiguous security. The leitmotiv of papers cited here is that ambiguity makes the asset less attractive. As a result investors require a higher premium for holding the asset and sell it if ambiguity becomes to severe.

1.3 The Aim of this Work

As mentioned above a common theme in the literature is that ambiguity averse preferences enforce a more conservative and cautious behavior. Ambiguity averse decision maker aims to reduce their exposure to uncertainty by preferring safer payoffs to risky/uncertain ones. For example, an ambiguity averse seller of a house accepts a lower bid for the house instead of waiting for uncertain (in his mind likely dark) future to come. The reason for the lower reservation bid is the use of the worst-case model for the future bid distribution – the model that gives him the lowest expected utility. The worst-case measure in this example is intuitive and easy to identify: choosing the prior that puts more weight on low bids at every node of the decision tree minimizes the utility over the whole set of models. Although the set of beliefs can be very complicated the decision maker here uses a rather simple one, sticking to it once chosen. He then acts as an SEU maximizer (although a very pessimistic one) and his optimal strategy follows by means of classical backward induction making the complex machinery of multiple prior backward induction and cautious modeling of the set of beliefs obsolete.[3]

This observation lead to a long standing belief that ambiguity aversion is just a form of risk aversion and its effects are observationally equivalent to risk aversion patterns. However, the hypothesis that every ambiguity averse decision maker acts as an SEU maximizer for an appropriately chosen measure that can be identified a priori is not true in general. The ability to reduce the above problem to a SEU problem for a fixed deterministic measure relies heavily on the structure of the payoff processes one considers.

The first objective of this work is to take a closer look at the key factors that ensure the simple deterministic form of the worst-case measure and to provide examples for payoffs where the worst-case measure is non determin-

[3] Riedel (2009) provides an extensive analysis of such problems which he calls monotone.

istic. Then, the worst-case measure changes over time as the state process is realized and the simple reduction argument used above fails to work. Basically, the simplicity of the worst-case measure in examples studied in the literature so far grounds on three assumptions: measurability, monotonicity and Markovian structure of the payoff process. We show that relaxing those assumptions generates stochastic worst-case measures that depend on the whole history of the process. We illustrate this fact by analyzing several exotic American options in a discrete binomial tree model in the spirit of Cox, Ross, and Rubinstein (1979). For example, if the effect of uncertainty changes over time as it is the case for barrier options the worst-case measure changes with it becoming path dependent. We observe similar effects for non-monotone payoffs (in the case of straddle) and non-measurable payoffs (in the case of shout options). In all these examples the worst-case measure that the decision maker uses is not known a priori and the problem cannot be reduced to an SEU problem. Here, we can distinguish an ambiguity averse decision maker from an extremely pessimistic SEU maximizer by choices he makes while exercising the option.

A second fact widely assumed in the literature is that ambiguity aversion by making the future less valuable leads to earlier stopping. In the above example of the house selling as in the job search example analyzed by Nishimura and Ozaki (2004) an increase in ambiguity decreases the bid /the reservation wage that is accepted by the agent. In this sense the agents stops earlier. To test the hypothesis of earlier stopping we model and solve a best choice problem first introduced by Gardner (1960) in the multiple prior framework: An ambiguity averse decision maker aims to choose the best among a fixed number of applicants that appear sequentially in a random order. The only information available to the decision maker are the relative ranks of applicants seen so far. The decision needs to be taken immediately and once rejected applicants cannot be recalled. The main challenge of the analysis is to formulate a model that accommodates ambiguity aversion in a meaningful way. We do so by assuming ambiguity that the probability to meet a candidate – a relatively top applicant — is not not exactly known but comes from a specified interval. We show that our model covers the classical secretary problem, but also other interesting classes of problems. Our main result is that the stopping rule is simple, i.e. it is optimal to reject a constant known number of applicants and then to take the first candidate appearing, as in the case of classical secretary problem discussed by Gardner (1960). However, unlike examples above an increase in ambiguity does not lead to earlier stopping in general. The reason for this is that the payoff obtained from stopping is ambiguous itself. Thus, stopping does not necessarily reduce ambiguity. As a result it might be optimal to delay stopping to avoid the ambiguity attached to payoff obtained from stopping.[4] Moreover, as with American options the worst-case measure of the problem is stochastic again due to the non-measurability of the payoff.

While the first two chapters work along similar lines analyzing optimal decisions in dynamic settings the last chapter changes the perspective. Instead of analyzing the effects of ambiguity for the decision process of an individual agent we consider a static partial equilibrium model of an economy where some agents are ambiguity averse. The question is similar. Many authors noted that an ambiguity averse investor rather chooses safer assets to avoid uncertainty while selecting a portfolio. They require an additional premium to be rewarded for the ambiguity if they hold the ambiguous asset. This way, one can explain the equity premium puzzle[5] or flights from market with increased ambiguity.[6] Again, the suggestion made in the literature is that

[4]Kort and Trojanowska (2007) made a similar point with a different example.

[5]See Mehra and Prescott (1985) for the puzzle and Epstein and Wang (1994) for the model with ambiguity aversion.

[6]See Caballero and Krishnamurthy (2007).

ambiguity aversion ".. acts as an extra risk aversion..."[7] and leads to lower equilibrium price. Cao, Wang, and Zhang (2005) were first to notice that increasing the ambiguity aversion can lower equity premium since agents that require a high ambiguity premium leave the market. As a result the ambiguity premium paid on the market in equilibrium decreases leading to higher prices. In the last chapter of this work we make a similar point and show that an increase in ambiguity can lead to lower ambiguity premia if market participants are heterogeneous. In our model agents differ not only in their attitude towards uncertainty but also in their beliefs about the return of the market security. If the divergence of opinions and the degree of ambiguity is large enough, optimistic investor will demand a large amount of the security. However, ambiguity averse decision maker will refuse to supply the security if its price happens to lie in the no-trade interval. The demand of optimistic agents on the market then drives the premia down due to pronounced demand. We point out that ambiguity not only affects the demand but also the supply side of the economy. This way, the equity premium can decrease not because agents stop buying the security but because the agents stop selling it and thus slacken the supply.

The main chapters of this thesis, each of which self contained in notation, are based on three articles. The first two consider optimal stopping behavior of an agent facing uncertainty. Chapter 2 is coauthored with Frank Riedel and analyzes the best choice problem under uncertainty. Chapter 3 based on a joint work with Joerg Vorbrink studies the structure of worst-case measures for several American options. The last chapter of this work changes the perspective and analyzes overpricing in a static partial equilibrium model.

To this point we have given a brief outline of the general context and developments which lead to this work. Since the questions and topics treated in the following chapters differ a more detailed scientific placement of this

[7] Epstein and Wang (1994)

work will be discussed in each chapter separately.

Chapter 2

A Best Choice Problem under Ambiguity

2.1 Introduction

The classical secretary problem is one of the most popular problems in the area of Applied Probability, Economics, and related fields. Since its appearance in 1960 a rich variety on extensions and additional features has been discussed in the scientific[1] and popular[2] literature.

In the classical secretary problem introduced by Gardner (1960) an employer sequentially observes N girls[3] applying for a secretary job appearing in a random order. She can only observe the relative rank of the current girl compared to applicants already observed and has no additional information on their quality. The applicants can be strictly ordered. Immediately after

[1]Freeman (1983) gives an overview of the development until the eighties. Ferguson (1989) contains historical anecdotes. A lot of material is covered in Berezovski and Gnedin (1984).

[2]Gardner's treatment is the first instance here. A most recent example is the treatment in a book about "love economics" by a German journalist Beck (2005). It plays also a role in psychological experiments, see Todd (2009).

[3]In the following we will use *applicant* and *girl* interchangeably meaning applicants of both sexes.

the interview the employer has to accept the girl or to continue the observation. Rejected applicants do not come back. Based on this information the agent aims to maximize the probability of finding the best girl[4].

Most of the classical literature in this field assumes that the girls come in random order where all orderings are equally likely. The solution is surprisingly nice. It prescribes to reject a known fraction of girls, approximately $\frac{1}{e}$, and to accept afterwards the next *candidate*, i.e. a girl with relative rank 1. Such stopping rules are called simple. This strategy performs very well: Indeed, the chance of success is approximately $36,8\% \approx \frac{1}{e}$ for large N.

This nice and surprising result is based on the strong assumption that the girls arrive randomly, all possible orderings being equally likely. There are good reasons to care about the robustness of this assumption. From a subjective point of view, the decision maker might not feel secure about the distribution of arrivals — she might face ”ambiguity”. Even if we take a more objective point of view, we might want to perform a sensitivity analysis of the optimal rule. While there is certainly a degree of randomness in such choice situations, it is not obvious that the arrival probability would be independent of the girl's quality, e.g. It might well be that more skilled applicants find open jobs earlier[5]. In this paper, we present a way of dealing with these questions by embedding the best choice problem into a multiple prior framework as introduced by Gilboa and Schmeidler (1989) for the static case, and extended to dynamic frameworks by Epstein and Schneider (2003b). The agent works here with a class of possible prior distributions instead of a single one, and uses the minimax principle to evaluate her payoffs. We use then the general theory for optimal stopping under ambiguity developed in

[4]This is an extreme utility function, of course. On the other hand, the analysis based on this extreme assumption serves as a benchmark for more general utility functions. The results are usually similar, at least in the Bayesian setting, see Ferguson (2006), e.g.

[5]Another obvious way of introducing ambiguity concerns the number of applicants. This question is not pursued here; see Engelage (2009) for a treatment of best choice problems with an unknown and ambiguous number of applicants.

Riedel (2009) to analyze the model.

Our main result is that the optimal stopping rule is still *simple*, or a cutoff rule. The agent rejects a certain number of girls before picking the next applicant that is better than all previous ones — in the literature on best choice problems, such applicants are usually called *candidates*. The optimal strategy thus consists in building up a "database" by checking a certain number of applicants, and to take the first candidate thereafter[6]. We are able to obtain an explicit formula for the optimal threshold that determines the size of the database.

In best choice problems, ambiguity can lead to earlier or later stopping compared to the Bayesian case, in contrast to the analysis in Riedel (2009) where ambiguity leads to earlier stopping. The reason for this is that the original payoff process in best choice problems is not adapted. Indeed, when the employer accepts a candidate, she does not know if that candidate is the best among all applicants. She would have to observe all of them to decide this question. She thus uses her current (most pessimistic) belief about the candidate indeed being the best applicant. Two effects work against each other then. On the one hand, *after* picking a candidate, the agent's pessimism leads her to believe that the probability of better candidates to come is very high — this effect makes her cautious to stop. On the other hand, *before* acceptance, she uses a very low probability for computing the chance of seeing a candidate. This effect makes her eager to exercise her option. We illustrate these effects with three classes of examples.

In general, the optimal threshold can be quite different from the classical case (and in this general sense, the 37 %–rule described above is not robust). When the highest probability of finding a candidate decays sufficiently fast, the threshold– number of applicants–ratio can be very close to zero; indeed,

[6]Optimal stopping rules need not be simple. For example, in the situation with incomplete information about the number of objects a Bayesian approach does not lead to simple stopping rules in general, see Presman and Sonin (1975).

it is independent of the number of applicants for large N. In such a situation, one has rather an absolute than a relative threshold. Instead of looking at the first 37 % of applicants, one studies a fixed number of them before choosing the first relatively top applicant.

On the other hand, if the probability of finding a candidate at time n is in the interval $[\frac{\gamma}{n}, \frac{1}{\gamma n}]$ for some parameter $\gamma \in (0,1)$, the threshold– number of applicants–ratio can converge to any positive number between 0 and 1. For $\gamma \to 1$, we obtain again the 37 % – rule. In this sense, the classical secretary problem is robust.

Last not least, we give an example where the ambiguity about applicants being candidates remains constant over time. This example can be viewed as the outcome of independent coin tosses with identical ambiguity as described in Epstein and Schneider (2003a)[7]. The aim is to pick the last 1 in this series of zeros and ones. We show that the agent optimally skips all but a finite number of applicants. In this case, we the ratio converges to 1 for large N. Different parametrization of this example show that ambiguity can lead to earlier stopping (when the probability of finding a candidate are known to be small) as well as later stopping compared to the Bayesian case.

On the modeling side, our approach succeeds in finding a model that allows to introduce ambiguity into best choice problems. Note that one has to be careful when introducing ambiguity into dynamic models because one can easily destroy the dynamic consistency of the model[8]. To do so, we reformulate the classical secretary problem in the following way. The agent observes a sequence of ones and zeros, where 1 stands for "the current applicant is the best among the candidates seen so far". The agent gets the payoff of 1 if she stops at a 1 and there is no further 1 coming afterwards. In the secretary problem, the probability of seeing a 1 at time n is $1/n$ as all

[7]See also the examples of this type discussed in Riedel (2009).

[8]See Epstein and Schneider (2003b) for a general discussion of time–consistency in multiple prior models, and Riedel (2009) for the discussion in an optimal stopping framework.

orderings are equally likely. We then allow for ambiguity by introducing an interval $[a_n, b_n]$ for this probability; finally, we construct a time–consistent model by pasting all these marginal probabilities together as explained in Epstein and Schneider (2003b).

The analysis of the stopping problem proceeds in three steps. In a first step, we have to derive an equivalent model with adapted payoffs — note that the payoff function is not adapted here because the agent's payoff depends on the events that occur after stopping. We pass to adapted payoffs by taking conditional expectations prior by prior; it is not at all clear that this leads to the same ex ante payoff, though. Time–consistency and the corresponding law of iterated expectations for multiple priors[9] ensure this property. In the second step, we compute explicitly the relevant minimal conditional expectations. After having stopped, the agent uses the *maximal* probability for seeing a 1 afterwards. Intuitively, the agent's pessimism induces him to suppose that the best candidate is probable to come later after having committed herself to an applicant. After this, we have arrived at an optimal stopping problem that can be solved with the methods developed in Riedel (2009). Indeed, the problem at hand turns out to be a *monotone* problem: the worst–case measure can thus be identified as the measure under which the probabilities of seeing a candidate are minimal (until the time of stopping, of course). It then remains to solve a classical Bayesian stopping problem, and we are done.

The paper is organized as follows: Section 2 introduces the model and provides the stepwise solution as well as the main theorem. Section 3 contains three classes of examples that allow to discuss in more detail the effects of ambiguity in best choice problems.

[9]See, e.g., Riedel (2009), Lemma 11.

2.2 Best Choice Problems under Ambiguity

Let us start with formalizing the classical best choice problem in a way that allows a natural generalization to ambiguity[10]. In the classical secretary problem, the agent observes sequentially the relative ranks of applicants, say $R_1 = 1$ for the first one, $R_2 \in \{1, 2\}$ for the second, $R_3 \in \{1, 2, 3\}$ for the third and so on. The random ordering implies that the random variables $(R_n)_{n \leq N}$ are independent[11]. As we are only interested in finding the best girl, we can discard all applicants with a relative rank higher than 1, and call *candidates* those girls that are relatively top at the moment. Let us introduce the $0-1$–valued sequence $Y_n = 1$ if $R_n = 1$ and $Y_n = 0$ else. The random variables (Y_n) are also independent, of course, and we have

$$P[Y_n = 1] = \frac{1}{n}$$

because all permutations are equally likely[12].

A *simple* stopping rule first rejects $r - 1$ applicants and accepts the next candidate, if it exists, i.e.

$$\tau(r) = \inf \{k \geq r | Y_n = 1\}$$

with $\tau(r) = N$ if no further candidate appears after applicant $r-1$. One uses independence of the $(Y_n)_{n \leq N}$ and monotonicity of the value function to show that optimal stopping rules must be simple, see Section 2.2.1 below for the argument in our context. It then remains to compare the expected success of the different simple rules. The event that girl n is a candidate and also the best of all girls means that no further girl has relative rank 1. In terms of our variables $(Y_n)_{n \leq N}$, this means that $Y_n = 1$ and $Y_k = 0$ for all $k > n$.

[10]A similar treatment of the classical secretary problem is due to Bruss (2000)

[11]See Ferguson (2006) or Chow, Robbins, and Siegmund (1971) for the technical details.

[12]Note that the independence of relative ranks ensures that we do not loose information by conditioning on the σ-algebra generated by $(Y_n)_{n \leq N}$ instead of $(R_n)_{n \leq N}$.

The success of a simple strategy is then

$$\begin{aligned}
\phi(r) := P[\tau(r) \text{ picks the best girl}] &= \sum_{n=r}^{N} P[\tau(r) = n, \text{girl } n \text{ is best}] \\
&= \sum_{n=r}^{N} P[Y_r = 0, \ldots, Y_{n-1} = 0, Y_n = 1, Y_k = 0, k > n] \\
&= \sum_{n=r}^{N} \prod_{j=r}^{N} P[Y_j = 0] \, \frac{P[Y_n = 1]}{P[Y_n = 0]} \\
&= \prod_{j=r}^{N} \frac{j-1}{j} \sum_{n=r}^{N} \frac{1/n}{1 - 1/n} = \frac{r-1}{N} \sum_{n=r}^{N} \frac{1}{n-1} \, .
\end{aligned}$$

The sum approximates the integral of $1/x$, so the value is approximately

$$\phi(r) = \frac{r-1}{N} \log \frac{N}{r-1} \, .$$

The maximum of the function $-x \log x$ is in $1/e$, so we conclude that the optimal r is approximately $[N/e] + 1$.

2.2.1 Best Choice under Ambiguity

Formulation of the Problem

We generalize now the above model to ambiguity by allowing that the probabilities

$$P[Y_n = 1 | Y_1, \ldots, Y_{n-1}] \in [a_n, b_n]$$

for all histories $Y_1, \ldots, Y_{n-1}$ come from an interval instead of being a known number. Throughout the paper, we assume that $0 < a_n \leq b_n < 1$. Note that the bounds on conditional probabilities a_n, b_n are allowed to depend on time, but not on the realized path. For example, more skilled applicants are likely to apply sooner. If you think of the search where the hiring executive is visiting several universities while looking for the secretary, she might adjust her beliefs about meeting a skilled applicant depending on the university she is visiting. In the classical problem the probability of finding a top applicant also varies with time, there we have

$$P[Y_n = 1|Y_1, \dots, Y_{n-1}] = a_n = b_n = \frac{1}{n}.$$

Modeling ambiguity in dynamic settings requires some care if one wants to avoid traps and inconsistencies. We view the random variables $(Y_n)_{n \leq N}$ as outcomes of independent, but ambiguous experiments where in the nth experiment the distribution of Y_n, i.e. the number $P[Y_n = 1]$ is only known to come from an interval $[a_n, b_n]$. From these marginal distributions, the agent has to construct all possible joint distributions for the sequence $(Y_n)_{n \leq N}$. She does so by choosing any number $p_n \in [a_n, b_n]$ after having observed $Y_1, \dots, Y_{n-1}$. A possible prior then takes the form

$$P[Y_1 = 1] = 1 \tag{2.1}$$

because the first applicant is always a candidate, and

$$P[Y_n = 1|Y_1, \dots, Y_{n-1}] = p_n \in [a_n, b_n] \tag{2.2}$$

for a predictable sequence of one–step–ahead probabilities p_n. Note that we allow p_n to depend on the past realizations of $(Y_1, \dots, Y_{n-1})$. For a time–consistent worst–case analysis this is important because different one–step–ahead probabilities might describe the worst case after different histories. From now on, we work with class $\mathcal{P}$ of all probability measures that satisfy (2.1) and (2.2) for a given sequence $0 < a_n \leq b_n < 1$, $n = 1, \dots, N$. For more on the foundations of dynamic decisions under ambiguity, we refer the reader to Epstein and Schneider (2003b) and Epstein and Schneider (2003a), see also Riedel (2009).

The astute reader might now wonder why we speak about independent realizations if the conditional probabilities are allowed to depend on past observations. Independence in a multiple prior setting is to be understood in the sense that the interval $[a_n, b_n]$ is independent of past observations, just as it means that the conditional probability of the event $\{Y_n = 1\}$ given the past observations is independent of these observations in classical probability. In this sense, the agent does not learn from past observations about the degree of ambiguity of the nth experiment.

We are now ready to formulate our optimization problem. Based on the available information the agent chooses a stopping rule $\tau \leq N$ that maximizes the expected payoff which is 1 if she happens to find the best girl, and 0 else. A way to describe this in our model is as follows: applicant n is the best if she is a candidate (she has to be relatively best among the first n), and if she is not topped by any subsequent applicant: in other words, we have $Y_n = 1$ and there is no further candidate afterwards, or $Y_k = 0$ for $k > n$. Let us define

$$Z_n = \begin{cases} 1 & \text{if } Y_n = 1, Y_k = 0, k > n \\ 0 & \text{else} \end{cases} .$$

The agent aims to choose a stopping rule τ that maximizes

$$\inf_{P \in \mathcal{P}} \mathbb{E}^P[Z_\tau]. \tag{2.3}$$

Reformulation in Adapted Payoffs

The next problem that we face is that the sequence $(Z_n)_{n \leq N}$ is not adapted to the filtration generated by the sequence $(Y_n)_{n \leq N}$ because we do not know at the time when we pick an applicant if she is best or not. As in the classical case, we therefore take first conditional expectations of the rewards $(Z_n)_{n \leq N}$ before we can apply the machinery of optimal stopping theory. In the multiple prior framework we thus consider

$$X_n = \operatorname*{ess\,inf}_{P \in \mathcal{P}} \mathbb{E}[Z_n | \mathcal{F}_n] \,.$$

where $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$. In the Bayesian framework, it is relatively easy to show that the expected payoffs $\mathbb{E}[Z_\tau] = \mathbb{E}[X_\tau]$ are the same for all stopping times τ. In the multiple prior framework, this is less obvious. Indeed, the identity

$$\inf_{P \in \mathcal{P}} \mathbb{E}^P[Z_\tau] = \inf_{P \in \mathcal{P}} \mathbb{E}^P[X_\tau]$$

does require a condition on the set of priors which has become known as *rectangularity* or *m–stability*. In our model, this condition is satisfied[13], and we have

Lemma 1. *For all stopping times $\tau \leq N$ we have*

$$\inf_{P \in \mathcal{P}} \mathbb{E}^P[Z_\tau] = \inf_{P \in \mathcal{P}} \mathbb{E}^P[X_\tau].$$

We can thus reformulate our problem as

$$\text{maximize } \inf_{P \in \mathcal{P}} \mathbb{E}^P[X_\tau] \tag{2.4}$$

over all stopping times $\tau \leq N$.

Reduction to a Monotone Problem

We are now in the position to apply the general theory of optimal stopping with multiple priors as developed in Riedel (2009). To this end, let us first have a closer look at the payoffs $(X_n)_{n \leq N}$. It is clear that $X_n = 0$ if we do not have a candidate at n, i.e. $Y_n = 0$, so we need only to focus on the case $Y_n = 1$. We are then interested in the minimal (conditional) probability that all subsequent applicants are no candidates. It is quite plausible (but requires a proof, of course) that the probability is minimal under the measure $\bar{P}$ where the probabilities for being a candidate are maximal, $\bar{P}[Y_n = 1 | Y_1, \ldots, Y_{n-1}] = b_n$. Under this measure, the $(Y_n)_{n \leq N}$ are independent (because the conditional probabilities for $Y_n = 1$ are independent of past observations, but see the proof of Lemma 2 for the details), and we thus have

[13] Compare Epstein and Schneider (2003b) or Riedel (2009), Section 4.1.

Lemma 2. *The payoffs $(X_n)_{n\leq N}$ satisfy*

$$\begin{aligned} X_n =& Y_n \cdot \min_{P\in\mathcal{P}} P[Y_{n+1}=0,\ldots,Y_N=0] \\ =& Y_n \cdot \prod_{k=n+1}^{N}(1-b_k) \\ =:& Y_n \cdot B_n \end{aligned} \tag{2.5}$$

The agent faces now a sequence of adapted payoffs where for each $n \leq N$ the payoff Z_n is monotone in Y_n (indeed, linear). The random variables $(Y_n)_{n\leq N}$ are independent under the measure Q where the conditional probabilities for a candidate are

$$Q[Y_n = 1|Y_1,\ldots,Y_{n-1}] = a_n\,. \tag{2.6}$$

Moreover, the probabilities of finding a candidate are smallest under this measure in the whole class $\mathcal{P}$ in the sense of first–order stochastic dominance. We are thus in a situation that is called a monotone problem in Riedel (2009). The general theory there tells us that the optimal stopping rule with multiple priors coincides with the optimal stopping rule under the measure Q – the worst-case measure.

Theorem 1. *The optimal stopping rule τ^* for 2.4 is the same as the optimal stopping rule for the Bayesian problem*

$$\textit{maximize } \mathbb{E}^Q[X_\tau]. \tag{2.7}$$

over all stopping times $\tau \leq N$ where Q is as defined in 2.6.

Optimal Stopping under the Worst–Case Measure Q

We are now back to a classical optimal stopping problem under the measure Q. A standard argument shows that optimal stopping rules must be simple. It works as follows. From classical optimal stopping we know that it is optimal to stop when the current payoff X_n is equal to the current value of the problem

$$v_n := \sup_{\tau \geq n} \mathbb{E}^Q \left[X_\tau | X_1, \ldots, X_n \right] .$$

The independence of the $X_1, \ldots, X_N$ under Q implies that the value of the problem after having rejected $n-1$ applicants is independent of the past observations, i.e.

$$v_n = \sup_{\tau \geq n} \mathbb{E}^Q [X_\tau]. \tag{2.8}$$

The sequence $(v_n)_{n \leq N}$ is decreasing as we maximize over a smaller set of stopping times. On the other hand, the sequence $(B_n)_{n \leq N}$ defined as

$$B_n := \prod_{k=n+1}^{N} (1 - b_k)$$

is increasing in n. Now suppose that it is optimal to take a candidate n. We have then $B_n = v_n$; therefore, we get

$$B_{n+1} \geq B_n = v_n \geq v_{n+1} ,$$

and it is also optimal to stop when a candidate appears at time $n+1$. We conclude that optimal stopping rules are simple.

Lemma 3. *The optimal stopping rule τ^* is simple, i.e. there exists a number $1 \leq r^* \leq N$, s.t.*

$$\tau^* = \tau(r^*) = \inf\{n \geq r^* | Y_n = 1\}.$$

In the next step we compute the optimal threshold r^* maximizing (2.8) over all simple strategies. Let us denote by

$$\phi(r) := \mathbb{E}^Q [X_{\tau(r)}]$$

the payoff from starting to search at applicant r. We then have

$$\phi(N) := \mathbb{E}^Q (X_{\tau(N)}) = a_N \tag{2.9}$$

and

$$\phi(r) = a_r \cdot B_r + (1 - a_r) \cdot \phi(r+1) \tag{2.10}$$

for $r < N$.

While our recursive formula for $\phi(r)$ is useful for numerical computations, we record also the explicit solution of this linear difference equation. To simplify the interpretation of this expression, we introduce two concepts.

Definition 1. *For each $n \le N$ we call*

$$\alpha_n = \frac{1 - a_n}{1 - b_n} = 1 + \frac{b_n - a_n}{1 - b_n}$$

the degree of ambiguity and

$$\beta_n = \frac{a_n}{1 - b_n}$$

the ambiguous odds of applicant n.

The first ratio α_n measures the relative ambiguity persisting at the time n. The term tends to 1 as the size of the interval $[a_n, b_n]$ decreases. In case of $a_n = b_n$ the node n is completely unambiguous and the decision maker faces only risk at n. Similarly, one can think of the product $\prod_{k=n}^{N} \alpha_k$ as the cumulated ambiguity persisting between n and N. The model is unambiguous if and only if $\prod_{k=1}^{N} \alpha_k = 1$. Note, that we call the ratio $p/(1-p)$ the odds for a zero–one bet. In a similar way, the ration β_n measures the odds of seeing a candidate at time n where we now use the (nonlinear) probability induced by our ambiguity model.

The solution of the linear difference equation 2.10 with boundary condition 2.9 is given by

$$\begin{aligned} \phi(r) =& B_{r-1} \left(\beta_n + \alpha_n \beta_{n+1} + \cdots \prod_{k=r}^{n-1} \alpha_k \beta_N \right) \\ =& B_{r-1} \cdot \left(\sum_{n=r}^{N} \beta_n \prod_{k=r}^{n-1} \alpha_l \right) . \end{aligned} \tag{2.11}$$

Let us check now that ϕ has a unique maximizer. From our recursion 2.10, we get that $\phi(r) - \phi(r+1) \geq 0$ is equivalent to $w_r \leq 1$ for

$$w_r := \frac{\phi(r+1)}{B_r} = \sum_{n=r}^{N} \beta_n \prod_{k=r}^{n-1} \alpha_k \,. \tag{2.12}$$

As $\alpha_k > 1$ and $\beta_n > 0$, the sequence $(w_r)_{r\leq N}$ is strictly decreasing. Thus, $(\phi(r))_{r\leq N}$ is increasing as long as $w_r > 1$ and decreasing afterwards, which shows that it has a unique maximum.

The maximizer is determined by

$$r^* = \inf\{r \geq 1 | w_r \leq 1\} \tag{2.13}$$

The optimal threshold r^* is determined by the weighted average of ambiguous odds weighted with the ambiguity persisting between r and N. Equation (2.12) and Equation (2.13) completely characterize the solution.

We summarize our findings in the following theorem.

Theorem 2. 1. *The optimal stopping rule for 2.3 is simple, i.e. the agent first observes r^* candidates and takes then the first candidate that appears;*

2. *The optimal threshold r^* for the cutoff is given via* (2.12) *and* (2.13).

2.3 Comparative Statics

In this section we use the sequence $(w_r)_{r\leq N}$ and the variables $(\alpha_n)_{r\leq N}$ and $(\beta_n)_{r\leq N}$ defined above to analyze the effects of ambiguity on stopping and the structure of the stopping strategy τ^*.

As it was shown in Riedel (2009), an ambiguity averse decision maker behaves like a Bayesian decision maker under a special worst-case probability measure constructed via backward induction. We have seen in the preceding section how to construct this measure, and that the optimal stopping rule is still simple. The central question by analyzing the effect of ambiguity is now the threshold r^*. In case of monotone problems where the payoff is known at the time of decision such as *House Selling Problem* or *Parking Problem* discussed in Riedel (2009) ambiguity leads to earlier stopping. The use of the worst-case measure lowers the value of the Snell envelope and forces the agent to stop earlier. The situation differs here because the agent faces actually two kinds of uncertainty. On the one hand, there is payoff uncertainty in the adapted version of the problem because the probability distribution of Y_n is not known. This effect leads to earlier stopping because it reduces the expected value from waiting. On the other hand, ambiguity also affects the chances that a better applicant is going to come after the current candidate. This ambiguity induces the agent to wait longer because she believes *after* stopping that candidates are going to appear with high probability. The two effects work against each other, and we thus proceed to study more detailed models in which we can disentangle them[14]. In addition, we compute the value of the threshold r^* and show that asymptotically, the relative fraction of applicants that the agent lets go by can assume any value between 0 and 1.

2.3.1 Ambiguous Secretary Problem

Our first example is the multiple prior version of the classical secretary problem. The decision maker is uncertain about the real distribution of the orderings for reasons explained in the introduction but has no additional information on the quality of the applicants. Doubting her strategy she aims to know what happens if she changes the measure slightly. Instead of $P[Y_n = 1] = \frac{1}{n}$, the ambiguity averse decision maker assumes that the probability lies in an

[14] A similar point has been made in a completely different model in Nishimura and Ozaki (2007) when there is uncertainty about the timing and uncertainty about the value from stopping.

N / γ	1	0.9	0.8	0.7	0.6	0.4	0.3	0.2	0.1
5	3	3	3	3	3	3	5	5	5
10	4	4	5	5	5	5	6	7	10
50	19	19	19	20	20	22	24	27	34
100	38	38	38	38	39	43	46	53	65
500	185	185	186	189	193	210	227	257	316
1000	369	369	372	376	385	419	453	513	630

Table 2.1: Absolute values of the threshold r^* for different values of N and levels of ambiguity γ. The threshold is increasing with ambiguity. The agent waits longer before accepting a candidate when ambiguity increases.

interval near by $\frac{1}{n}$, i.e.

$$P[Y_n = 1|\mathcal{F}_{n-1}] \in \left[\frac{\gamma}{n}, \frac{1}{\gamma n}\right]$$

for appropriately chosen $\gamma < 1$[15], $2 \leq n \leq N$. The parameter γ measures the level of uncertainty, as it tends to 1 the uncertainty about the probability of finding a candidates vanishes. In the limit, for $\gamma = 1$ we are back in the classical case.

We can use the analysis of the preceding section to compute the thresholds r^* that depends on γ and N, of course. Typical values are tabulated in Table 2.1 and 2.2 for the absolute and relative values of the threshold, resp. It is interesting to see that one waits longer as ambiguity increases. The effect of missing a potentially better applicant outweighs the lower expectation from ambiguity. We get here a potentially testable implication: the more uncertain the agent is, the longer she should wait before taking a decisive action in a best choice problem.

The following result gives exact boundaries for the optimal threshold depending upon γ and N.

[15]Of course one has to choose γ in a way that ensures that $P[Y_n = 1|\mathcal{F}_{n-1}] \in (0, 1)$ for all $2 \leq n \leq N$.

N / γ	1	0.9	0.8	0.7	0.6	0.4	0.3	0.2	0.1
5	60%	60%	60%	60%	60%	60%	100%	100%	100%
10	40%	40%	50%	50%	50%	50%	60%	70%	100%
50	38%	38%	38%	40%	40%	44%	48%	54%	68%
100	38%	38%	38%	38%	39%	43%	46%	53%	65%
500	37%	37%	37%	38%	39%	42%	45%	51%	63%
1000	37%	37%	37%	38%	39%	42%	45%	51%	63%

Table 2.2: Relative values of the threshold r^* for different values of N and levels of ambiguity γ. Also the relative threshold is increasing with ambiguity.

Theorem 3. *For given $\gamma > \frac{1}{2}$ and N, the optimal threshold $r^*(\gamma, N)$ satisfies*

$$e^{-\frac{1}{\gamma}} \leq \frac{r^*(\gamma, N)}{N} \leq e^{-\frac{2\gamma}{1+\gamma}} + \frac{3}{N} . \tag{2.14}$$

In particular, the secretary problem is robust in the sense that

$$\lim_{N\to\infty, \gamma\uparrow 1} \frac{r^*(\gamma, N)}{N} = \lim_{N\to\infty} \frac{r^*(0)}{N} = e^{-1} . \tag{2.15}$$

2.3.2 Independent Coins with Identical Ambiguity

Our example corresponds to the *i*ndependent indistinguishably distributed case introduced in Epstein and Schneider (2003a). Here, the probability to meet a candidate remains constant over time. More generally, this is the case, where the decision maker does not know if the experiment changes over time. At the same time she has no reason to distinguish between periods. To express the uncertainty about the coin the agent uses a class of measures in each period.

We consider the following bet: We observe an ambiguous coin being tossed N times and we win if we stop at the last time $\{head\}$ appears in the sequence. With this setup we are in the situation of the ambiguous best choice problem where the probabilities for $\{head\}$ remain constant over time:

$$P[n\text{-th toss is a head}|\mathcal{F}_{n-1}] \in [p - \epsilon, p + \epsilon]$$

for $\epsilon \geq 0$, chosen such that $0 < p - \epsilon < p + \epsilon < 1$. To get a feeling for the problem, let us start with the pure risk case, $\epsilon = 0$. In this case, we get

$$w_r = \beta(N - r) = \frac{p}{1-p}(N - r)$$

and the optimal threshold is the first r such that

$$N - r \leq \frac{1-p}{p}\,.$$

In this problem, it is optimal to focus solely on the last $\left[\frac{1-p}{p}\right] + 1$ applicants, irrespective of the total number of applicants.

Let us now come to the ambiguous case. From Equation 2.12, we obtain for the degree of ambiguity $\alpha = \frac{1-p+\epsilon}{1-p-\epsilon} > 1$ and ambiguous odds $\beta = \frac{p-\epsilon}{1-p-\epsilon}$

$$w_r = \sum_{k=r}^{N} \beta \prod_{l=r}^{k-1} \alpha = \beta\,\frac{\alpha^{N-r+1} - 1}{\alpha - 1}\,.$$

The threshold r^* is given by the first r such that

$$\alpha^{N-r} \leq 1 + \frac{\alpha - 1}{\beta} = \frac{p+\epsilon}{p-\epsilon}\,.$$

We learn from this that the agent focuses only on the last

$$k(p,\epsilon) \simeq \frac{\log \frac{p+\epsilon}{p-\epsilon}}{\log \frac{1-p+\epsilon}{1-p-\epsilon}}$$

applicants. This quantity is independent of N.

In this case we observe *memoryless stopping*: The decision about stopping does not depend on the number of the options already observed. Only the number of options left matters. Consequently, we obtain

$$\lim_{N\to\infty} \frac{r^*(N)}{N} = 1\,.$$

This example also allows us to demonstrate that ambiguity can lead both to earlier as well as to later stopping. For $p < \frac{1}{2}$, the quantity $k(p,\epsilon)$ is increasing; consequently, the agent stops earlier when ambiguity increases. For $p = 1/2$, $k(p,\epsilon)$ is independent of ϵ and ambiguity does not influence the stopping behavior. For $p > 1/2$, the agent stops later, in general.

2.3.3 Finite Stopping

In our last example we consider the case where the probability to meet a candidate falls very fast. Here, the value of waiting decreases very fast and becomes zero at some point. In this situation the future becomes worthless and interviewing additional candidates does not improve the expected payoff. Even if the pool of applicants is infinite the decision will be made in finite time. Here, we can compute the maximal amount of applicants that need to be interviewed in order to decide optimally.

To see how it works we first consider the value of waiting for a fixed number of candidates N and a given one-step-ahead probabilities $[a_n, b_n]$. Now we add an applicant with $P[Y_{N+1} = 1|\mathcal{F}_N] \in [a_{N+1}, b_{N+1}]$. Clearly, adding applicants does not decrease the value of the problem. As we vary the number of applicants now, let us write w_r^N for the crucial sequence that determines the threshold $r^*(N)$. Clearly, w_r^N is increasing in N and the value of the threshold $r^*(N+1) \geq r^*(N)$. Now assume that $w_r^\infty := \lim_{N\to\infty} w_r^N$ exists. Then we can find $R \in \mathbb{N}$ s.t. $w_R^\infty < 1$ and therefore $w_R^N < 1$ for all N sufficiently large. Therefore, the value of the threshold $r^*(N)$ cannot exceed R. As $r^*(N)$ is an increasing, but bounded sequence of integers, it has to be constant from some point on, $r^*(N) = R$ for N sufficiently large.

In other words, the number of applicants does not matter here for large pools of applicants. The agent first studies a fixed number of applicants before taking the next candidate.

Lemma 4. *If*

$$w^\infty := \lim_{N\to\infty} w_1^N \tag{2.16}$$

exists, then

1. *The value of the threshold $r^*(N)$ is bounded by a constant $R \in \mathbb{N}$ and for sufficiently large $N \in \mathbb{N}$, we have $r^*(N) = R$,*

2. The fraction of rejected candidates converges to zero, i.e.

$$\lim_{N\to\infty} \frac{r^*(N)}{N} = 0\,.$$

Let us reflect a moment under what condition the series $w^\infty = \sum_{k=1}^\infty \beta_k \prod_{l=r}^{k-1} \alpha_l$ is finite. By d'Alembert's ratio test, this is the case if we have

$$\limsup_{n\to\infty} \frac{1-a_n}{a_n}\frac{a_{n+1}}{1-b_{n+1}} < 1\,.$$

This condition holds true, e.g., when both (a_n) and (b_n) converge fast, say exponentially, to zero.

In this section we analyzed the observation period for different sets of measures. Depending on the structure of the set $\mathcal{P}$ the observation period converges to a constant $c \in (0,1)$ as in the case of the ambiguous secretary problem. Or it can converge to zero making the future worthless as in the finite stopping case. In the opposite case of memoryless stopping the observation period tends to 1, assigning zero value to the past.

2.4 Conclusion

We provide a closed form solution for the best choice problem in the multiple prior framework. An appropriate version of backward induction leads to the solution if the set of priors is time-consistent. Due to time-consistency most of classical arguments remain valid, the stopping rule is simple. The closed form solution allows to analyze the impact of ambiguity on the stopping behavior. Additionally, we show the robustness of the classical secretary problem in the multiple prior framework. A natural next step is to generalize the utility function. Additionally, one might extend the model to infinite settings.

Chapter 3

Exercise Strategies for American Exotic Options under Ambiguity

3.1 Introduction

The increasing trade volume of exotic options both in the plain form and as component of more sophisticated products motivates the more precise study of these structures. The OTC nature of these contracts allows for almost endless variety of products which comes at the price of tractability and evaluation complexity. The payoff of the option is often conditioned on an event during the lifetime leading to a path dependent structure which is challenging to evaluate.

Most of the literature on this field concentrates on hedging or replication analyzing the hedging strategy of the seller or deriving the no arbitrage price. This analysis is sufficient in the case of European options as it also captures the problem of the buyer. However, in the case of American options the task of the buyer holding the option in her portfolio differs structurally from the hedging problem of the seller. Unlike the bank/the market the holder of the option is not interested in the risk neutral value of the option but aims to exercise the claim optimally realizing highest possible utility. This valuation in general needs not to be related to the market value of the option as it reflects the personal utility of the holder which depends on investment horizon and objectives and also on the risk attitude of the holder.

Given a stochastic model in discrete time, such as the Cox, Ross, and Rubinstein (1979) (CRR) model one can easily solve the problem of the buyer using dynamic programming. However, classical binomial tree models impose the assumption of a unique given probability measure driving the stock price process. This assumption might be too strong in several cases as it requires perfect understanding of the market structure and complete agreement on one particular model.

As an example we consider an asset manager holding an American claim in her portfolio. Her exercise decision depends on the underlying model for the stock price process derived from filtering using past stock price observations. As the model cannot be determined perfectly she faces model uncertainty. Being unable to completely specify the model the asset manager rather uses multiple prior model instead of choosing one particular model. If the uncertainty cannot be resolved and the accurate model specification is impossible more robust strategies are to be preferred as they perform well even if the model is specified slightly incorrect.[1]

Also a risk controlling unit assigning value and riskiness to the portfolio chosen by the manager uses rather a multiple prior models in order to test for model robustness and to measure model risk. Taking several models into account while performing portfolio distress tests allows to check for

[1] Several authors discussed the robust portfolio optimization problem in the multiple prior context. However, most of works on this field only consider investments in plain vanilla products.

the sensitivity of the portfolio to model misspecification. Again in a situation of model uncertainty more robust riskiness assignment is desirable as it minimizes model risk/uncertainty[2].

Similar reasoning can be applied to accounting issues. An investment funds manager making his annual valuation is interested in the value of options in the book that are not settled yet. In case the company applies coherent risk measures as standard risk evaluation tool for future cash flows on the short side, it is plausible to use a multiple prior model evaluating long positions. Finally, a private investor holding American claims in his account might exhibit ambiguity aversion in the sense of Ellsberg paradox or Knightian uncertainty. Such behavior may arise from lack of expertise or bad quality of information that is available to the decision maker.

Although for different reasons, all the market participants described above face problems that should not be analyzed in a single prior model and can be formulated as multiple prior problems. In this paper we analyze the problem of the holder of an American claim facing model uncertainty that results in a multiple prior model. We characterize optimal stopping strategies for the buyer that assesses utility to future payoffs in terms of minimal expectation and study how the multiple prior structure affects the stopping behavior.

Multiple prior models have gained much attention in recent studies. Hansen and Sargent (2001) considered the multiple prior models in the context of robust control, Karatzas and Zamfirescu (2003) approached the problem from game theoretical point of view. Delbaen (2002) introduced the notion of coherent risk measures which mathematically corresponds to the approach used in this paper.

The decision theoretical model of multiple priors was introduced by Gilboa and Schmeidler (1989) and further developed to dynamical settings by Epstein and Schneider (2003b). The methods we use in this paper rely heavily

[2]See Cont (2006) for an extensive discussion on the issue of model risk

on these works.
Epstein and Schneider (2003a) applied the multiple prior model to financial markets and Epstein and Schneider (2003b) addressed the question of learning under uncertainty. Riedel (2009) considered the general task to optimally stop an adapted payoff process in a multiple prior model and showed that backward induction fails in general. He imposed more structure on the set of priors that ensured that the solution can be found using an appropriate form of backward induction. The cornerstone of the method is the time-consistency of the set of priors which allows the decision maker to change her beliefs about the underlying model as the time evolves. If the set of priors is time-consistent one can proceed as in the classical case[3] computing the value process of the stopping problem – the multiple prior Snell envelope. It is then optimal to stop as soon as the payoff process reaches the value process. Additionally, the ambiguous optimal stopping problem corresponds to a classical optimal stopping problem for a measure $\hat{P}$ – the so-called worst-case measure[4].

As an application of the technique Riedel (2009) solves the exercise problem for the buyer of an American put and call in discrete time. A similar problem was analyzed by Nishimura and Ozaki (2007) and later by Kort and Trojanowska (2007), they considered the optimal investment decision for a firm in continuous time with infinite time horizon under multiple priors which can be related to the perpetual American call. In this paper we follow the lines of Riedel (2009) and analyze several exotic options that have a second source of uncertainty from the perspective of the buyer in a multiple prior setting. We focus on the discrete time version of the problem and develop an ambiguous version of the CRR model. Instead of assuming that the distribution of up- and down- movements of the underlying is known to the buyer we allow the probability of going up on a node to lie in a appropriately modeled

[3]See Snell (1952),Chow, Robbins, and Siegmund (1971) for more detailed analysis.
[4]See Riedel (2009), Föllmer and Schied (2004), Karatzas and Kou (1998).

set[5].

This leads to a set of models that agree on the size of up- and down-movement but disagree on the mean return. In this ambiguous binomial tree setting which was first analyzed in Epstein and Schneider (2003a) we aim to apply standard Snell reasoning to evaluate the options. Due to the above mentioned duality result it is enough to calculate the worst-case measure $\hat{P}$ and then to analyze the classical problem under $\hat{P}$. However, the worst-case measure depends highly on the payoff structure of the claim and needs to be calculated for each option separately. If the payoff satisfies certain monotonicity conditions the worst-case measure is easy to derive. The direction/effect of uncertainty is the same for all states of the world and the worst-case measure is then independent on the realization of the stock price process leading to a statical structure that resembles classical models. In the case of more sophisticated payoffs this stationarity of the worst-case measure breaks down and the worst-case measure changes over time depending on the realization of the stock price process. This is due to the fact that uncertainty may affect the model in different ways changing the beliefs of decision maker and so the worst-case measure according to the effect that is dominating. This ability to react to information by adjusting the model and to choose the model depending on the payoff is the main structural difference between the classical single measure model and the multiple prior model considered here.

We identify additional sources of uncertainty that lead to the dynamical and path-dependent structure of the worst-case measure. We also analyze the impact of different effects of uncertainty on the overall behavior and the resulting model highlighting differences between the single prior models used in the literature and the multiple prior model introduced here.

In our analysis we decompose the claims in monotone parts as the worst-case measure for monotone problems is well known. We then analyze each

[5] Alternatively one might assume that the probability of an up-movement is known while the size of the increment is not. This approach turns out to be equivalent to ours.

claim separately deriving the worst-case measure conditioned on monotonicity. To complete the analysis we paste the measures obtained on subspaces together using time-consistency. This idea is closely linked to the method of pricing derivatives using digital contracts introduced by Ingersoll (2007) and also used by Buchen (2004). However, this literature focuses on European style options and does not cover the dynamical structure analyzed here.

In the case of barrier options the value of the option is conditioned on the event of reaching a trigger. Unlike the plain vanilla option case, the lifetime of an barrier option become uncertain as it depends on the occurrence of the trigger event. This leads to an additional source of uncertainty causing a change in the monotonicity of the value function when the stock price hits the barrier. For example, in the case of an up-and-in put the ambiguity averse decision maker assumes the returns to be low and chooses therefore the measure with the lowest drift before the stock price reaches the barrier. After hitting the barrier she obtains a plain vanilla put option monotone in the underlying and uses therefore the measure with the highest drift. Similar behavior can be observed for other types of barrier options.

The second group of options we focus on are the dual expiry options. Here, the strike of the option is not known at time zero as it is being determined as a function of the underlying's value on a date different from the issue date of the option – the first expiry. Therefore, additional to the uncertainty about the final payoff the decision maker faces uncertainty about the value of the strike before first expiry date.

In the case of shout options the first expiration date, the so-called shout date/freeze date, is determined by the buyer. Here, the investor has to call the bank if she aims to fix the strike. Therefore, the buyer of an shout option faces two stopping problems: First, she has to determine the optimal shouting time in order to set the strike optimally and then the to stop the payoff process optimally. The holder of an shout put gets an put after shouting and thus, anticipates high returns on the stock after shouting. Before shouting however he owns a claims whose value is increasing in the price of the underlying which results in low returns anticipated before shouting.

Finally, we analyze options whose payoff function consists of two monotone pieces. Typical examples are straddles and strangles. The buyer of such options presumes a change in the underlying's price but is not sure about the direction of the change. Depending on the value of the underlying the option pays off a call or a put, so as a consequence the actual payoff function becomes uncertain. Here, one can decompose the value of the option in an increasing and a decreasing leg. The buyer of the option changes her beliefs about the returns every time the value switches from decreasing to increasing part of the value function. So, an ambiguity averse buyer of a straddle presumes the stock price to go down in hausse phases and up in baisse phases.

An outline of the paper is as follows. Section 2 introduces the discrete model which is in this form due to Riedel (2009). Section 3 recalls the solution for payoffs monotone in underlying's price introduced in Riedel (2009) and builds the base for the following analysis. Section 4 provides the solution for barrier options options, and Section 5 develops the solution for multiple expiry. Finally, Section 6 discusses U–shaped payoffs.

3.2 Financial Markets and Optimal Stopping under Ambiguity in discrete Time

We first introduce the basic theoretical setup to evaluate options in multiple prior model. This model has the CRR model as the starting point and was already developed in Riedel (2009) and can be seen as a version of the IID model introduced in Epstein and Schneider (2003a) with a different objective. At the same time the model is the discrete time version of the κ–ambiguity model in Epstein and Chen (2002).

Having established the model we discuss the market structure and recall the decision problem of the buyer and the solution method – the multiple prior backward induction introduced by Riedel (2009).

3.2.1 Stochastic Structure

To set up the model we start with a classical binomial tree. For a fixed maturity date $T \in \mathbb{N}$ we consider a probability space $(\Omega, \mathcal{F}, P_0)$ where $\Omega = \otimes_{t=1}^{T} \mathcal{S}$ with $\mathcal{S} = \{0, 1\}$ is the the set of all sequences with values in $\{0, 1\}$, $\mathcal{F}$ is the σ–field generated by all projections $\epsilon_t : \Omega \to \mathcal{S}$ and P_0 denotes the uniform on $(\Omega, \mathcal{F})$. By construction, the projections $(\epsilon_t)_{t=1,...,T}$ are independent and identically distributed under P_0 with $P_0(\epsilon_t = 1) = \frac{1}{2}$ for all $t \leq T$. Furthermore, we consider the filtration $(\mathcal{F}_t)_{t=0,...,T}$ generated by the projections $(\epsilon_t)_{t=1,...,T}$ where $\mathcal{F}_0$ is the trivial σ–field – $\{\emptyset, \Omega\}$. The event $\epsilon_t = 1$ represents an up-movement on a tree while the complementary event denotes the down-movement[6].

Additionally, we define a convex set of measures $\mathcal{Q}$ in the following way: We fix an interval $[\underline{p}, \overline{p}] \subset (0, 1)$ for $\underline{p} \leq \overline{p}$ and consider all measures whose conditional one–step–ahead probabilities, i.e. the conditional probability of going up on a node of the tree remain within the interval $[\underline{p}, \overline{p}]$ for every $t \leq T$, i.e.

$$\mathcal{Q} = \big\{ P \in \mathcal{M}_1(\Omega) | P(\epsilon_t = 1 | \mathcal{F}_{t-1}) \in [\underline{p}, \overline{p}], \forall t \leq T \big\} \qquad (3.1)$$

The set $\mathcal{Q}$ is generated by the conditional one-step-ahead correspondence assigning at every node $t \leq T$ the probability of going up. In particular, $\mathcal{Q}$ contains all product measures defined via $P_p(\epsilon_{t+1} = 1 | \mathcal{F}_t) = p$ for a fixed $p \in [\underline{p}, \overline{p}]$ and all $t < T$. In the following we denote by $\overline{P} = P_{\overline{p}}$ and by $\underline{P} = P_{\underline{p}}$.

Clearly, the state variables $(\epsilon_t)_{t=1,...,T}$ are independent under all product measures, correlated in general, however. To see this consider the measure

[6] We assume $\mathcal{F}_{T+1} := \mathcal{F}_T$ and $\inf \emptyset := \infty$.

P^τ defined via
$$P^\tau(\epsilon_{t+1} = 1|\mathcal{F}_t) = \begin{cases} \overline{p} & \text{if } t \leq \tau \\ \underline{p} & \text{else} \end{cases}$$
for a stopping time $\tau \leq T$. As the one-step-ahead probabilities remain in the interval $[\underline{p}, \overline{p}]$ the so defined measure P^τ belongs to $\mathcal{Q}$ for all stopping times $\tau \leq T$. At the same time the probability of going up on a node depends on the realized path through the value of τ and $(\epsilon_t)_{t=1,\dots,T}$ are correlated.

The above example reveals an important structural feature of $\mathcal{Q}$: The set of measures is stable under the operation of decomposition in marginal and conditional part. Loosely speaking, it allows the decision maker to change the measure she uses as the time evolves in an appropriate manner. In the example above, the decision maker first uses the measure $\overline{P}$ until an event indicated by the stopping time τ and then changes to $\underline{P}$. Mathematically, this property is equivalent to an appropriate version of the Law of Iterated Expectation and is closely linked to the idea of backward induction. The concept has gained much attention in the recent literature and was also discussed under different notions by Delbaen (2002), Epstein and Schneider (2003a), Föllmer and Schied (2004) and Riedel (2009).

The following lemma summarizes crucial properties of the set $\mathcal{Q}$.

Lemma 5. *The set of measures defined as in* (3.1) *satisfies the following properties*

1. *$\mathcal{Q}$ is compact and convex,*
2. *all $P \in \mathcal{Q}$ are equivalent to P_0,*
3. *$\mathcal{Q}$ is time-consistent in the following sense: Let $P, Q \in \mathcal{Q}$, $(p_t)_t, (q_t)_t$ densities of P, Q with respect to P_0. For a fixed stopping time τ define the measure R via*
$$r_t = \begin{cases} p_t & \text{if } t \leq \tau \\ \frac{p_\tau q_t}{q_\tau} & \text{else} \end{cases}$$
then $R \in \mathcal{Q}$.

Due to Lemma 5 we can identify the set $\mathcal{Q}$ with the set of the density processes with respect to the measure P_0. In the following we denote by D the density process of $P \in \mathcal{Q}$ with respect to P_0, i.e. $D_t = \frac{dP}{dP_0}|_{\mathcal{F}_t}$ for $P \in \mathcal{Q}$, $t \leq T$. A more detailed analysis of the structure of D can be found in Riedel (2009).

3.2.2 The Market Model

Within the above introduced probabilistic framework we establish the financial market in the spirit of the CRR model. We consider a market consisting of two assets: a riskless bond with a fixed interest rate $r > -1$ and a risky stock with multiplicative increments. For given model parameters $0 < d < 1 + r < u$ and $S_0 > 0$ the stock S evolves according to

$$S_{t+1} = S_t \cdot \begin{cases} u & \text{if } \epsilon_{t+1} = 1 \\ d & \text{if } \epsilon_{t+1} = 0 \end{cases} .$$

Without loss of generality, we assume $u \cdot d = 1$[7].Then, for every $t \leq T$ the range of possible stock prices is finite and bounded, we denote by

$$E_t = \{S_0 \cdot u^{t-2k} | k \in \mathbb{N}, -t \leq k \leq t\}$$

the set of possible stock prices at time t. Moreover, the filtration generated by the sequence $(S_t)_{t=0,\dots,T}$ coincides with $(\mathcal{F}_t)_{t=0,1,\dots,T}$ and every realized path $(s_1, \dots, s_t)$ of S can be associated with a realization of $(\epsilon_s)_{s \leq t}$.

As the state variables are not independent under every probability measure $P \in \mathcal{Q}$ in our model the increments of S are correlated in general. The probability of an up-movement depends on the realized path but stays within the boundaries $[\underline{p}, \overline{p}]$ for every $P \in \mathcal{Q}$.

Economically, our model describes a market where the market participants are not perfectly certain about the asset price dynamics. In order to

[7]This is a common assumption when dealing with exotic options in binomial models, see Cox and Rubinstein (1985) for a detailed discussion.

express this uncertainty investors use a class of measures constructed above. The set $\mathcal{Q}$ is the set of possible models the decision maker takes into account. Different choices of $P \in \mathcal{Q}$ correspond to different models. With our specification mean return on stock is uncertain and as one can easily see, $\overline{P}$ corresponds to the highest mean return at every node, while $\underline{P}$ corresponds to the lowest mean return on stock on every node. The specification of $\mathcal{Q}$ is a part of the model and in practice may arise from regulation policies or be imposed by the bank accounting standards, result from statistical consideration or just reflects the degree of ambiguity aversion. The length of the interval $[\underline{p}, \overline{p}]$ determines the range of possible models. As the interval decreases the model converges to the classical binomial tree model and we obtain the classical CRR model as a special case of our model by choosing $\underline{p} = \overline{p}$.

The use of a set of models especially allows for correlated returns. This gives the decision maker the possibility to adjust the models as the path is realized and new information arrives. Now the economical implication of time-consistency of $\mathcal{Q}$ becomes clear. Due to this property the multiple prior decision maker is allowed to use the measure $P_1 \in \mathcal{Q}$ until an event indicated by a stopping time τ and then to change his beliefs about the right model using P_2 after τ. The multiple prior decision maker is allowed to adjust the model she uses responding to the state of the market. However, this notion is not the same as classical Bayesian learning as the decision maker has too little information or market knowledge to learn the real distribution. While in the learning process the decision maker updates the model adjusting the set of possible models, here the investor keeps the set of possible models fixed not excluding any of the possible models as the time evolves but choses a particular model at every point of the time reconsidering her choice when new information arrives.

3.2.3 The Decision Problem

In this setting we consider an investor holding an exotic option in her portfolio. As most of the exotic options are OTC[8] contracts there is usually no functioning market for these derivatives or the trading of claims involves high transaction costs. Therefore, in absence of a trading partner the buyer is forced to hold the claim until maturity, so we exclude the possibility of selling the acquired contracts concentrating purely on the exercise decision of the investor. In our analysis we mainly concentrate on institutional investors already holding the derivatives in the portfolio. Therefore, it is plausible to assume risk neutral agents who discount future payoff by the riskless rate.

Remark 1. *When having an ambiguity averse private investor in mind it seems natural also to introduce risk aversion and to discount by individual discount rate δ. As these considerations do not change the structure of the worst-case measure obtained here, we omit this possibility maintaining risk neutrality.*

We consider an American claim $A : \Omega \to \mathbb{R}_+$ written on the asset S and maturing at T that pays off $A(t, (S_s)_{s \leq t})$ when exercised at time t. The investor holding A in her portfolio aims to maximize her expected payoff by choosing an appropriate exercise strategy, i.e. the best time to exercise the contract. As the expectation in our multiple prior setting is not uniquely defined the ambiguity averse decision maker maximizes her minimal expected payoff, i.e.

$$\text{maximize } \inf_{P \in \mathcal{Q}} E^P A(\tau, (S_s)_{s \leq \tau}) \text{ over all stopping times } \tau \leq T. \tag{3.2}$$

The choice of the exercise strategy according to the worst possible model corresponds to conservative value assignment. It treats long book positions

[8]OTC deals are contracts that are traded *o*ver the counter with a counterparty and not through a centralized liquid trade exchange.

in the same way as the coherent risk measures treats short positions[9]. The value of the multiple prior problem $U^{\mathcal{Q}}$ stated in (3.2) is lower or equal than the value of the problem U^P for every possible model $P \in \mathcal{Q}$. Therefore, this notion minimizes the model risk as the model misspecification within $\mathcal{Q}$ increases the value of the claim.

Remark 2. 1. *The problem of the long investor stated in* (3.2) *differs structurally from the task of the seller of the option. The seller of the American claim needs to hedge claim against every strategy of the buyer. To obtain the hedge she solves the optimal stopping problem under the equivalent martingale measure P^*. In the binomial tree the unique equivalent martingale measure P^* is completely determined by parameters r, u and d and does not depend on the mean return. See Hull (2006) for a more detailed analysis. The situation is different for the buyer as she solves the optimal stopping problem under the physical measure taking the mean return into account and being interested in personal utility maximization rather than in risk neutral valuation. Although the buyer and the seller use different techniques assigning value to the options and obtaining different values for the claim there is no contradiction to no arbitrage condition because of the American structure of the claims considered here.*

2. *It is usual to evaluate claims in the book that are not settled yet using mark-to-market approach. The value of the option is then set to be equal to the market price. This makes sense if markets are well functioning or if the investor intends to sell the option on the secondary market rather than hold it until maturity. However, this approach may value the claims wrongly in times of distressed markets or if there is no market for the security at all as it was seen and still is seen at financial markets these days. Multiple prior value assignment through $U^{\mathcal{Q}}$ is an*

[9] Mathematically, our model is equivalent to a representation of coherent risk measures. See Delbaen (2002) or Riedel (2009) for more detailed analysis.

alternative to the mark-to-market accounting as it provides conservative value assignment by using the worst possible scenario but protects the book value from too pessimistic or overoptimistic views of the market that are due to expectations and do not reflect fundamentals. However, $U^{\mathcal{Q}}$ is not the price for the option but rather the private value for the investor that may differ from the market view.

3.2.4 The Solution Method

If $\mathcal{Q}$ is a singleton the problem stated in (3.2) can be solved using classical dynamic programing methods. One defines backwards the value process of the problem – the Snell envelope – and stops as soon as the value process reaches the payoff process. This technique fails to hold in the multiple prior setting[10]. Riedel (2009) extended backward induction to the case of time-consistent multiple priors stating sufficient conditions for the Snell arguments to hold.

Theorem 4 (Riedel (2009))**.** *Given a set of measures satisfying conditions stated in Lemma 5 and a payoff process $X = (X_t)_{t=0,\dots,T}$ bounded by a $\mathcal{Q}$-uniformly integrable variable Z, define the multiple prior Snell envelope $U^{\mathcal{Q}}$ recursively by*

$$\begin{aligned} U_T^{\mathcal{Q}} &= X_T \\ U_t^{\mathcal{Q}} &= \max\{X_t, \inf_{P\in\mathcal{Q}} \mathbb{E}^P(U_{t+1}^{\mathcal{Q}}|\mathcal{F}_t)\} \text{ for } t < T \end{aligned} \tag{3.3}$$

Then,

1. *$U^{\mathcal{Q}}$ is the smallest multiple prior $\mathcal{Q}$-supermartingale [11] dominating the payoff process X.*

[10] See Riedel (2009) for an example.

[11] Given a set of measures $\mathcal{Q}$, a multiple prior supermartingale with respect to $\mathcal{Q}$ is an adapted process, say S, satisfying $S_t \geq \text{ess}\inf_{P\in\mathcal{Q}} \mathbb{E}^P(S_{t+1}|\mathcal{F}_t)$ for $t \in \mathbb{N}$.

2. *$U^{\mathcal{Q}}$ is the value process of the multiple prior stopping problem for the payoff process X, i.e.*

$$U_t^{\mathcal{Q}} = \sup_{\tau \geq t} \inf_{P \in \mathcal{Q}} \mathbb{E}^P(X_\tau | \mathcal{F}_t)$$

3. *An optimal stopping rule is then given by*

$$\tau^{\mathcal{Q}} = \inf\{t \geq 0 | U_t^{\mathcal{Q}} = X_t\}.$$

The above result ensures the existence of the solution of problem (3.2). Moreover, as shown by several authors (for example Föllmer and Schied (2004), Karatzas and Kou (1998), Riedel (2009)) problem (3.2) is equivalent to a single prior problem for a measure $\hat{P} \in \mathcal{Q}$, i.e. the value function of the multiple prior problem

$$U^{\mathcal{Q}} = U^{\hat{P}} \qquad P_0\text{-a.s.}. \tag{3.4}$$

The measure $\hat{P}$ is called worst–case measure and can be constructed via backward induction by choosing the worst one-step-ahead probability at every node of the tree and pasting the so obtained densities together at time 0. The worst-case measure is stochastic in general and depends on the payoff process.

Due to equality (3.4) the optimal stopping strategies $\tau^{\mathcal{Q}}$ of the multiple prior problem and $\tau^{\hat{P}}$ of the problem for the prior $\hat{P}$ coincide. Therefore, the problem can be solved in two steps. In the first step one identifies the worst–case measure $\hat{P}$ and solves the classical problem under $\hat{P}$ in the second step. This technique allows to make use of solutions already obtained in the classical case. For problems where no closed form solution is available the technique reduces numerical complexity by reducing the task to a single model problem where methods are well developed.

Analyzing the exotic options we use this techniques emphasizing the behavioral interpretation of the worst–case measure and highlighting the difference between classical models and the multiple prior approach.

3.2.5 Options with Monotone Payoffs

In this subsection we provide the solution for claims whose payoffs are monotone in the underlying's price at each time and satisfy the Markov property. The results of this section build the foundation for the analysis of more complicated payoffs in the next sections and were stated in Riedel (2009).

We consider a discounted American claim maturing at T and paying off $X_t = A(t, S_t)$ when exercised at t.

Theorem 5 (Riedel (2009))**.** *1. If the payoff function of the claim $A(t, S_t)$ is increasing in S_t for all t, then the multiple prior Snell envelope is*

$$U^{\mathcal{Q}} = U^{\underline{P}}$$

where $\underline{P}$ is the measure defined by the density

$$\underline{D}_t = 2^t \prod_{u \le t} \underline{p}\epsilon_u + (1 - \underline{p})(1 - \epsilon_u)$$

and the holder of the claim uses the optimal stopping rule given by

$$\tau^{\underline{P}} = \inf\{t \ge 0 : A(t, S_t) = U_t^{\underline{P}}\}.$$

2. If $A(t, S_t)$ is decreasing in S_t for all t, the multiple prior Snell envelope is

$$U^{\mathcal{Q}} = U^{\overline{P}}$$

where $\overline{P}$ is the measure defined by the density

$$\overline{D}_t = 2^t \prod_{u \le t} \overline{p}\epsilon_u + (1 - \overline{p})(1 - \epsilon_u)$$

and an optimal stopping rule under ambiguity is given by

$$\overline{\tau} = \inf\{t \ge 0 : A(t, S_t) = U_t^{\overline{P}}\}.$$

The key to this result is the fact that $\underline{P}$ (or $\overline{P}$ resp.) is the worst probability measure in the sense of first–order stochastic dominance and that the payoff is a monotone function of the underlying stock price. These results help us finding the worst-case measure for more complicated payoffs. Using appropriate decompositions we represent the options as monotone claims. For those monotone claims we can identify the worst-case measure using Theorem 5. Pasting the so obtained measures together we construct the desired worst-case measure.

3.3 Barrier Options

Barrier options are among most traded exotic options and are often used as components of more sophisticated structured derivatives. The payoff of a barrier claim depends on the stock price reaching the barrier prespecified in the contract: *in* options are activated when the stock price hits the barrier while *out* options become worthless at the barrier. Depending on the direction of the stock price change one distinguishes between *up* and *down* options indicating whether the stock price reaches the barrier from below or above. Combining barrier features yields four different barrier types each of which helps expressing different views on the market. Once activated an in-barrier option is then a plain vanilla call or put with known payoff structure[12].

The knock-in/knock-out feature of the options lowers the price which has to be paid by the buyer. In return, the buyer is exposed to the risk of the barrier event that makes the option worthless. This singularity of the payoff at the barrier makes the barrier option interesting from the mathematical point of view and challenging to evaluate.

Unlike plain vanilla options discussed in the previous section barrier options are path-dependent as the payoff depends on the stock process reaching the barrier. Thus, the Markovian reasoning that was essential for the results cannot be applied directly. However, conditioned on the event of reaching the

[12]Combining put and call payoffs with the four different barrier types introduced above gives eight different barrier options.

barrier, the option is Markovian again. Also the second important ingredient for the above result – the monotonicity of the payoff at every node fails to hold here in general. To overcome this difficulty we extend the Theorem 2 to hold not for all $t \leq T$ but on stochastic intervals.

To formalize the ideas stated above mathematically we denote by

$$\tau_H^u : \Omega \longrightarrow [0, T+1], \qquad \tau_H^u(\omega) := \inf\{t \geq 0 : S_t(\omega) \geq H\} \wedge T + 1.$$

the first upcrossing time at H for a given barrier[13] $H \in E_t$ for some $t \leq T$. τ_H^u is the first time the stock price reaches the barrier H from below and will be important for up-options. Similarly we define τ_H^d – *the first downcrossing time at* H useful for down-options. While the knock-in event activates the option the knock-out event terminates it. Thus, the life time of the option is determined by a stochastic event marked by above defined stopping times and is not known from ex ante perspective. The time frame in which the option can be exercised is then given by a stochastic interval $[\tau_1, \tau_2[$ defined by:

$$[\tau_1, \tau_2[\ := \left\{(s, \omega) \in [0, T] \times \Omega \,\middle|\, \tau_1(\omega) \leq s < \tau_2(\omega)\right\}.$$

We will often write with slight abuse of notation $\mathbb{1}_{[\tau_1,\tau_2[}(\omega)$ instead of $\mathbb{1}_{[\tau_1,\tau_2[}(t, \omega)$.

With notations introduced above we are now ready to state the result that will be the key for the analysis of payoffs with barrier feature.

Theorem 6. *Let $H_1 < H_2$ be barriers defining first upscrossing times τ_1^u and τ_2^u, respectively. Consider a bounded payoff process $X = (X_t)_{0,\dots,T}$ defined by*

$$X_t = x(t, S_t, \tau_1^u, \tau_2^u) = A(t, S_t)\, \mathbb{1}_{[\tau_1^u, \tau_2^u[}(t)$$

1. *If $A(t, \cdot)$ is decreasing in S_t for all $t \leq T$, the multiple prior Snell envelope is*

$$U^{\mathcal{Q}} = U^{\hat{P}}$$

[13]Here, we assume that the barrier lies in the set of possible stock prices for some $t \leq T$. This assumption is not crucial and can be relaxed easily, see Hull (2006) for more detailed review.

where a worst-case measure $\hat{P}$ is defined by the density

$$\hat{D}_t := 2^t \prod_{u \le t \wedge \tau_1^u} \big(\varepsilon_u \underline{p} + (1-\varepsilon_u)\big)(1-\underline{p}) \prod_{u \in]\tau_1^u, t]} \big(\varepsilon_u \overline{p} + (1-\varepsilon_u)(1-\overline{p})\big)$$

for all $t \le T$. The holder of the claim uses the optimal stopping rule given by

$$\hat{\tau} = \inf \left\{ t \in [\tau_1^u, \tau_2^u[\;\middle|\; X_t = U_t^{\hat{P}} \right\}.$$

2. *If $A(t,\cdot)$ is increasing in S_t for all $t \le T$, the multiple prior Snell envelope is*

$$U^{\mathcal{Q}} = U^{\hat{P}}$$

and the worst-case measure $\hat{P}$ is defined by the density

$$\hat{D}_t := 2^t \prod_{u \le t, \tau_2^u \wedge t \text{ or } S_u < H_2 \cdot d} \big(\varepsilon_u \underline{p} + (1-\varepsilon_u)(1-\underline{p})\big) \prod_{u \le t, S_u \ge H_2 \cdot d} \big(\varepsilon_u \overline{p} + (1-\varepsilon_u)(1-\overline{p})\big)$$

for all $t \le T$. An optimal stopping rule under ambiguity is given by

$$\hat{\tau} = \inf \left\{ t \in [\tau_1^u, T] \;\middle|\; X_t = U_t^{\hat{P}} \right\}.$$

The proof will be given in appendix.

Remark 3. 1. *Note that the worst-case measure is not unique since every measure is optimal on events where the option has knocked out already. In the theorem above we arbitrarily picked a measure on events, where the payoff is constant.*

2. *A similar theorem that provides the worst-case measure for claims conditioned on downcrossing times is stated and proven in the Appendix. Also mixtures of both theorems including up- and downcrossing times can be stated easily. Finally, analogous results hold for claims that consist of sequences of barrier options. We will use the last result while analyzing ladder options.*

3.3.1 Simple Barrier Options

With the above theorem the worst-case measures for barrier options follow as simple corollaries. Here, we first analyze in-options. In-options need to reach a barrier before they can be exercised. Thus, additional to the stock price uncertainty there is knock-in uncertainty since there is a risk that the option never knocks in leaving the buyer without a chance to exercise the claim. We can see in options as claims having a fixed maturity but a stochastic start determined by the knock-in event.

We start our discussion with the up-and-in put analyzing it in detail and sketching the proof. However, we will omit the proofs and keep the discussions short for remaining options as the reasoning is similar. An American up-and-in put with strike K, and maturity T needs to be knocked-in from below at the barrier H before it can be exercised by the buyer. When exercised it pays

$$X_t = (K - S_t)^+/(1+r)^t \mathbb{1}_{[\tau_H^u, T]}$$

for all $t \in [0, T]$.

Corollary 1 (Up-and-in put)**.** *For an American up-and-in put option with data as specified above the ambiguity averse agent uses the measure $\hat{P}$ defined by the density*

$$\hat{D}_t := 2^t \prod_{u \leq t \wedge \tau_H^u} \big(\varepsilon_u \underline{p} + (1-\varepsilon_u)(1-\underline{p})\big) \prod_{u \in]\tau_H^u, t \wedge T]} \big(\varepsilon_u \overline{p} + (1-\varepsilon_u)(1-\overline{p})\big) \quad \textit{for } t \leq T.$$

Hence, the value of the option at time t from the perspective of the ambiguity averse buyer is given by

$$U_t^{\mathcal{Q}} = U_t^{\hat{P}} = E^{\hat{P}}[X_{\hat{\tau}} \mid \mathcal{F}_t], \tag{3.5}$$

where $\hat{\tau}$ is an optimal stopping time given by

$$\hat{\tau} = \inf \left\{ t \in [\tau_H^u, T] \;\middle|\; X_t = U_t^{\hat{P}} \right\}.$$

Proof. We apply Theorem 6 part 1. Set $\tau_1^u := \tau_H^u$ and $\tau_2^u := T+1$. The discounted payoff process is given by $X_t = (K-S_t)^+/(1+r)^t \mathbb{1}_{[\tau_H^u, T+1[}$ for all $t \in [0,T]$. Since $A(t,S_t) := (K-S_t)^+/(1+r)^t$ is monotone decreasing in S_t for each t, Theorem 6 Part 1 applies. □

As a consequence the value of the claim is given by

$$\begin{aligned} U_t^Q = E^{\hat{P}}[X_{\hat{\tau}}|\mathcal{F}_t] &= E^{P_0}\left(X_{\hat{\tau}} \frac{\hat{D}_{\hat{\tau}}}{\hat{D}_t} \middle| \mathcal{F}_t \right) \\ &= E^{P_0}\left[E^{P_0}\left(X_{\hat{\tau}} \frac{\hat{D}_{\hat{\tau}}}{\hat{D}_{\tau_H^u}} \middle| \mathcal{F}_{\tau_H^u} \right) \frac{\hat{D}_{\tau_H^u}}{\hat{D}_t} \middle| \mathcal{F}_t \right] \\ &= E^{\underline{P}}\left[E^{\overline{P}}\left[X_{\hat{\tau}} | \mathcal{F}_{\tau_H^u} \right] \big| \mathcal{F}_t \right], \end{aligned} \tag{3.6}$$

Especially at $t=0$ we get

$$\begin{aligned} U_0^Q &= E^{\underline{P}}\left[E^{\overline{P}}\left[X_{\hat{\tau}} | \mathcal{F}_{\tau_H^u} \right] \big| \mathcal{F}_0 \right] \\ &= E^{\underline{P}}\left[E^{\overline{P}}[(K-S_{\hat{\tau}})^+/(1+r)^{\hat{\tau}-\tau_H^u} \big| \mathcal{F}_{\tau_H}]/(1+r)^{\tau_H^u} \right] \\ &= E^{\underline{P}}\left(\sum_{i=0}^{T} E^{\overline{P}}_{\{\tau_H^u = i\}}[(K-S_{\hat{\tau}})^+ \mathbb{1}_{\{\tau_H^u = i\}}/(1+r)^{\hat{\tau}-\tau_H^u}]/(1+r)^{\tau_H^u} \right) \\ &= \sum_{i=0}^{T} E^{\overline{P}}_{\{\tau_H^u = i\}}[(K-S_{\hat{\tau}})^+/(1+r)^{\hat{\tau}-i}]/(1+r)^i \underline{P}(\tau_H^u = i) \\ &= \sum_{i=0}^{T} E^{\overline{P}}_{\{\tau_H^d = i\}}[(K-S_{\hat{\tau}})^+/(1+r)^{\hat{\tau}-i}]/(1+r)^i \frac{J_H}{i} \binom{i}{\frac{i+J_H}{2}} \underline{p}^{\frac{i+J_H}{2}} (1-\underline{p})^{\frac{i-J_H}{2}}, \end{aligned} \tag{3.7}$$

where J_H is the positive integer such that $H = S_0 u^{J_H}$. For a derivation of the formula used in the last line see Feller (1968). The expectation in the last line denotes the value under ambiguity of an American plain vanilla put starting at time i with initial price of the underlying $S_i = H$.

The buyer of an up-and-in put uses a measure that is stochastic and has a non-stationary structure. She changes her belief about the stock return when the underlying hits the barrier anticipating low returns before the knock-in and high returns thereafter. The economic reason for this intuitively appealing behavior is the presence of an additional uncertainty caused by the knock-in. Before that there is uncertainty about whether and when the option can be exercised. Once it is resolved there is uncertainty about the final payoff as in the plain vanilla case. These uncertainty effects work in opposite direction causing a change of the measure once the first effect disappears. As a result the worst-case measure is a pasting of $\underline{P}$ and $\overline{P}$ at τ_H^u and thus cannot be fixed a priori.

Similar reasoning applies to a down-and-in call with barrier $H < S_0$ paying of

$$X_t = (S_t - K)^+/(1+r)^t \, \mathbb{1}_{[\tau_H^d, T]}$$

for all $t \leq T$:

Corollary 2 (Down-and-in call). *The ambiguity averse agent uses the following prior $\hat{P}$ given by the density*

$$\hat{D}_t := 2^t \prod_{u \leq t \wedge \tau_H^d} \big(\varepsilon_u \overline{p} + (1-\varepsilon_u)(1-\overline{p})\big) \prod_{u \in]\tau_H^d, \, t \wedge T]} \big(\varepsilon_u \underline{p} + (1-\varepsilon_u)(1-\underline{p})\big) \quad \text{for } t \leq T.$$

Similar to an up-and-in barrier put option a down-and-in barrier call equals a plain vanilla call option once the underlying has hit the barrier level H. As in (3.6) we can derive an analogous formula for the value process $(U_t^{\mathcal{Q}})_{t=0,\dots,T}$ of the down-and-in call option. For $t \leq \tau_H^d$ we obtain

$$U_t^{\mathcal{Q}} = E^{\overline{P}} \left[E^{\underline{P}} \left[X_{\hat{\tau}} | \, \mathcal{F}_{\tau_H^d} \right] \big| \, \mathcal{F}_t \right], \tag{3.8}$$

where $\hat{\tau}$ is an optimal stopping time for this considered problem under the measure $\hat{P}$. Assuming $\underline{p}u + (1-\underline{p})d \geq 1 + r$ we get that $\hat{\tau} = T$ is optimal, see Corollary 4.6 in Riedel (2009).

The situation changes if one considers out-options. Here, the option can be exercised by the buyer immediately after it is issued. However, once the knock-out level is reached the buyer forfeits the exercise right. Here, the issuance date of the option is fixed while the expiration is stochastic.

For an up-and-out call paying of $X_t := (S_t - K)^+/(1+r)^t\,\mathbb{1}_{[0,\tau_H^u[}$ for all $t \leq T$ we then get

Corollary 3 (Up-and-out call)**.** *The ambiguity averse buyer of an up-and-out call uses the measure $\hat{P}$ defined by the density*

$$\hat{D}_t := 2^t \prod_{u\leq t, u\leq \tau_H^u\wedge t \text{ or } S_u<H\cdot d} \left(\varepsilon_u \underline{p} + (1-\varepsilon_u)(1-\underline{p})\right) \prod_{u\leq t, S_u\geq H\cdot d} \left(\varepsilon_u \overline{p} + (1-\varepsilon_u)(1-\overline{p})\right)$$

for all $t \leq T$. In particular,

$$\tau := \inf\left\{t < \tau_H^u | S_t \geq \frac{H\cdot d}{(1+r)^{T-t}} + K\left(1 - \frac{1}{(1+r)^{T-t}}\right)\right\}$$

is optimal.

Proof. The agent uses the stated prior density due to Theorem (6) part 2. The early exercise payoff at each time is bounded from above by $H \cdot d - K$. Therefore, early exercise at time t is optimal if

$$\begin{aligned}
&(S_t - K)(1+r)^{T-t} \geq H\cdot d - K\\
\iff &S_t \geq \frac{H\cdot d}{(1+r)^{T-t}} + K\left(1 - \frac{1}{(1+r)^{T-t}}\right).
\end{aligned}$$

See also Reimer and Sandmann (1995). □

Note that the early exercise condition is always satisfied if $S_t = H \cdot d$. Hence, the decision maker always exercises the option when there is knock-out danger meaning that the option's underlying might knock-out in the next period. As a consequence, the decision maker does not directly experience changes of the conditional one-step-ahead probabilities after the exercise. The measure thus changes only once at the exercise. As in previous cases the worst-case measures can be chosen arbitrarily after the exercise. However, given the option is not exercised yet, the worst-case measure switches on the events $\{S_t = H \cdot d\}$. Thus, the worst-case measure and its uniqueness not only depends on the realization of the stock price process but also on the strategy used by investor.

Remark 4. *Assuming additionally in Corollary 3 that the inequality* $\underline{p}u + (1-\underline{p})d > 1+r$ *is satisfied, the American up-and-out call is exactly exercised the first time when* $S_t = H\cdot d$. *This can be derived by the following reasoning: The value of the American up-and-out call under ambiguity being still alive at a fixed time t with* $S_t \leq H\cdot d^2$ *is larger or equal to*

$$\begin{aligned}
U_t \geq & \frac{1}{1+r} E^{\hat{P}}\left((S_{t+1}-K)^+ \mid S_t \leq H\cdot d^2\right) = \frac{1}{1+r} E^{\underline{P}}\left((S_{t+1}-K)^+ \mid S_t \leq H\cdot d^2\right)\\
&= \frac{1}{1+r}\left((S_t\cdot u-K)^+\underline{p} + (S_t\cdot d-K)^+(1-\underline{p})\right)\\
&\geq \max\left\{\frac{1}{1+r}\left((S_t\cdot u-K)\underline{p} + (S_t\cdot d-K)(1-\underline{p})\right), 0\right\}\\
&\geq \max\left\{\left(S_t - \frac{K}{1+r}\right), 0\right\} \geq (S_t-K)^+ \quad \textit{for all } S_t \leq H\cdot d^2.
\end{aligned}$$

The first inequality follows by assumption $\underline{p}u + (1-\underline{p})d > 1+r$. *This shows that the sufficient condition for early exercise is not satisfied for all* $S_t \leq H\cdot d^2$. *Thus, in this case early exercise is only optimal the first time when the price equals* $H\cdot d$.

The analysis of down-and-out put is similar to the exercises we performed above and will therefore be omitted here. There again, the worst-case measure is a pasting of $\overline{P}$ and $\underline{P}$ at the barrier. The remaining barrier options not discussed here (there are four left) are covered by the above theorems. However, the discussion of the worst-case measure for them is even more simple and can be reduced to the monotone case. There, the knock-in/knock-out uncertainty and the stock price uncertainty work in the same direction for those options making them monotone. They can be analyzed in the monotone setting introduced by Riedel (2009).

3.3.2 Multiple Barrier Options

The above reasoning can also be applied to options endowed with more than one barrier. As mentioned in Remark 3 one can use the theorems to obtain the worst-case measure for options with both a knock-in and a knock-out barrier level, or for out-options having an additional barrier level which replaces the original one after some time. A feature similar to this is typical for the express certificate structure that is very common in the German structured derivatives market.

In the following we analyze ladder options and focus on the special case of an up-and-out ladder call option expiring at time T with two barrier levels H_1 and H_2. We assume $S_0 < H_1 < H_2$. This claim resembles a single up-and-out barrier call option with the additional feature that after some prespecified date $t_1 \in (0, T)$ the knock-out barrier changes from H_1 to the higher level H_2. Hitting H_1 after t_1 does not lead to a knock-out. The stock price has to reach H_2 to become worthless after t_1. The option thus pays of

$$\begin{aligned} X_t &= \begin{cases} (S_t - K)^+/(1+r)^t, & \text{if } t \leq t_1 \text{ and } t < \tau_{H_1} \\ (S_t - K)^+/(1+r)^t, & \text{if } t > t_1, t < \tau_{H_2} \text{ and } t < \tau_{H_1} \\ 0, & \text{else} \end{cases} \\ &= (S_t - K)^+/(1+r)^t \, \mathbb{1}_{[0, \tau_{H_1} \wedge t_1[} + (S_t - K)^+/(1+r)^t \, \mathbb{1}_{[t_1, \tau_{H_2} \wedge \tau_{H_1}[} \end{aligned}$$

whereas $\tau_{H_1} := \inf\{t \in [0, t_1] \big| S_t = H_1\} \wedge T + 1$ and $\tau_{H_2} := \inf\{t \in]t_1, T] \big| S_t = H_2\} \wedge T + 1$. Here, $[t_1, t_1[$ is defined as the empty set.

As we will see, the additional barrier impacts the form of the worst-case measure by influencing the optimal exercise strategy of the buyer. While the buyer of a single barrier up-and-out call cashes the option as soon as $H \cdot d$ is reached this might be not optimal here since she can get an even higher payoff after t_1. The higher return after the first barrier is dropped increases the value option making waiting more attractive.

In order to represent the density of the worst-case measure we need the following stopping times that indicate nodes at which a knock-out is possible in the next period:

$$\sigma_i := \inf\{t \in [\sigma_{i-1} + 1, \tau_{H_1} \wedge t_1 - 1[\; \big| \; S_t = H_1 \cdot d\} \wedge T + 1$$

for all $1 \leq i < t_1$ with the notation $\sigma_0 := -1$.

Corollary 4 (Ladder call option)**.** *Let all data be given as above, in particular, let us suppose the strict inequality of Corollary 3. Additionally, assume that for all* $t \leq t_1$

$$H_1 \cdot d/(1+r)^t < \mathbb{E}^{\overline{P}}(U_{t+1}|S_t = H_1 \cdot d) \tag{3.9}$$

Then the ambiguity averse buyer of this ladder option uses the measure $\hat{P}$ *specified by the density*

$$\hat{D}_t := 2^t \prod_{u \leq \tau_{H_2} \wedge t:\ u \neq \sigma_i + 1} \big(\varepsilon_u \underline{p} + (1-\varepsilon_u)(1-\underline{p})\big) \prod_{u \leq T:\ u = \sigma_i + 1 \text{ or } u \in]\tau_{H_2}, t]} \big(\varepsilon_u \overline{p} + (1-\varepsilon_u)(1-\overline{p})\big)$$

for all occurring $1 \leq i < t_1$.

Proof. We can also apply the second part of the theorem to this special situation since the time interval $[0, T]$ is divided into two disjoint intervals and $A(t, S_t) := (S_t - K)^+/(1+r)^t$, increases in S_t for all $t \leq T$ on both intervals. Thus, applying the theorem on both subintervals yields the density for the ambiguity averse agent. From t_1 on same arguments as in the case of an usual up-and-out call option (see Corollary 3) lead to the worst-case measure along with an optimal stopping time. □

In the case of the up-and-in call the measure changes only once if the buyer exercises the option optimally. The situation differs here because it can be optimal to hold on to the option even if it can knock-out in the next period. The worst-case measure here is thus the one with the lowest mean returns at nodes from which the option can be knocked out, i.e. nodes with $S_t = H_1 \cdot d$ for $t < t_1$ and $S_t = H_2 \cdot d$ afterwards. Once the stock price is below this threshold the conditional probability switches back to the marginal probability with the lowest returns. Of course, the so defined measure is also a worst-case measure for the case of the up-and-in call.

3.4 Multiple Expiry Options

While in the previous section uncertainty came from stochastic events that couldn't be controlled by the investor, here we consider claims where the holder can partially control and switch off one source of uncertainty. We analyze claims where the owner of the option has the right to modify the contract conditions resetting the strike or the maturity in a predefined way. New conditions of the contract depend on the underlying's value at switching dates and are not known to the buyer at time zero. Therefore, additional to the uncertainty about future underlying's value the decision maker faces uncertainty about future contract conditions while evaluating the option a priori. The expiry dates can be predefined points in time (forward start options) or random dates chosen by the buyer or seller of the contract (shout options).

Such options can be seen as a sequence of claims where every claim expires at a predefined date and pays off a new born claim expiring at the next expiry date. In the case of European claims the expiry dates are deterministic corresponding to forward start options. In the case of shout options we face American claims leading to stochastic expiry dates. In general, multiple expiry options can be entitled with any number of expiry dates, here, we consider dual expiry options where contract conditions change exactly once [14].

3.4.1 Shout Options

Shout options are contracts that give the buyer the right to reset the strike at a date chosen by her. The event of resetting the contract features is called shouting and gives the structure its name. The reset right allows the

[14] Kwok and Wu (2004) analyze shout options with infinite number of shout possibilities and establishes a relation to lookback options.

investor to benefit from market movements by choosing a favorable strike. At the same time she can lock in already realized profits ensuring against an unfavorable stock movement.
Shout options are often used by professional investors as a cheaper alternative to lookback options. Whereas the buyer of the lookback option has the right to sell the stock at the maximal price the owner of the shout option has to call her bank and to freeze the price at which she can sell/buy at any time $\sigma \leq T$ prior to maturity. Mathematically, the buyer faces an optimal stopping problem, aiming to set the strike optimally.

In the following we analyze shout floors. The same analysis can be performed for call options.

At time zero the buyer receives a put option with an unspecified strike. This strike is set to be the stock price at a date chosen by the buyer. Thus, the buyer has to call his bank and to freeze the strike when she thinks that the strike is favorable. The buyer shouts once at $\sigma \leq T$ fixing the strike at S_σ. At the expiry date she receives a payoff that corresponds to the payoff profile of an European put i.e. $(S_\sigma - S_T)^+$. Thus, the buyer of this shout option has to solve the following problem

$$\text{Maximize } \min_{P \in \mathcal{Q}} \mathbb{E}^P((S_\sigma - S_T)^+/(1+r)^T) \text{ over all stopping times } \sigma \leq T \tag{3.10}$$

Note, that unlike the American put, the exercise date is fixed but the birth date has to be determined optimally by the buyer. Determining the optimal starting time/shouting time constitutes the optimal stopping problem for the single shout option. The task is to optimally start the payoff process rather then stop it which can be seen as purchasing a new issued European option with a fixed maturity. We will maintain this parallel during our analysis.

However, we cannot apply our standard theory of backward induction to the problem stated in (3.10) because the payoff $(S_\sigma - S_T)^+/(1+r)^T$ obtained from stopping at any stopping time $\sigma \leq T$ depends on the value of the stock at maturity and is for this reason not adapted to the filtration $(\mathcal{F}_t)_{t=1,\ldots,T}$ generated by the path. To overcome this difficulty we condition the payoff on the available information and consider the following payoff process

$$X_t = \min_{P \in \mathcal{Q}} \mathbb{E}^P((S_t - S_T)^+/(1+r)^T | \mathcal{F}_t). \tag{3.11}$$

For every $t \leq T$ we can interpret X_t as discounted multiple prior value of the shout floor if shouted at t. At the same time it corresponds to the value of an at-the-money European put issued at t and maturing at T evaluated under multiple priors.[15]

With the same arguments that we used to prove Lemma (1) we get

Lemma 6. *For all stopping times $\sigma \leq T$ we have*

$$\min_{P \in \mathcal{Q}} \mathbb{E}^P((S_\sigma - S_T)^+/(1+r)^T) = \min_{P \in \mathcal{Q}} \mathbb{E}^P(X_\sigma) \tag{3.12}$$

Therefore, we can reformulate the problem stated in (3.10) equivalently in the following way

$$\text{Maximize } \min_{P \in \mathcal{Q}} \mathbb{E}^P((X_\sigma) \text{ over all stopping times } \sigma < T \tag{3.13}$$

where the payoff process X is defined via (3.11). Thus, the optimal stopping time found for (3.13) is also optimal for the problem (3.10) and the values of the two problems coincide. Again, we can interpret the problem as optimal investment in a put with a fixed investment horizon.

As in Chapter 2 we solve the problem in two steps: first we compute X_t – the explicit value of the shout option freezed at t for all $t \leq T$ and derive the worst-case measure after shouting. In the second step, we identify the worst-case measure before shouting reducing the problem to the single prior case.

[15] Strictly speaking, the value of the European put issued at t and maturing at T differs from the expression 3.11 by a discount term

Lemma 7. *The adapted payoff process corresponding to the optimal shouting problem* (3.13) *satisfies*

$$X_t = x(S_t, t, \overline{P}) \tag{3.14}$$

and is homogeneous of degree 1 in S_t for every $t \leq T$. Moreover, the optimal shouting rule σ^* for the problem is the same as the optimal stopping rule for the Bayesian problem

$$\textit{Maximize } \mathbb{E}^{\underline{P}}(X_\sigma)$$

under the measure $\underline{P}$ over all shouting times $\sigma < T$.

Proof. To compute X_t for a fixed $t \leq T$ we note that the uncertainty about the strike is resolved at the time of shouting. The strike becomes a constant and as a consequence the claim becomes a plain vanilla European put. As the payoff of the put is decreasing in S_t for all $t \leq T$ by Theorem 5 we conclude that the worst-case measure is given by $\overline{P}$ and we have

$$\begin{aligned} X_t =& \min_{P \in \mathcal{Q}} \mathbb{E}^P \left((S_t - S_T)^+ / (1+r)^T | \mathcal{F}_t \right) \\ =& \mathbb{E}^{\overline{P}}((S_t - S_T))^+ / (1+r)^T | \mathcal{F}_t). \end{aligned}$$

Additionally, under $\overline{P}$ the increments of the underlying between t and $T - \Delta(S_t, S_T)$ are independent for all $t \leq T$ which leads to

$$\begin{aligned} X_t =& S_t \cdot \mathbb{E}^{\overline{P}}((1 - \Delta(S_t, S_T))^+ / (1+r)^T | \mathcal{F}_t) \\ =:& S_t \cdot g(\tau) \end{aligned} \tag{3.15}$$

where $\tau = T - t$ and

$$g(\tau) = (1+r)^{-T} \cdot (1 - \overline{p})^\tau \sum_{k=0}^{k^*(\tau)} \binom{\tau}{k} \left(\frac{\overline{p}}{1 - \overline{p}} \right)^k \left(1 - d^{\tau - 2k} \right)$$

with $k^*(t) =: \max \left\{ k : k < \frac{\tau}{2} \right\}$.

The above equation provides the value of the embedded option contained in the shout contract maturing at T at the time of shouting. At the same time it corresponds to the value of the at-the-money European put issued at $t \leq T$ and maturing at T.

The buyer of a shout option uses $\overline{P}$ to evaluate the option after shouting. Moreover, the value of a freezed shout floor is homogeneous of degree one in the current stock price S_t.

As $g(\tau) > 0$ for all $\tau > 0$ for every $t \leq T$ $x(\cdot, t)$ is increasing in S_t. Again using Theorem (5) we conclude that the worst-case measure of problem (3.13) is given by $\underline{P}$. □

Remark 5. *It might be surprising at the first sight that the value of the option is monotone increasing in the strike since the opposite is true for put options. The reason for this is the fact that the strike is not a constant at the moment of issuance of the option. The value of the claim at the time of shouting is increasing with respect to the difference between strike and the current stock price. Economically, a higher S_t at the time of shouting increases the strike of the new born option enlarging the in-the-money region of the option.*

From the discussion above we can already derive the worst-case measure for shout claims:

Corollary 5 (Shout put)**.** *A risk-neutral buyer of an single shout floor option uses the optimal stopping rule for the prior $\hat{P}$ given by the density*

$$\hat{D}_t = 2^t \prod_{v=1}^{\sigma^* \wedge t} \left(\underline{p} \cdot \epsilon_v + (1 - \underline{p}) \cdot (1 - \epsilon_v)\right) \prod_{v=\sigma^*+1}^{t} \left(\overline{p} \cdot \epsilon_v + (1 - \overline{p}) \cdot (1 - \epsilon_v)\right)$$

where σ^* denotes the stopping time solving (3.13)*.*

Summing up, we conclude that the value of the shout floor is given by

$$U_t^Q = \begin{cases} \mathbb{E}^{\underline{P}} \left(\mathbb{E}^{\overline{P}} \left((S_\sigma^* - S_T)^+ / (1+r)^T | \mathcal{F}_\sigma^* \right) | \mathcal{F}_t \right), & \text{if } t < \sigma \\ \mathbb{E}^{\overline{P}} \left((S_\sigma^* - S_T)^+ / (1+r)^T | \mathcal{F}_t \right) & \text{else} \end{cases}$$

The decision maker changes her beliefs about mean returns at the first expiry date. Before shouting and freezing the strike she presumes low returns of the stock that keeps the in-the-money region of the option small and decreases the value of the embedded put; after shouting she receives a put option and therefore changes her belief — being pessimistic, she now presumes that the risky asset will have high returns. This change of beliefs causes the difference in the values of the classical result and the multiple prior result.

To complete the analysis it remains to solve the optimal stopping problem for X under the worst-case measure. The classical solution for the continuous time setting was provided by Kwok and Wu (2004). To our knowledge binomial tree analysis has not been conducted for shout options yet.

Lemma 8. *Denote by* $\underline{\mu}$ *the mean return under* $\underline{P}$ *i.e.* $\underline{\mu} = \underline{p} \cdot u + (1 - \underline{p}) \cdot d$ *and by* x^* *the maximum of the function* $g(\tau) \cdot \mu^{T-\tau}$ *where* $g(\tau)$ *is defined as above. Then an optimal stopping time is given by*

$$\sigma^* = \inf\{t \geq 0 | g(\tau) \cdot \underline{\mu}^{T-\tau} = x^*\}$$

If the maximum x^* *is unique, then all* $\sigma^* \leq t \leq T - 1$ *are optimal.*

Proof. To prove the lemma we use the generalized parking technique introduced by Lerche, Keener, and Woodroofe (1994). For all $t \leq T$, $\tau = T - t$ we have

$$\mathbb{E}^{\underline{P}}(S_t \cdot g(\tau)) = \mathbb{E}^{\underline{P}}\left(\frac{S_t}{\underline{\mu}^t} \cdot g(\tau)\underline{\mu}^t\right) \leq \mathbb{E}^{\underline{P}}\left(\frac{S_t}{\underline{\mu}^t} \cdot x^*\right)$$

and equality holds for the maximizer t^*. Now since $\frac{S_t}{\underline{\mu}^t}$ is a $\underline{P}$-martingale we get for all stopping times $\sigma < T$

$$\mathbb{E}^{\underline{P}}\left(\frac{S_\sigma}{\underline{\mu}^\sigma} \cdot x^*\right) = S_0 \cdot x^*$$

and therefore

$$V_0 = \sup_{\sigma < T} \mathbb{E}^{\underline{P}}\left(S_\sigma \cdot g(\tau)\right) = \mathbb{E}^{\underline{P}}(S_t^* \cdot g(T - t^*))$$

where t^* satisfies $g(T - t^*) \cdot \underline{\mu}^{t^*} = x^*$. □

The optimal stopping rule is deterministic and does not depend on the level of the stock price S. This follows from the homogeneity of the payoff in S. However, the time of stopping depends highly on the model parameters $u, \overline{p}, \underline{p}$. We suspect that function $g(\tau) \cdot \mu^{T-\tau}$ is quasiconcave and thus have a unique maximum but we are not able to prove it. However, can state a sufficient condition for immediate stopping.

Corollary 6. *In the above situation we have $\sigma^* = 0$ if $1 - \overline{p} \geq \underline{\mu}$.*

While in the classical CRR market the stopping time depends only on the one step mean return, in the multiple prior model the relation of $\underline{p}$ and $\overline{p}$ plays a crucial role.

3.5 Quasi-convex Payoffs

In the last section we consider options whose payoffs consist of two monotone parts. Typical examples are options having U–shaped payoff including straddles, strangles or short option strategies. Investors buying such options are speculating on change in the underlying's value without specifying the direction of the change. Depending on the actual price of the underlying falling or rising stock increases the profit of the investor. To illustrate this idea consider a straddle: by exercising the straddle at $S_t > K$ the buyer gets a payment of $(S_t - K)$ which corresponds to a call. Otherwise, if $S - t < K$ she gets the payoff of a put – $(K - S_t)$. The first part of the payoff is often called the call leg, the second the put leg. As the monotonicity on both legs of the payoff is different, the worst-case measure is also different depending on the actual stock price at the moment of the valuation.

Remark 6. *Mathematically, payoffs described above correspond to quasi-convex/quasi-concave payoff functions. Note, that we still deal with functions discretely defined and have to be careful when using the term quasi-convex. Strictly speaking, the notion we use corresponds to discrete convexity studied* *intensively in the context of indivisible goods (see for example Murota (1998) for a general introduction). In one dimensional setting discrete convexity reduces to the following: A set $E \subset \mathbb{N}$ is convex if all points in E are contained in the convex hull of E. The definition of quasi-convex is then straight forward.*

We show that the Snell envelope $U_t^{\mathcal{Q}}$ at time $t \leq T$ is a quasi-convex function in S_t if the claim is Markovian.

Lemma 9. *If the discounted payoff function $A(t, S_t)$ is quasi-convex in its second variable for every $t \leq T$, then the Snell envelope $U_t^{\mathcal{Q}}$ is given by a quasi-convex function $v(t, x)$, i.e. given $S_t = x_t$*

$$U_t^{\mathcal{Q}} = v(t, x_t) = \sup_{\tau \geq t} \min_{P \in \mathcal{Q}} \mathbb{E}^P(A(\tau, S_\tau)|S_t = x_t)$$

Proof. We have to show that for every $t \leq T$ the value function $v(t, \cdot)$ depends only on the value of the stock at time t and that the quasi-convexity of the payoff function carries over to the value function. We do it via backward induction.

Before applying backward induction we note that in our case a function $g : E \to \mathbb{N}$ is quasi-convex if and only if there exists a $\hat{x} \in E$ such that $g(x) \geq g(\hat{x})$ holds for all $x \in E$.

Now to the backward induction. For $t = T$ we clearly have for all possible values of $S_T = x_T$

$$U_T^{\mathcal{Q}} = A(T, x_T)$$

where $A(T, \cdot)$ is a quasi-convex function.

For $t+1 < T$ we assume that for any value of $S_{t+1} = x_{t+1} \in E_{t+1}$ the value function $v(t + 1, \cdot)$ is quasi-convex function depending only on the current value of the stock. Because of quasi-convexity there exists a unique minimum m_{t+1} and a unique

$$\hat{x}_{t+1} = \inf\{x_{t+1} \in E_{t+1}|v(t + 1, x_{t+1}) = m_{t+1}\}.$$

The function $v(t + 1, \cdot)$ is decreasing on the set $\{x_{t+1} \leq \hat{x}_{t+1}\}$ and increasing on the set $\{x_{t+1} \geq \hat{x}_{t+1}\}$.

In $t < T$ we then have for any value $S_t = x_t$

$$\begin{aligned} U_t^{\mathcal{Q}} &= \max\{A(t, S_t), \min_{P \in \mathcal{Q}} \mathbb{E}^P \left(U_{t+1}^{\mathcal{Q}} | \mathcal{F}_t\right)\} \\ &= \max\{A(t, x_t), \min_{P \in \mathcal{Q}} \mathbb{E}^P \left(U_{t+1}^{\mathcal{Q}} | S_t = x_t\right)\} \\ &= \max\left\{A(t, x_t), \hat{p}_{t+1} v(t+1, x_t \cdot u) + (1 - \hat{p}_{t+1}) v(t+1, x_t \cdot d)\right\} \\ &= v(t, x_t) \end{aligned} \tag{3.16}$$

where $\hat{p}_{t+1} \in [\underline{p}, \overline{p}]$ is the marginal of the worst-case measure $\hat{P}$ at time t. Since $v(t+1, \cdot)$ is independent of the realized past by assumption, the minimizer $\hat{p}_{t+1}$ depends only on the value of x_t. This proves that the value function at time t $v(t, \cdot)$ depends only on current value of the underlying.

To prove quasi-convexity we analyze the structure of the continuation value in equation (3.16)

$$u(t, x_t) := \hat{p}_{t+1} v(t+1, x_t \cdot u) + (1 - \hat{p}_{t+1}) v(t+1, x_t \cdot d)$$

for different values of $S_t = x_t$.

On the set

$$E_t^d = \{x_t \in E_t | x_t \leq \hat{x}_{t+1} \cdot d\} \tag{3.17}$$

we have

$$x_t \cdot d < x_t \cdot u < \hat{x}_{t+1}$$

and therefore using the induction hypothesis we can conclude that the function $u(t+1, \cdot)$ is decreasing as a convex combination of two increasing functions. Similarly, for all

$$E_t^i = \{x_t \in E_t | x_t \geq \hat{x}_{t+1} \cdot u\} \tag{3.18}$$

we have $\hat{x}_{t+1} < x_t \cdot d < x_t \cdot u$ and the function increases on the above set with the same argument.

Because of the binomial tree structure of the state space and the fact that

$$E_{t+1} = \{E_t \cdot u^k | k \in \{-1; 1\}\}$$

equations (3.17) and (3.18) partition the set of possible values of S_t and E_t can be written as

$$E_t = \{x_t \in E_t | x_t \leq \hat{x}_{t+1} \cdot d\} \cup \{x_t \in E_t | x_t \geq \hat{x}_{t+1} \cdot u\}$$

Because of monotonicity of $u(t, \cdot)$ on E_t^d and E_t^i the minimum of $u(t, \cdot)$ is unique. This shows that the function $u(t, \cdot)$ is quasi-convex.

To complete the proof we recall that $A(t, x_t)$ is quasi-convex by assumption. Thus, the function defined by equation (3.16) is a quasi-convex function as maximum of two quasi-convex functions. Clearly, the value function at time t depends only on the current stock price and given $S_t = x_t$ we can write $U^{\mathcal{Q}}$ as a function $v(t, x_t)$. □

The quasi-convexity of the value function implies that for every $t \leq T$ we can separate the space E_t on which the value of the claim is monotone allowing to determine the worst-case measure. The decomposition point is the minimizer of the value function $\hat{x}_t$ which is constructed in the proof of Lemma 9.

Having analyzed the shape of the value function we now can compute the worst-case measure with the following argument. If asset prices are low, the value function is decreasing. Therefore, with the same argument as for simple American options, one can show that $\overline{P}$ is the worst-case measure here. In the other region on the contrary, $\underline{P}$ is the worst-case measure. At a predefined level $\hat{x}_t$ the investor changes his beliefs and so the mean return on stock under the measure. We then have the following

Lemma 10 (Straddle)**.** *The buyer of a straddle uses the optimal stopping rule for the measure $\hat{P}$ with density*

$$\hat{D}_t = 2^t \prod_{v \leq t, S_v \in E_v^i} \big(\underline{p} \cdot \epsilon_v + (1 - \underline{p}) \cdot (1 - \epsilon_v)\big) \prod_{v \leq t, S_v \in E_v^d} \big(\overline{p} \cdot \epsilon_v + (1 - \overline{p}) \cdot (1 - \epsilon_v)\big).$$

Proof. We consider the value function on the continuation region where for a given $S_t = x_t$ we have $U_t^{\mathcal{Q}} = v(t, x_t)$

$$v(t, x_t) = \min_{p_{t+1} \in [\underline{p}, \overline{p}]} (p_{t+1} v(t + 1, x_t \cdot u) + (1 - p_{t+1}) v(t + 1, S_t \cdot d)$$

As $v(t,\cdot)$ is decreasing on E_t^d, the worst-case measure on this set is given by $\overline{P}$. With the same argument the worst-case measure $\hat{P}$ is $\underline{P}$ on E_t^i, i.e.

$$\hat{P}[\epsilon_{t+1} = 1|\mathcal{F}_t] = \begin{cases} \underline{p} & \text{on } \{x_t \geq \hat{x}_{t+1} \cdot u\} \\ \overline{p} & \text{on } \{x_t \leq \hat{x}_{t+1} \cdot d\} \end{cases} . \tag{3.19}$$

where $\hat{x}_{t+1}$ is the minimizer of $v(t+1,\cdot)$. Using the definition of $\underline{p}$ and $\overline{p}$ and pasting the densities together one obtains the result. □

Under $\hat{P}$ the process (S_t) becomes mean-reverting in an appropriate sense pushing S_t down if it is high and up if it is low. This corresponds to the intuition: the ambiguity averse decision maker anticipates low mean returns in bull market phases and high mean returns when the stock value is low. Unlike previous cases the uncertainty about the payoff function here cannot be resolved before T in general. The change of the measure occurs every time the stock price crosses the critical value $\hat{x}_t$ forcing the decision maker to change her beliefs about mean returns.

3.6 Conclusion

This paper studies the worst-case measures that arise if one considers various American options in a framework that allows for model uncertainty in discrete time. The imprecise information about the correct probability measure driving the stock price process in the market generates different models with varying conditional one-step-ahead probabilities used by the buyer. The buyer is then allowed to change the measure, and so the model she uses and to assign the value to the claim according to the worst possible model. While the solution for plain vanilla options is straightforward in the model the situation differs if the payoff of the option becomes more sophisticated. The effect of uncertainty differs over time leading to a dynamical structure of the worst-case measure. This paper analyzes different effects of uncertainty highlighting the structural difference between the standard models used in

Finance and the multiple prior models: the buyer of the option adapts her beliefs to the state of the world and the overall effect of model uncertainty. A natural next step is to extend the theory to continuous market models and to analyze exotic options in that framework.

Chapter 4

Ambiguity Aversion and Overpricing

4.1 Introduction

The main goal of the Chapter is to highlight the impact of ambiguity on market prices in a static market model with heterogeneous agents. We show that the presence of ambiguity averse agents can lead to higher prices of the asset.

Classical asset pricing literature relies heavily on the assumption that agents have homogeneous correct expectations about future returns. Lintner (1969) first analyzes a partial equilibrium model where agents have heterogeneous beliefs about the profitability of the asset. It turns out that the equilibrium price corresponds to the weighted average of opinions of market participants. While bullish (optimistic) investors demand security, bearish (pessimistic) investors supply it by shortselling. In equilibrium the price reflects the average opinion on the market. In this setting, Miller (1977) and Jarrow (1980) analyze the effect of short selling constraints on the equilibrium price and show that short selling constraint together with heterogeneous expectations may lead to overpricing. Since pessimistic agents cannot express their beliefs by selling the asset short, the price is biased upward and reflects the opinion of more optimistic agents. If this trend is not corrected over several periods a bubble can arise.

This idea gave rise to a series of papers analyzing bubbles and speculative overpricing caused by heterogeneous beliefs. Harrison and Kreps (1978) constructed a speculative bubble model in discrete time, Scheinkman and Xiong (2003) modeled overpricing as a consequence of overconfidence in continuous time. However, all papers explaining bubbles by heterogeneity of agents rely on the impossibility of short selling. This restriction is set exogenously and is justified by the market structure, high costs of short selling or regulation. The assumption which seems reasonable in some situations is hard to support in general. Most of developed financial markets explicitly allow short selling, a vast majority of stocks traded on exchanges is shortable at low cost[1]. At the same time short supply, i.e. the supply generated through short selling constitutes only a small fraction of the market. Based on an empirical analysis Lamont and Stein (2004) come to the conclusion that "..the problem is not too much short selling in falling markets ... but rather, too little in rising markets". To explain this phenomena authors refer to internal restrictions set by companies' chartas or reluctance to sell short. However, there is no model rationalizing this reluctance to sell short.

In this Chapter we relax the assumption of impossibility of short selling and establish the reluctance to sell the asset short through preferences. The reluctance to short sell is not an exogenously given attitude as before but comes as optimal behavior of agents that have certain preferences. To model this we use ambiguity averse preference in the sense of Knight, axiomatized by Gilboa and Schmeidler (1989).

Already Miller (1977) in his paper referred to Knightian uncertainty and pointed out that "[i]n practice, uncertainty, divergence of opinion about a

[1] A review of short selling constraints over the world can be found in Saffi and Sigurdsson (2008).

security's return, and risk go together". Thus, in markets where agents have different views on the asset at least some market participants are likely to experience uncertainty and behave like ambiguity averse agents. We formalize this idea and analyze a market where heterogeneous agents face risk and uncertainty.

We assume that some market participants are subjective utility maximizer (SEU) that differ in their expectation on the security's return while others have minimax preferences axiomatized by Gilboa and Schmeidler (1989). An ambiguity averse decision maker (AA) uses a class of models instead of one to assess utilities to future payoff streams and commits to a position only if the expected utility of this position is positive for all models she considers. In this setting both types of agents determine their demand for the risky/uncertain security by solving their utility maximization problem. Subjective utility maximizers then demand the security if the price is below their subjective expected return and supply it otherwise. In any case they hold a position of the security except for the knife edge case. The picture is different for ambiguity averse investors: there exists an interval of prices within which it is optimal for them to hold zero position in the security. This so called no trade interval first studied by Dow and Werlang (1992) arises because the expected utility of a short and long position is positive for some but not for all models the investors takes into account. As a result the ambiguity averse investor refuses to participate in the market at all. This has two implications: First, the agents do not demand the asset and the risk has to be taken by fewer investors. This leads to higher risk premia required by investors to hold the asset. This effect has been studied extensively in the non-participation literature. On the other hand, the ambiguity averse agents also refuse to short the asset and fail to generate short supply. Thus the supply is lower compared to a market with SEU agents only. This can lead to higher equilibrium prices. Even though short selling is not forbidden, an upward biased price (compared to the average risk adjusted valuation) can arise if subjective utility maximizers are overly optimistic and bid the price up. As a result an increase in ambiguity may force ambiguity averse decision maker to stop short selling and thus increase the equilibrium price. This effect does not arise in the previous equilibrium models with ambiguity where an increase in ambiguity decreases demand and leads to lower equilibrium prices.

The result is an extension of the series of papers following the ideas of Miller (1977). Unlike previous papers on overpricing we do not impose the short selling assumption exogenously. Here, it is a result of a rational utility maximization of agents having minimax preferences.

Several papers investigate the impact of ambiguity on the portfolio and investment choice and its consequences for markets. Epstein and Wang (1994) use ambiguity aversion and no trade interval to explain non participation in the markets and portfolio inertia. It is showed to be a reason for underdiversification in Uppal and Wang (2003) and market incompleteness in Mukerji and Tallon (2001). Caballero and Krishnamurthy (2007) highlights the role of ambiguity in flights to quality. In the latter model an increase in ambiguity that is unrelated to fundamental value causes a sell out in the security pressing the price down and resulting in a flight to quality. Ui (2009) studies a model of non-participation in a financial market with finitely many different potential investors who exhibit heterogeneous levels of ambiguity and obtain private signals. As Cao, Wang, and Zhang (2005) and Easley and O'Hara (2009) they also note that equity premium can decrease if non-participation arises. This happens because investors that exhibit higher levels of ambiguity aversion and therefore demand a higher premium for holding the asset leave the market. This decreases the average premium required to hold the security. The heterogeneity in this models refers to different degrees of uncertainty aversion of the investors and not to their point estimate of the returns of the asset.

Our approach differs from this literature in two aspects. First, non-participation literature concentrates on the interaction of ambiguity averse and perfectly rational agents having correct beliefs for the asset return. In contrast, we study the interaction of ambiguity averse decision makers with heterogeneous risk averse agents. As a result we do not assume anyone being rational and having correct beliefs. The prices on the market reflect solely average beliefs and not necessarily fundamental values. All our predictions about the changes of the equilibrium price are made with respect to the average opinion and not with respect to the fundamental value.

Second, previous models on ambiguity aversion in equilibrium models suggest that the presence of ambiguity averse decision makers results in lower prices since ambiguity aversion increases the premium required by the investor to hold the asset. In our model ambiguity can cause a price increase. As more agents become ambiguity averse or ambiguity about the return of investment increases, more investors become reluctant to short the asset, thus lowering the supply of the asset. If SEU agents are optimistic enough the effect of lower supply is stronger than the effect of lower demand. In equilibrium we get higher prices.

Our approach can also be seen in the spirit of limits of arbitrage studied by Shleifer and Vishny (1997). There, rational arbitrageurs refuse to correct a bubble and to take advantage of an arbitrage because this arbitrage is risky. They only step in if the mispricing is high enough to be rewarded for the risk they take. Our story is similar: the arbitrage here is not only risky but also ambiguous. Moreover, given the price is in the no trade interval, a short position in the security is arbitrage for some models an ambiguity averse decision maker takes into account but not in all. Similar to Shleifer and Vishny (1997) overpricing can persist due to its ambiguous nature.

4.2 The Model

4.2.1 Setup

We consider a two period exchange economy with two assets: one ambiguous and one riskfree. The risky asset is traded at $t = 0$ and pays a liquidating[2] dividend x at time $t = 1$. The supply of the asset is fixed at Q while demand is determined by the maximization problem of investors. The riskfree asset is traded in infinite supply at a given riskfree rate [3].

There is a continuum of risk averse agents in the economy that share the same v NM index defining the CARA utility with risk aversion coefficient γ:

$$u(x) = -e^{-\gamma x}. \tag{4.1}$$

Investors that take prices as given differ in their beliefs about returns on stock and their attitude towards uncertainty. There are two types of investors: subjective utility maximizer (SEU agents) that maximize their expected utility under their subjective belief and ambiguity averse decision maker (AA agents) that take a class of models into account since they do not trust the validity of one particular model.

While all market participants agree that the dividend in the next period is normally distributed with volatility σ they disagree about the expected return μ of the asset. This disagreement might be a result of overconfidence in own ability to evaluate signals as the model of Scheinkman and Xiong (2003) suggests or be the result of the use of different models.

The range of possible drifts is given by $[\underline{\mu}, \overline{\mu}]$, investors are distributed across this interval according to a distribution M having a density m.

[2]To set up a meaningful static model we assume that the risky asset is withdrawn from the market and has zero value after paying the dividend at time $t = 1$.

[3]We consider a partial equilibrium model and set the riskfree rate as exogenous.

4.2.2 Maximization Problem

Subjective expected utility investors

Heterogeneous risk averse agents agree to disagree about the asset's return. Each risk averse investor has a point estimate μ_i for the expected return which he uses for his evaluations. For expositional simplicity we assume that risk averse investors are optimists, having a belief above some threshold $\hat{\mu} \in [\underline{\mu}, \overline{\mu}]$. From the modeling point of view this assumption seems reasonable since agents that are optimistic about the returns on a new security are also likely to be confident about the choice of their model. An example for this type of behavior was the overoptimism combined with a high level of confidence during the Internet bubble. The assumption made here is not crucial for the validity of the result but simplifies greatly the analysis. We will discuss how to relax it later on.

Given the belief μ_i the agent maximizes her expected utility. Due to the shape of the utility function endowments of agents do not affect their demand for the risky asset. The problem of the SEU investor with belief μ_i then reads

$$\text{Maximize } \mathbb{E}^i(-\exp(-\gamma d_i^s(x-p))) \text{ over } d_i \in \mathbb{R} \tag{4.2}$$

where p denotes the equilibrium price of the asset and d_i^s the number of risky asset in the portfolio. The expectation is taken with respect to her personal belief μ_i. Note that unlike the seminal paper of Miller (1977) we do allow for short selling, i.e. negative values of d_i. Thus, depending on the equilibrium price, the agent can be either supplier or demander of the risky asset.

Standard techniques show that the demand of the investor i is given by

$$d_i^s = \frac{\mu_i - p}{\gamma\sigma^2} \tag{4.3}$$

An SEU agents is a net demander of the asset if the price is below the mean return and a net supplier if the price is above. In any case the optimal position in the asset is nonzero except for the knife edge case $\mu_i = p$. Thus, SEU agents are always active on the market trading in one or other direction and generating both demand and short sell supply depending on the individual belief μ_i.

Using the mean value theorem the aggregate demand of SEU investors can be calculated as

$$D^s = \int_{\hat{\mu}}^{\overline{\mu}} d_i m(d\mu_i) \tag{4.4}$$

$$= \int_{\hat{\mu}}^{\overline{\mu}} \frac{\mu_i - p}{\gamma\sigma^2} m(d\mu_i) \tag{4.5}$$

$$= \frac{\underline{\mu} + k^* - p}{\gamma\sigma^2} (1 - M(\hat{\mu})) \tag{4.6}$$

where $k^* \in [\hat{\mu} - \underline{\mu}, \overline{\mu} - \hat{\mu}]$. The demand of SEU agents is determined by the weighted average of opinions of SEU agents $\underline{\mu} + k^*$ and the mass of the SEU agents in the economy $(1 - M(\hat{\mu}))$.

Ambiguity averse investors

An ambiguity averse decision maker is uncertain about the right model and takes a set of models into account. Instead of using their own model ambiguity averse decision maker use all models that they see on the market. More precisely, they build a belief about the return and if the belief is below the ambiguity threshold $\hat{\mu}$ they use all estimates they see on the market to assess the profitability. Thus, the set of priors used by ambiguity averse agents is given by

$$\mathcal{Q} = \{P : x_P \sim \mathcal{N}(\mu, \sigma^2), \mu \in [\underline{\mu}, \overline{\mu}]\} \tag{4.7}$$

Being ambiguity averse she maximizes her minimal expected payoff, i.e.

$$\text{Maximize } \inf_{P \in \mathcal{Q}} \mathbb{E}^P(-\exp(-\gamma d_i(x - p))) \text{ over } d_i \in \mathbb{R} \tag{4.8}$$

where $\mathcal{Q}$ is defined by (4.7). It is known from Dow and Werlang (1992) that the demand function of the ambiguity averse investor is continuous and has kinks at $p = \underline{\mu}$ and $p = \overline{\mu}$. The ambiguity generates a so called no trade interval, in which the agents refuse to trade the risky asset. The exact expression for the demand function of an ambiguity averse agent in our setting is given by

$$d^a = \begin{cases} \frac{\underline{\mu}-p}{\gamma\sigma^2} & \text{if } p < \underline{\mu} \\ 0 & \text{if } \underline{\mu} < p < \overline{\mu} \\ \frac{\overline{\mu}-p}{\gamma\sigma^2} & \text{if } p > \overline{\mu} \end{cases} . \tag{4.9}$$

The ambiguity averse investor demands the asset if the equilibrium price is low enough. However, if the price is sufficiently high, i.e. $p > \underline{\mu}$ the ambiguity averse agents refuses to invest in risky asset. Unlike the SEU agent who starts short selling as soon as she stops buying, ambiguity averse agent is also reluctant to short sell the asset at a price $p \in [\underline{\mu}, \overline{\mu}]$. Thus, ambiguity has two effects: first, if the price is above $\underline{\mu}$ ambiguity averse agents stop investing in the asset, decreasing aggregate demand, on the other hand, they do not start short selling thus decreasing supply of the asset.

Clearly, the aggregate demand of ambiguity averse agents is then given by

$$D^a = \int_{\underline{\mu}}^{\hat{\mu}} d^a m(d\mu_i) \tag{4.10}$$

4.2.3 Equilibrium Analysis

Before we perform equilibrium analysis we note that the demand is only positive if $p < \overline{\mu}$ since all agent aim to sell the asset if $p > \overline{\mu}$. Thus, in equilibrium ambiguity averse agents will never short the asset since the price for the risky asset will never exceed the most optimistic valuation $\overline{\mu}$ in equilibrium. Therefore, the demand of an ambiguity averse investor in equilibrium is

$$d^a = \max\left\{\frac{\underline{\mu} - p}{\gamma\sigma^2}, 0\right\} \tag{4.11}$$

Then, the aggregate demand of ambiguity averse decision makers amounts to

$$D^a = \int_{\underline{\mu}}^{\hat{\mu}} \max\left\{\frac{\underline{\mu} - p}{\gamma\sigma^2}, 0\right\} m(d\mu_i) \tag{4.12}$$

$$= M(\hat{\mu}) \max\left\{\frac{\underline{\mu} - p}{\gamma\sigma^2}, 0\right\} \tag{4.13}$$

Ambiguity aversion has the equilibrium effect of preventing short selling by ambiguity averse agents. Aggregate demand of ambiguity averse agents corresponds to the demand in Miller (1977) where short selling restrictions where imposed exogenously. While ambiguity averse agents are internally constrained and only act as demander of the asset, SEU agents sell short if the price is high enough. The aggregate demand for the risky asset in the economy is given by the demand of the two groups of investors

$$D = D^s + D^a \tag{4.14}$$

In equilibrium,

$$D = Q \tag{4.15}$$

and we have the following lemma:

Lemma 11. *Under above conditions there exists a unique equilibrium in the market for the risky/uncertain asset with equilibrium price given by*

$$p = \begin{cases} \underline{\mu} + k^*(1 - M(\hat{\mu})) - Q\gamma\sigma^2 & \textit{if } k^*(1 - M(\hat{\mu})) < Q\gamma\sigma^2 \\ \underline{\mu} + k^* - \frac{Q\gamma\sigma^2}{1 - M(\hat{\mu})} & \textit{else} \end{cases} . \tag{4.16}$$

The first value in the above equation corresponds to the price when the ambiguity averse agents demand positive amounts of the asset, i.e. $p < \underline{\mu}$ and no short selling takes place. This price would arise in an unconstrained economy with unconstrained SEU maximizers $M(\hat{\mu})$ of them having belief $\underline{\mu}$ and the price equals the average opinion of all market participants .

The second value is the constrained equilibrium price when ambiguity averse agents do not demand the risky asset. Since the risk adjusted return of a long position in the worst case scenario is negative ambiguity averse agents do not demand the security and stay away from the market leaving it to the overly optimistic investors. However, they are reluctant to short sell the asset since the worst case return of the short position is negative as well. The price reflects the valuation of SEU agents only leading to a higher equilibrium price than in the pure heterogeneous expectation case. The effect of the reluctance to short sell is twofold. On the one hand, ambiguity averse agents stop demanding the asset, decreasing the scarcity of the asset by shifting the demand downwards. This potentially decreases the price. On the other hand, the agents with optimistic beliefs demand higher amounts of the asset causing higher prices.

4.3 Comparison with Miller (1977)

In this section we compare our results to the findings of Miller (1977). Recall that Miller (1977) assumed that heterogeneous investors are uniformly distributed across

$$[\underline{\mu}, \overline{\mu}] = [\hat{\mu} - k, \hat{\mu} + k] \tag{4.17}$$

The riskfree rate is zero and short selling is not allowed in this market. The maximization problem of the investor with belief μ_i then becomes

$$\text{Maximize } \mathbb{E}^i(-\exp(-\gamma d_i(x - p))) \text{ over } d_i \in \mathbb{R}^+ \tag{4.18}$$

and the individual demand amounts to

$$d^s = \max\left\{\frac{\mu_i - p}{\gamma\sigma^2}, 0\right\} \tag{4.19}$$

Aggregating over all investors and solving for equilibrium yields

$$p = \begin{cases} \hat{\mu} - Q\gamma\sigma^2 & \text{if } k < \gamma\sigma^2 Q \\ \hat{\mu} + k - 2\sqrt{kQ\gamma\sigma^2} & \text{else} \end{cases}. \tag{4.20}$$

The first value in the above equation denotes the equilibrium price that arises if the short selling constraint does not bind. The second value is the price resulting in the constrained equilibrium. Here, some of investors aim to hold negative amounts of the security but are prevented from it by the short selling constraint. It can be easily checked that the constrained price is higher than the unconstrained.

To compare our results we assume that the set of all possible beliefs is given by (4.17), where $\hat{\mu}$ denotes the ambiguity threshold and M is the uniform. Using Lemma 4.2.3 we can compute the equilibrium price as

$$p = \begin{cases} \frac{4\mu - k}{4} - Q\gamma\sigma^2 & \text{if } k < \frac{4}{3}Q\gamma\sigma^2 \\ \frac{2\mu + k}{2} - 2Q\gamma\sigma^2 & \text{else} \end{cases}.$$

Note that the price in our setting is always below the price in the setting of Miller (1977). This happens for two reasons. First, the average expected return in our model is lower due to the presence of ambiguity averse investors. This leads to a lower price in the unconstrained equilibrium. Second, while the model of Miller completely excludes short selling, some short selling takes place in our model in the restricted equilibrium. The opinion of moderate investors that aim to go short is contained in the restricted price of our model. Since some short selling is executed by moderate investors with belief μ_i such that $p > \mu_i > \hat{\mu}$, the overpricing is not as severe as in the model of Miller.

4.4 Comparative Statics and Sensitivity Analysis

In the next section we analyze how a change of parameters changes the equilibrium price. Our main goal is to study the impact of ambiguity on the equilibrium price in the meaningful way.

4.4.1 Sensitivity with respect to ambiguity increase

While classical equilibrium models[4] with ambiguity predict that an increase in ambiguity lowers the equilibrium prices the situation differs here. The two main factors for the sensitivity analysis is the distribution of opinions and the ambiguity threshold. First we analyze the sensitivity of the price with respect to changes in ambiguity and distribution of opinions. We study how the change in ambiguity threshold affects the equilibrium price, assuming that the distribution remains the same. This can happen due to an exogenous shock on the market such as an unexpected market outcome causing more agents to doubt their models.

Lemma 12. *Denote by* $p^u = \underline{\mu} + k^*(1 - M(\hat{\mu})) - Q\gamma\sigma^2$ *the unrestricted equilibrium price and by* $p^c = \underline{\mu} + k^* - \frac{Q\gamma\sigma^2}{1-M(\hat{\mu})}$ *the restricted equilibrium price. Then* p^u *and* p^c *satisfy*

1. $\frac{dp^u}{d\hat{\mu}} < 0$ *and*
2. $\frac{dp^c}{d\hat{\mu}} > 0$

An increase in ambiguity caused by an increase of the ambiguity threshold lowers the prices in the unconstrained equilibrium. This result is intuitive, since all investor participate in the market and demand the asset. The decrease in $\hat{\mu}$ decreases the average expected return and thus the aggregate demand. This result is in line with the literature. In the constrained equilibrium however, the picture is different. Here, an increase in ambiguity leads to higher prices on the market. This happens for the following reason. On the one hand the increase of the ambiguity threshold forces some market participants into the no trade interval and the market becomes smaller. At the same time moderate agents who where willing to short sell before are now

[4] See Epstein and Wang (1994) and references therein.

in the no trade interval and fail to generate short sale supply. If the concentration of optimists in the market is high enough they again can absorb the security available at the market and bid the price up. In extreme case where $\hat{\mu}$ is high enough there is no short selling at all and only the most optimistic investors determine the price.

Not only the ambiguity threshold is important for the equilibrium price but also the distribution of agents within intervals. In the following we analyze the impact of the change of the distribution of the agents along $[\underline{\mu}, \overline{\mu}]$. For example, due to a unfavorable outcome caused by a shock some agents that previously had the belief $\mu_i > \hat{\mu}$ may start to doubt their model and become ambiguity averse while others keep the belief μ_i and remain insensitive to ambiguity. More ambiguity on the market thus means in this case, that although the ambiguity threshold remains constant, the mass of ambiguity averse investors changes.

In this case the direction of the price change depends heavily on the shape of the distribution of beliefs before and after the shock.

Lemma 13. *Let P and Q be distributions on $[\underline{\mu}, \overline{\mu}]$ with absolutely continuous densities f, g. Assume that P and Q satisfy the hazard rate condition, i.e.*

$$\frac{f(x)}{1-F(x)} \leq \frac{g(x)}{1-G(x)} \text{ for all } \underline{\mu} \leq x \leq \overline{\mu} \tag{4.21}$$

i.e. $P \succeq_{hr} Q$. Denote by p_P resp. p_Q the price on the market where the agents are distributed according to P, Q resp. Then

1. *$p^c_P \geq p^c_Q$ and*
2. *$p^u_P \geq p^u_Q$*

One could also think of an increase in ambiguity by means of an increase of $\overline{\mu}$ or decrease of $\underline{\mu}$. However this analysis is not meaningful without specific assumptions on the underlying distribution m. For this reason we omit this analysis here.

4.4.2 Sensitivity to changes of Q, γ, σ^2

The standard market factors work in the usual direction decreasing the equilibrium price.

Lemma 14. *Denote by* $p^u = \underline{\mu} + k^*(1 - M(\hat{\mu})) - Q\gamma\sigma^2$ *the unrestricted equilibrium price and by* $p^c = \underline{\mu} + k^* - \frac{Q\gamma\sigma^2}{1-M(\hat{\mu})}$ *the restricted equilibrium price. Then* p^u *and* p^r *satisfy*

1. $\frac{dp^u}{dQ} < 0$ *and* $\frac{dp^c}{dQ} < 0$
2. $\frac{dp^u}{d\gamma} < 0$ *and* $\frac{dp^c}{d\gamma} < 0$
3. $\frac{dp^u}{d\sigma^2} < 0$ *and* $\frac{dp^c}{d\sigma^2} < 0$

These sensitivity results are in line with the standard theory and the economics intuition carries over as well. Investors decrease their demands if risk aversion resp. volatility increase, thus, as a result prices decrease as well. Since this effects are well understood we keep the discussion short and omit the proof.

4.5 Extensions and Robustness Checks

4.5.1 Uncertainty about volatility

In the generic model we assumed that all agents agree on the volatility of the underlying asset. We can extend the model easily to the case with heterogeneous beliefs about volatility. The essence of result does not change much.

Again agents in the economy differ in their beliefs about the return of the asset. Assume that the asset is normally distributed according to (μ, σ) where beliefs about the actual value of (μ, σ) are given by

$$\mathcal{P} := \{P : x_P \sim \mathcal{N}(\mu, \sigma^2), \text{ s.t. } (\mu, \sigma^2) \in [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}^2, \overline{\sigma}^2]\}$$

Every agent is endowed with a belief (μ, σ) from the above interval. The set of all possible beliefs can be partitioned into two regions: an ambiguity averse region $A \subset [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}^2, \overline{\sigma}^2]$ and a subjective region S satisfying

$$\mathcal{A} + \mathcal{S} = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}^2, \overline{\sigma}^2].$$

All agents that have a belief in $\mathcal{S}$ maximize their subjective utility given their belief while all agents in $\mathcal{A}$ maximize the minimal expected utility. The distribution M of agents across the interval is now two dimensional and we maintain the assumption that M has a density.

The analysis for the individual demand of an subjective utility investors carries over from the single dimensional case and we get

$$d_i^s = \frac{\mu_i - p}{\gamma \sigma_i^2} \tag{4.22}$$

resulting in aggregate demand

$$D^s = \int_{\mathcal{S}} d_i^s dm \tag{4.23}$$

$$= \int_{\mathcal{S}} \frac{\mu_i - p}{\gamma \sigma_i^2} dm \tag{4.24}$$

Using an appropriate version of the Mean Value Theorem we can show that

$$D^s = M(\mathcal{S}) \cdot \left(\frac{\mu^* - p}{\gamma (\sigma^*)^2} \right)$$

The effect of heterogeneity in the volatility may either decrease or increase the demand of SEU agents depending on the resulting average volatility σ^*. For the demand of ambiguity averse investors we only need to note that the highest variance minimizes their expected return. As in the standard literature on ambiguity averse portfolio choice we then get

$$d^a = \begin{cases} \frac{\underline{\mu} - p}{\gamma \overline{\sigma}^2} & \text{if } p < \underline{\mu} \\ 0 & \text{if } \underline{\mu} < p < \overline{\mu} \\ \frac{p - \overline{\mu}}{\gamma \overline{\sigma}^2} & \text{if } p > \overline{\mu} \end{cases} . \tag{4.25}$$

Similarly for the aggregate demand:

$$D^a = \int_{\mathcal{A}} \max\left\{\frac{\underline{\mu} - p}{\gamma\overline{\sigma}^2}, 0\right\} m(d\mu_i) \tag{4.26}$$

$$= M(\mathcal{A}) \max\left\{\frac{\underline{\mu} - p}{\gamma\overline{\sigma}^2}, 0\right\} \tag{4.27}$$

Here, the uncertainty about the volatility reduces the demand of the ambiguity averse agents. However, only the mean return is essential for the decision to short sell or buy the asset. From this point on the model essentially reduces to the single dimensional case. In the same manner as above we can perform the equilibrium analysis.

4.5.2 N ambiguous independent assets

Already Jarrow (1980) in his paper investigated the effect of adding securities to Miller's model. It turns out that the answer depends on the distribution of assets. If the assets are correlated substitution effects influence demand and prices of securities and the effect of short selling constraint may go in both directions. However, if market participants agree on volatility of the assets, the results of Miller (1977) carry over to the multiple asset case.

We can easily extend our model to the multiple asset case. In case of independent assets the analysis does not change much. Due to independence of assets and the form of the utility function the demand for each asset is determined separately for SEU agents. The same holds true for ambiguity averse agents. Thus, the price for each asset is set independently and the equilibria can be analyzed one by one with the same technique as above. However, we cannot distinguish anymore between unconstrained and constrained equilibria since some assets might be in the constrained equilibrium while other in the unconstrained.

4.5.3 Different distribution of preferences

In our model we assumed that all agents with belief more pessimistic than the ambiguity threshold $\hat{\mu}$ are not only pessimistic about returns of the stock but also about the model they use. This assumption simplifies the analysis considerably but is not essential for the result. The main result still holds true if we relax the assumption that there are no SEU maximizer with a belief more pessimistic than $\hat{\mu}$. In this case although ambiguity averse agents still refuse to go short for prices within $[\underline{\mu}, \overline{\mu}]$ SEU agents with pessimistic belief can reflect their opinion by going short. However, the overpricing still can occur if the short sell supply of pessimistic agents is not big enough to offset the overoptimistic demand.

4.5.4 Market Crashes and Panics

The model we derived can also be used to explain panics and crashes. Those can happen if heterogeneous agents become pessimistic and want to sell the security causing a sell out of security. As in the optimistic case, ambiguity averse refuse to correct this overreaction of the price due to their ambiguity. In this way pessimistic agents can bring the prices to crash.

4.6 Discussion and Conclusion

In this Chapter we analyzed the impact of ambiguity on the equilibrium price on markets with heterogeneous agents. Agents' sensitivity to both risk and uncertainty may impose short selling constraint on their portfolio. This short selling constraint in turn affects the equilibrium price in an economy with heterogeneous agents by increasing the equilibrium price. While the effect of short selling constraint itself was already known, the Chapter rationalize the short selling constraint on markets through preferences. The model considered here has also interesting implications for the regulation. After the beginning of the financial crisis a lively discussion has started on how to regulate the markets better in order to prevent investors to take huge risks and to avoid bubbles. One of suggestions made by the theorists was to impose a minimax regulation. Within this kind of regulations agents on the market have to consider several models instead one when assessing riskiness to a future payoff. The claim is then acceptable if its return is nonnegative under all models the agents consider. This imposes a more conservative value assignment preventing investors from excessive risk taking. However, this kind of regulation can have side-effects highlighted in this Chapter. Agents regulated in the above sketched way behave as ambiguity averse investors in our model. If the regulation is imposed only locally they might be investors on the global market who behave like SEU agents in our model. As we have seen in the model such kind of interaction may lead to overpricing if beliefs are heterogeneous. Thus, a minimax regulation although conservative form of regulation may help to generate bubbles if it is not established in a careful way.

Chapter 5

Concluding Remarks

Within the three chapters of this thesis we have studied several problems arising in the context of decision problems under Knightian uncertainty. Each chapter discusses its respective topic in detail and ends with a conclusion summarizing its results. Nevertheless for completeness we will briefly restate our achievements at this point:

First, we modeled and analyzed the best choice problem under ambiguity. We presented it in a way that allows to introduce ambiguity in the model in an intuitive way and also covers a broader class of problems. The main insight is that the optimal strategy of the problem under ambiguity has the same structure as the problem under risk. The stopping rule is simple again. However, the stopping here can occur later or earlier compared to the risky problem. The reason for this is the non-measurability of the payoff obtained from stopping.

In the third chapter of this work we analyze several exotic American options in a discrete financial market under ambiguity. Here, the holder of the option has to solve a multiple prior stopping problem in order to assign value to the payoff and to determine her optimal strategy. Unlike many examples analyzed in literature the payoffs we are considering cannot be reduced to a single measure. The worst-case measures arising here are path-dependent and change over time as the state process is realized. The reason for this involved structure is path-dependency, the non-monotonicity or the non-measurability of the considered payoff.

In the last chapter we change the perspective and consider a static financial market with heterogeneous investors. We show that when investors differ in their beliefs and their attitude towards ambiguity an increase in ambiguity can lead to a price increase. The reason for the lower premium paid in equilibrium is that ambiguity averse agents leave the market and fail to generate the short sale supply that is provided in the market by pessimistic agents. The result suggests that the coherent risk regulation that is supported by the scientific community as one of the lessons from the crisis can have side effects on markets with heterogeneous agents.

Appendix A

Proofs for Chapter 2

A.1 Proof of Lemma 1

Proof. Fix a stopping time τ with values in $\{1, \ldots, N\}$. Our set of priors is compact, hence we can choose $Q^k \in \mathcal{Q}$ that minimize $E^P X_k 1_{\{\tau=k\}}$. Time–consistency of the set of priors implies that there exists a measure $Q \in \mathcal{Q}$ such that

$$\sum_{k=1}^{N} E^{Q^k} X_k 1_{\{\tau=k\}} = E^Q \sum_{k=1}^{N} X_k 1_{\{\tau=k\}} = E^Q X_\tau \,,$$

see, e.g., Lemma 8 in Riedel (2009). It follows that we have

$$\inf_{P \in \mathcal{P}} E^P X_\tau = \inf_{P \in P} E^P \sum_{k=1}^{N} X_k 1_{\{\tau=k\}} = \sum_{k=1}^{N} \inf_{P \in \mathcal{Q}} E^P X_k 1_{\{\tau=k\}} \,.$$

By applying the law of iterated expectations for time–consistent multiple priors, this quantity is

$$= \sum_{k=1}^{N} \inf_{P \in P} E^P Z_k 1_{\{\tau=k\}}$$

and by applying time–consistency again, we get

$$= \inf_{P \in P} E^P Z_\tau \,.$$

□

A.2 Proof of Lemma 2

Proof. To show the independence we have to show that

$$\operatorname*{ess\,inf}_{P\in\mathcal{P}} P(Y_1 = y_1, \dots, Y_N = y_N) = \prod_{i=1}^{N} \hat{P}(Y_i = y_i)$$

for a $\hat{P} \in \mathcal{P}$, $y_i \in \{0, 1\}$ for $1 \le i \le N$
Because of definition of $\mathcal{P}$ all events of above kind have positive probability under every $P \in \mathcal{P}$, i.e. $P(Y_1 = y_1, \dots, Y_N = y_N) > 0$ for all sequences (y_i) with $y_i \in \{0, 1\}$ and all $P \in \mathcal{P}$. Therefore, using Bayes' rule and the fact that one-step-ahead probabilities $[a_n, b_n]$ depend only on time we get

$$\begin{aligned}
\min_{P\in\mathcal{P}} P[Y_1 = y_1, \dots, Y_N = y_N] &= \min_{P\in\mathcal{P}} P[Y_N = y_N|\mathcal{F}_{N-1}]P[Y_i = y_i, i < N] \\
&= \min_{P\in\mathcal{P}} \prod_{n=1}^{N} P[Y_n = y_n|\mathcal{F}_{n-1}] \\
&= \prod_{n=1}^{N} \min_{x_n\in[\alpha_n,\beta_n]} P_{x_n}[Y_n = y_n|\mathcal{F}_{n-1}] \\
&= \prod_{n=1}^{N} \min_{x_n\in[\alpha_n,\beta_n]} P_{x_n}[Y_n = y_n]
\end{aligned}$$

where P_{x_n} denotes the measure defined via $P_{x_n}[Y_n = y_n|\mathcal{F}_{n-1}] = x_n$.
In the last part of the equation we used the fact that the worst one-step-ahead probabilities depends only on time and not on the realization $\omega \in \Omega$. □

A.2.1 Proof of Lemma 3

Proof. The value process of the stopping problem for the payoff process (X_n) is given via

$$U_t = \operatorname*{ess\,sup}_{\tau\ge t} \mathbb{E}^Q[X_\tau|\mathcal{F}_t] \tag{A.1}$$

$$= \sup_{\tau\ge t} \mathbb{E}^Q[Y_\tau \cdot B_\tau|Y_1, \cdot, Y_t] \tag{A.2}$$

$$= \sup_{\tau\ge t} \mathbb{E}^Q[Y_\tau \cdot B_\tau] \tag{A.3}$$

$$=v_t \tag{A.4}$$

Having established the independence we can conclude that the sequence v_t is nonincreasing as the set of the arguments decreases. On the other hand, the sequence (B_n) is increasing and we get using the principle of backward induction Using this monotonicity observation we can conclude that if it is optimal to stop at Now suppose that it is optimal to take a candidate r. We have then $B_r = v_r$; therefore, we get

$$B_{r+1} \geq B_r = v_r \geq v_{r+1}\,,$$

and it is also optimal to stop when a candidate appears at time $r+1$. We conclude that optimal stopping rules are simple. □

A.3 Proof of Theorem 3

Proof. We denote by w_n^γ the sequence corresponding to the problem with the level of ambiguity γ. Straightforward calculations show that

$$w_n^\gamma = \sum_{k=n}^{N} \frac{\gamma^2}{k\gamma - 1} \prod_{l=n}^{N} \left(1 + \frac{1-\gamma^2}{l\gamma - 1}\right) \tag{A.5}$$

To prove robustness we first show

$$e^{-\frac{1}{\gamma}} \leq \frac{r_\gamma^*}{N} \leq e^{-\frac{2\gamma}{1+\gamma}} + \frac{3}{N} \tag{A.6}$$

For the left-hand side of A.6:

$$w_n^\gamma = \sum_{k=n}^{N} \frac{\gamma^2}{k\gamma - 1} \prod_{l=n}^{N} \left(1 + \frac{1-\gamma^2}{l\gamma - 1}\right) \geq \sum_{k=n}^{N} \frac{\gamma^2}{k\gamma - 1} \tag{A.7}$$

$$\geq \sum_{k=n}^{N} \frac{\gamma}{k} \tag{A.8}$$

$$\geq \int_n^N \frac{\gamma}{k} dk \tag{A.9}$$

$$= \gamma \log\left(\frac{N}{n}\right) \tag{A.10}$$

For the threshold r_γ we obtain

$$1 = w_n^\gamma \geq \gamma \log\left(\frac{N}{n}\right) \tag{A.11}$$

$$\Leftrightarrow \tag{A.12}$$

$$\frac{r_\gamma^*}{N} \leq e^{-\frac{1}{\gamma}} \tag{A.13}$$

For the second inequality:

$$\begin{aligned}
w_n^\gamma &= \sum_{k=n}^{N} \frac{\gamma^2}{k\gamma - 1} \prod_{l=n}^{N} \left(1 + \frac{1-\gamma^2}{l\gamma - 1}\right) \\
&= \sum_{k=n}^{N} \frac{\gamma^2}{k\gamma - 1} \exp\left(\sum_{l=n}^{k-1} \ln\left(1 + \frac{1-\gamma^2}{l\gamma - 1}\right)\right) \\
&\leq \sum_{k=n}^{N} \frac{\gamma^2}{k\gamma - 1} \exp\left(\sum_{l=n}^{k-1} \frac{1-\gamma^2}{l\gamma - 1}\right) \\
&\leq \sum_{k=n}^{N} \frac{\gamma^2}{k\gamma - 1} \exp\left(\int_{l=n-1}^{k} \frac{1-\gamma^2}{l\gamma - 1} dl\right) \\
&\leq \sum_{k=n}^{N} \frac{\gamma^2}{k\gamma - 1} \left(\frac{k\gamma - 1}{(n-1)\gamma - 1}\right)^{\frac{1-\gamma^2}{\gamma}}
\end{aligned}$$

Using $\alpha := \frac{1-\gamma^2}{\gamma} - 1$ we obtain for $\gamma \geq 0.5$

$$\begin{aligned}
w_n^\gamma &\leq \int_{n-1}^{N} \frac{\gamma^2}{((n-1)\gamma - 1)^{\alpha+1}} (k\gamma - 1)^\alpha \, dk \\
&\leq \frac{\gamma^2}{1-\gamma^2} \left[\left(\frac{N}{n-3}\right)^{\alpha+1} - 1\right]
\end{aligned}$$

By setting $w_n^\gamma = 1$ we get

$$1 \leq \frac{\gamma^2}{1-\gamma^2} \left[\left(\frac{N}{n-3}\right)^{\alpha+1} - 1\right]$$

$$\Leftrightarrow$$

$$
\begin{aligned}
\frac{n-3}{N} &\leq \gamma^{\frac{2\gamma}{1-\gamma^2}} \\
&\leq \exp\left(\ln(\gamma)\frac{2\gamma}{1-\gamma^2}\right) \\
&\leq \exp\left(\frac{(\gamma-1)2\gamma}{1-\gamma^2}\right) \\
&\leq e^{-\frac{2\gamma}{1+\gamma}}
\end{aligned}
$$

□

A.3.1 Proof of Lemma 4

Proof. Because of boundedness of w^∞ there exists a $R \in \mathbb{N}$, s.t.

$$\sum_{k=n}^{\infty} \beta_k \prod_{l=1}^{k-1} \alpha_l \leq 1 \text{ for all } n \geq R$$

and it follows that

$$r^*(N) \leq R \text{ for all } N$$

and

$$\frac{r^*(N)}{N} \leq \frac{R}{N} \to 0 \text{ for } N \to \infty$$

□

Appendix B

Proofs for Chapter 3

B.1 Proof of Theorem 6

Proof. We give proof for decreasing A in S_t for all $t \leq T$. The second case works similarly. For notational simplicity we write $\omega_{(t)}$ for an element in $\bigotimes_{i=1}^{t}\{0,1\} \subseteq \Omega$. Furthermore, for a stopping time τ we introduce for each $t \leq T$ the restriction τ^t of τ to paths in Ω running up to time t:

$$\tau^t : \bigotimes_{i=1}^{t} \{0,1\} \longrightarrow [0,t] \cup \{T+1\}$$

$$\omega_{(t)} \longmapsto \tau^t(\omega_{(t)}) = \begin{cases} \tau(\omega_{(t)}), & \text{if } \tau(\omega_{(t)}) \leq t \\ T+1, & \text{if else} \end{cases} .$$

The restricted stopping are being used in order to be mathematically more exact.

We start the proof with

Lemma 15. *Let $(U_t^{\mathcal{Q}})_{t \leq T}$ be the multiple prior Snell envelope of X as defined in Theorem 6. Assume that $U_t^{\mathcal{Q}}$ is given by the function $u(t, S_t, \tau_1^t, \tau_2^t)$ for all $t \leq T$. Then for all $t \in [0, T-1]$ and all $k \in [1, T-t]$*

$$u(t, S, t, T+1) \geq u(t+k, S, t+k, T+1).$$

Proof. The inequality follows directly by the inequality

$$u(t, S, t, T+1) \geq u(t+k, S, t, T+1) = u(t+k, S, t+k, T+1).$$

The inequality always holds for claims of American style whose payoff does only depend on the underlying's price S at each time. For the special choice of τ_1^t and τ_2^t it therefore also holds for the considered claims of the theorem. The equality holds since the claim is already knocked-in. □

Using theory of multiple prior Snell envelope, see Riedel (2009), we show by backwards induction that $U_t^Q = u(t, S_t, \tau_1^t, \tau_2^t)$ for all t such that u has the following properties:

(i) for $t < \tau_1^t :\ u\left(t, \cdot, \tau_1^t(\cdot), \tau_2^t(\cdot)\right) \nearrow$ in $S \leq \bar{S}^1$,
where $\bar{S}_t^1$ is determined by $\tau_1^t(\bar{S}^1) = t$

(ii) for $t \in [\tau_1^t, \tau_2^t[:\ u\left(t, \cdot, \tau_1^t(\cdot), \tau_2^t(\cdot)\right) \searrow$ in S

(iii) for $t \geq \tau_2^t :\ u\left(t, \cdot, \tau_1^t(\cdot), \tau_2^t(\cdot)\right) = 0$ for all S.

First, note that u is well-defined due to the definition of the payoff process X. (u complies with the definition of a function since X_t which only depends on S_t, τ_1^t, and τ_2^t, does for each $t \leq T$.) For $t = T$ we have

$$\begin{aligned} U_T^Q(\cdot) = X_T(\cdot) &= \mathbb{1}_{[\tau_1^T, \tau_2^T[}(T, \cdot)\, A(T, S_T(\cdot)) \\ &= \begin{cases} 0, & \text{if } \tau_1^T = T+1 \text{ or } \tau_2^T \leq T \\ A(T, S_T), & \text{if } \tau_2^T = T+1 \text{ and } T \geq \tau_1^T \end{cases} \\ &= \begin{cases} 0 = u(T, S_T, \tau_1^T, \tau_2^T)\ \forall S_T, & \text{if } \tau_1^T = T+1 \text{ or } \tau_2^T = T \\ A(T, S_T) = u(T, S_T, \tau_1^T, T+1)\ \forall S_T, & \text{if } \tau_1^T \leq T < \tau_2^T \end{cases} . \end{aligned}$$

So, U_T^Q satisfies the representation and the properties by the assumptions on $X_T, A(T, \cdot)$, respectively.

In the induction step for $t < T$ we handle the different cases separately. First, assume $t \in [\tau_1^t, \tau_2^t[$, say $\tau_1^t(\omega_{(t)}) =: k \leq t$: Then

$$\begin{aligned} U_t^{\mathcal{Q}}(\omega_{(t)}) &= \max\left\{X_t(\omega_{(t)}), \min_{P \in \mathcal{Q}} [U_{t+1} | \mathcal{F}_t(\omega_{(t)})]\right\} \\ &\overset{\text{(IH)}}{=} \max\left\{X_t(\omega_{(t)}), \min_{p_{t+1} \in [\underline{p}, \overline{p}]}\right. \\ &\quad \left.\left\{p_{t+1} u(t+1, S_t u, k, \tau_2^{t+1}(\omega_{(t)}, 1)) + (1 - p_{t+1}) u(t+1, S_t d, k, \tau_2^{t+1}(\omega_{(t)}, 0))\right\}\right\}. \end{aligned}$$

By induction hypothesis and due to $\tau_2^{t+1}(\omega_{(t)}, 0) \geq \tau_2^{t+1}(\omega_{(t)}, 1)$, properties (ii) and (iii) for $t+1$ imply
$u(t+1, S_t d, k, \tau_2^{t+1}(\omega_{(t)}, 0)) \geq u(t+1, S_t u, k, \tau_2^{t+1}(\omega_{(t)}, 1))$. Therefore,

$$\begin{aligned} U_t^{\mathcal{Q}}(\omega_{(t)}) &= \max\left\{X_t(\omega_{(t)}), \overline{p} u(t+1, S_t u, k, \tau_2^{t+1}(S_t u)) + (1 - \overline{p}) u(t+1, S_t d, k, T+1)\right\} \\ &= \hat{U}_t(\omega_{(t)}). \end{aligned}$$

Hence, in this case $U_t^{\mathcal{Q}}$ is a function $u(t, S_t, \tau_1^t, \tau_2^t)$ which is decreasing in S since $A(t, \cdot)$ is decreasing in S by assumption, and $u(t+1, \cdot, k, \tau_2^t(\cdot))$ is monotone decreasing in S by induction hypothesis (property (ii), (iii), respectively).
Second, if $t \geq \tau_2^t(\omega_{(t)}) =: l < T$, and $\tau_1^t(\omega_{(t)}) =: k < l$:

$$\begin{aligned} U_t^{\mathcal{Q}}(\omega_{(t)}) &= \max\left\{X_t(\omega_{(t)}), \min_{p_{t+1} \in [\underline{p}, \overline{p}]}\right. \\ &\quad \left.(p_{t+1} u(t+1, S_t u, k, l) + (1 - p_{t+1}) u(t+1, S_t d, k, l))\right\} \\ &= 0, \end{aligned}$$

since $X_t(\omega_{(t)}) = 0$ by assumption and $u(t+1, \cdot, k, l) = 0$ by induction hypothesis (property (iii)).
Third, assume the case $t < \tau_1^t(\omega_{(t)}) = T+1$:
Then $X_t = 0$ and therefore we get in the <u>first case</u> when $\tau_1^{t+1}(\omega_{(t)}, 1) = T+1$

$$\begin{aligned} U_t^{\mathcal{Q}}(\omega_{(t)}) &= \min_{p_{t+1} \in [\underline{p}, \overline{p}]} \left\{p_{t+1} u(t+1, S_t u, \tau_1^{t+1}(\omega_{(t)}, 1), T+1)\right. \\ &\quad \left.+ (1 - p_{t+1}) u(t+1, S_t d, \tau_1^{t+1}(\omega_{(t)}, 0), T+1)\right\} \\ &= \underline{p} u(t+1, S_t u, T+1, T+1) + (1 - \underline{p}) u(t+1, S_t d, T+1, T+1) \end{aligned}$$

by induction hypothesis (property (i)). Hence, $p_{t+1} = \underline{p}$ and $u(t, \cdot, T{+}1, T{+}1)$ is increasing in S.
In the <u>second case</u> when $\tau_1^{t+1}(\omega_{(t)}, 1) = t + 1$:

$$
\begin{aligned}
U_t^{\mathcal{Q}}(\omega_{(t)}) &= \min_{p_{t+1}\in[\underline{p},\overline{p}]} \{p_{t+1}u(t+1, S_t u, t+1, T+1) \\
&\qquad + (1-p_{t+1})u(t+1, S_t d, T+1, T+1)\} \\
&= \underline{p}u(t+1, S_t u, t+1, T+1) + (1-\underline{p})u(t+1, S_t d, T+1, T+1) \\
&= u(t, S_t, T+1, T+1)
\end{aligned}
$$

again by induction hypothesis (property (i) since $S_t \cdot u = \bar{S}^1$) and we obtain $p_{t+1} = p_{\tau_1^{t+1}} = \underline{p}$. In order to show the monotonicity note that by induction hypothesis (property (i)) the last expression is greater or equal to $\underline{p}u(t+1, S_t, T+1, T+1) + (1-\underline{p})u(t+1, S_t dd, T+1, T+1)$ which again is equal to $u(t, S_t d, T+1, T+1)$ (see the first case).
Thus, for showing property (i) we just have to prove that $u(t, \bar{S}^1, t, T+1) \geq u(t, \bar{S}^1 d, T+1, T+1)$. Using property (i) of induction hypothesis we obtain

$$
\begin{aligned}
u(t, \bar{S}^1 d, T+1, T+1) &= \underline{p}u(t+1, \bar{S}^1, t+1, T+1) \\
&\qquad + (1-\underline{p})u(t+1, \bar{S}^1 \cdot d^2, T+1, T+1) \\
&\leq \underline{p}u(t+1, \bar{S}^1, t+1, T+1) \\
&\qquad + (1-\underline{p})u(t+1, \bar{S}^1, t+1, T+1) \\
&= u(t+1, \bar{S}^1, t+1, T+1) \\
&\leq u(t, \bar{S}^1, t, T+1).
\end{aligned}
$$

The last inequality is due to Lemma 15. This completes the proof and $(U_t^{\mathcal{Q}})$ satisfies the same recursion as $(\hat{U}_t)$. Thus, $(U_t^{\mathcal{Q}}) = (\hat{U}_t)$ follows and the worst-case measure $\hat{P}$ is specified by the density $\hat{D}_T$ as claimed.

An optimal stopping time is given by $\hat{\tau}$. This follows by general theory, see Riedel (2009). The time boundary σ of the optimal stopping rule is due to the claim's knock-out feature. □

B.2 Downcrossing Times Theorem

A result similar to Theorem 6 also holds for downcrossing times. The only difference is the monotonic behavior of X and U which changes for downcrossing times. As a consequence, the densities of the worst-case measures change. So, we will state the theorem without giving the proof since it would be almost a copy of of the above proof.

Theorem 7. *Take the same assumptions as in Theorem 6 except the one for τ_1 and τ_2 being now either down-crossing times or constant again. Thus, assume $S_0 > H_1 > H_2$.*

1. *If $A(t,\cdot)$ is decreasing in S for all t, the multiple prior Snell envelope is $U = U^{\hat{P}}$ and the worst-case measure $\hat{P}$ is given by the density*

$$\hat{D}_t := 2^t \prod_{u \leq t \wedge \tau_2:\ u \neq \sigma_i + 1} \big(\varepsilon_u \overline{p} + (1-\varepsilon_u)(1-\overline{p})\big) \prod_{u \leq t:\ u = \sigma_i + 1} \big(\varepsilon_u \underline{p} + (1-\varepsilon_u)(1-\underline{p})\big) \prod_{u \in \,]\tau_2, t \wedge T]} \big(\varepsilon_u \underline{p} + (1-\varepsilon_u)(1-\underline{p})\big)$$

   for all $t \leq T$ and all occurring $1 \leq i \leq T$. An optimal stopping rule under ambiguity is given by
   $\hat{\tau} = \inf\left\{t \in [\tau_1, \sigma_1] \,\middle|\, X_t = U_t^{\hat{P}}\right\} \wedge T.$

2. *If $A(t,\cdot)$ is increasing in S for all t, the multiple prior Snell envelope is $U = U^{\hat{P}}$ and the worst-case measure $\hat{P}$ is given by the density*

$$\hat{D}_t := 2^t \prod_{u \leq t \wedge \tau_1} \big(\varepsilon_u \overline{p} + (1-\varepsilon_u)(1-\overline{p})\big) \prod_{u \in \,]\tau_1, t \wedge T]} \big(\varepsilon_u \underline{p} + (1-\varepsilon_u)(1-\underline{p})\big)$$

   for all $t \leq T$. An optimal stopping rule under ambiguity is given by
   $\hat{\tau} = \inf\left\{t \in [\tau_1, \sigma_1] \,\middle|\, X_t = U_t^{\hat{P}}\right\} \wedge T.$

Appendix C

Proofs for Chapter 4

C.1 Proof of Lemma 12

The proof is simple calculations. For the first part we get

$$p^u = \left[\int_{\hat{\mu}}^{\overline{\mu}} \mu_i \cdot m(\mu_i) d\mu_i\right] \cdot (1 - M(\hat{\mu})) + \underline{\mu} \cdot M(\hat{\mu}) - Q\gamma\sigma^2$$

and thus

$$\frac{dp^u}{d\hat{\mu}} = -m(\hat{\mu})\hat{\mu} + m(\hat{\mu})\underline{\mu} < 0$$

For the second part we get

$$p^c = \left[\int_{\hat{\mu}}^{\overline{\mu}} \mu_i \cdot m(\mu_i) d\mu_i\right] - Q\gamma\sigma^2/(1 - M(\hat{\mu}))$$

and thus for the derivative

$$\frac{dp^c}{d\hat{\mu}} = -\hat{\mu} \cdot m(\hat{\mu})(1 - M(\hat{\mu})) + m(\hat{\mu})\left((\underline{\mu} + k^*)(1 - M(\hat{\mu})) - Q\gamma\sigma^2\right)/(1 - M(\hat{\mu}))^2$$

Now since $\hat{\mu} < \underline{\mu} + k^*$ we get the desired inequality.

C.2 Proof of Lemma 13

Proof. We first consider the restricted equilibrium price defined in (4.16)

$$\begin{aligned} p^c =& \int_{\hat{\mu}}^{\overline{\mu}} \mu \cdot f(\mu) d\mu - Q\gamma\sigma^2/(1-F(\hat{\mu})) \\ =& \mathbb{E}^P(\mu|\mu > \hat{\mu}) - \frac{Q\gamma\sigma^2}{1-F(\hat{\mu})} \end{aligned}$$

Since $P \succeq_{hr} Q$ we know that

$$\mathbb{E}^P(\mu|\mu > \hat{\mu}) \geq \mathbb{E}^Q(\mu|\mu > \hat{\mu})$$

and

$$(1 - F(\hat{\mu})) \geq (1 - G(\hat{\mu}))$$

which implies the result.

Now to the second part: From (4.16) we know that the price in the unrestricted equilibrium is given by

$$p^u = M(\hat{\mu}) \cdot \underline{\mu} + \int_{\hat{\mu}}^{\overline{\mu}} \mu \cdot m(\mu) d\mu - Q\gamma\sigma^2 \tag{C.1}$$

$$= \int_{\hat{\mu}}^{\overline{\mu}} (\mu - \underline{\mu}) \cdot m(\mu) d\mu - (Q\gamma\sigma^2 + \underline{\mu}) \tag{C.2}$$

where M denotes the distribution of agents and m its density. Now consider the prices on the market with distributions P and Q. Since (4.21) implies the first order stochastic dominance we get

$$\int_{\hat{\mu}}^{\overline{\mu}} (\mu - \underline{\mu}) \cdot f(\mu) d\mu \leq \int_{\hat{\mu}}^{\overline{\mu}} (\mu - \underline{\mu}) \cdot g(\mu) d\mu$$

□

Bibliography

ANSCOMBE, F., AND R. AUMANN (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199–205.

ARTZNER, P., F. DELBAEN, J.-M. EBER, AND D. HEATH (1999): "Coherent Measures of Risk," *Mathematical Finance*, 9, 203–228.

BECK, H. (2005): *Der Liebesökonom – Nutzen und Kosten einer Himmelsmacht*. Frankfurter Allgemeine Buch.

BEREZOVSKI, B., AND A. V. GNEDIN (1984): *Problems of Best Choice (in Russian)*. Akademia Nauk, Moscow.

BEWLEY, T. (2002): "Knightian Decision Theory: Part I," *Decisions Econom. Finance*, 25, 79 – 110, First version: 1986.

BOYLE, L. GARLAPPI, R. U., AND T. WANG (2009): "Keynes meets Markowitz: The tradeoff between familiarity and diversifcation," Unpublished manuscript.

BRUSS, F. T. (2000): "Sum the Odds to One and Stop," *The Annals of Probability*, 28(3), 1384–1391.

BUCHEN, P. (2004): "The Pricing of Dual-expiry Exotics," *Quantitative Finance, 4, 101-108*.

Caballero, R. J., and A. Krishnamurthy (2007): “Collective Risk Management in a Flight to Quality Episode,” *NBER Working Paper No. W12896*.

Cao, H. H., T. Wang, and H. H. Zhang (2005): “Model Uncertainty, Limited Market Participation, and Asset Prices,” *The Review of Financial Studies*, 18(4), 1219–1251.

Chow, Y., H. Robbins, and D. Siegmund (1971): *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin Comp., Boston.

Cont, R. (2006): “Model Uncertainty and its Impact on the Pricing of derivative Instruments,” *Mathematical Finance*, 16(3), 519–547.

Cox, J., and M. Rubinstein (1985): *Options Markets*. Prentice Hall.

Cox, J. C., S. A. Ross, and M. Rubinstein (1979): “Option Pricing: A Simplified Approach,” *Journal of Financial Economics*, 7, 229–263.

Delbaen, F. (2002): “The Structure of m-Stable Sets and in Particular of the Set of Risk neutral Measures,” *mimeo, ETH Zürich*.

Dow, J., and S. Werlang (1992): “Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio,” *Econometrica*, 60(1), 197–204.

Easley, D., and M. O’Hara (2009): “Ambiguity and Nonparticipation: The Role of Regulation,” *The Review of Financial Studies*, 22(5), 1817–1843.

Ellsberg, D. (1961): “Risk, Ambiguity, and the Savage Axioms,” *Quarterly Journal of Economics*, 75, 643–669.

Engelage, D. (2009): “Essays on Multiple Priors,” Ph.D. thesis, Bonn Graduate School of Economics, Bonn, Germany.

EPSTEIN, L., AND Z. CHEN (2002): "Ambiguity, Risk and Asset returns in Continuous Time," *Econometrica*, 70, 1403–1443.

EPSTEIN, L., AND J. MIAO (2003): "A Two–Person dynamic Equilibrium under Ambiguity," *Journal of Economic Dynamics and Control*, 27, 1253–1288.

EPSTEIN, L., AND M. SCHNEIDER (2003a): "IID: Independently and Indistinguishably Distributed," *Journal of Economic Theory*, 113(1), 32–50.

——— (2003b): "Recursive Multiple Priors," *Journal of Economic Theory*, 113, 1–31.

EPSTEIN, L. G., AND T. WANG (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62(2), 283–322.

FELLER, W. (1968): *An Introduction to Probability Theory and Its Applications. Volume 1, 3rd Edition*. Wiley.

FERGUSON, T. (2006): "Optimal Stopping and Applications," Electronic Text, http://www.math.ucla.edu/∼tom/Stopping/Contents.html, University of California, Los Angeles.

FERGUSON, T. S. (1989): "Who Solved the Secretary Problem?," *Statistical Science*, 4(3), 282–289.

FÖLLMER, H., AND A. SCHIED (2004): *Stochastic Finance: An Introduction in Discrete Time. 2nd Edition*. de Gruyter.

FREEMAN, P. (1983): "The secretary Problem and its Extensions — a Review," *International Statistical Review*, 51, 189–206.

GARDNER, M. (1960): *Scientific American*.

GILBOA, I., AND D. SCHMEIDLER (1989): "Maxmin Expected Utility with Non–Unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

Hansen, L., and T. Sargent (2001): "Robust Control and Model Uncertainty," *American Economic Review Papers and Proceedings*, 91, 60–66.

Harrison, J. M., and D. M. Kreps (1978): "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *The Quarterly Journal of Economics*, 92(2), 323–336.

Hull, J. C. (2006): *Options, futures, and other derivatives. 6th ed.* Prentice-Hall International Editions. Upper Saddle River, NJ: Prentice Hall.

Illeditsch, P. (2009): "Ambiguous information, risk aversion and asset pricing," *Unpublished manuscript, Wharton.*

Ingersoll, J. (2007): "Digital Contracts: Simple Tools for Pricing Complex Derivatives," *Journal of Business*, 73, 67–88.

Jarrow, R. (1980): "Heterogeneous Expectations, Restrictions on Short Sales, and Equilibrium Asset Prices," *The Journal of Finance*, 35(5), 1105–1113.

Karatzas, I., and S. Kou (1998): "Hedging American Contingent Claims with Constrained Portfolios," *Finance and Stochastics*, 2, 215–258.

Karatzas, I., and I. Zamfirescu (2003): "Game Approach to the Optimal Stopping Problem," Working Paper.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73(6), 1849–1892.

Knight, F. (1921): *Risk, Uncertainty, and Profit.* Hart, Schaffer, and Marx, Boston, MA.

Kogan, L., and T. Wang (2002): "A Simple Theory of Asset Pricing under Model Uncertainty," Working Paper.

Kort, P., and M. Trojanowska (2007): "The Worst Case for Real Options," *Mimeo.*

Kwok, Y.K. Dai, M., and L. Wu (2004): "Optimal shouting policies of options with strike reset rights," *Mathematical Finance*, 14(3), 383–401.

Lamont, O. A., and J. C. Stein (2004): "Aggregate Short Interest and Market Valuations," *The American Economic Review*, 94(2), 29–32.

Lerche, R., R. Keener, and M. Woodroofe (1994): "A Generalized Parking Problem," in *Statistical Decision Theory and Related Topics V*, ed. by J. B. S.S. Gupta, pp. 523–532. Springer Verlag.

Lintner, J. (1969): "The Aggregation of Investor's Diverse Judgments and Preferences in Purely Competitive Security Markets," *The Journal of Financial and Quantitative Analysis*, 4(4), 347–400.

Lutgens, F., and P. C. Schotman (2010): "Robust Portfolio Optimisation with Multiple Experts," *Review of Finance*, 14, 343–383.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74(6), 1447–1498.

Maenhout, P. J. (2004): "Robust Portfolio Rules and Asset Pricing," *The Review of Financial Studies*, 17(4), 951–983.

Mehra, R., and E. C. Prescott (1985): "The equity premium: A puzzle," *Journal of Monetary Economics*, 15(2), 145 – 161.

Miao, J., and N. Wang (2004): "Risk, Uncertainty, and Option Exercise," Working Paper.

Miller, E. M. (1977): "Risk, Uncertainty, and Divergence of Opinion," *The Journal of Finance*, 32(4), 1151–1168.

Mukerji, S., and J.-M. Tallon (2001): "Ambiguity Aversion and Incompleteness of Financial Markets," *The Review of Economic Studies*, 68(4), 883–904.

Murota, K. (1998): "Discrete Convex Analysis," *Mathematical Programming*, 83(1-3), 313–371.

Nishimura, K. G., and H. Ozaki (2004): "Search and Knightian Uncertainty," *Journal of Economic Theory*, 119, 299–333.

Nishimura, K. G., and H. Ozaki (2007): "Irreversible Investment and Knightian Uncertainty," *Journal of Economic Theory, 136, 668-684.*

Peng, S. (1997): "BSDE and Related g–Expectation," in *Backward Stochastic Differential Equation*, ed. by N. El Karoui, and L. Mazliak, no. 364 in Pitman Research Notes in Mathematics Series. Addison Wesley Longman Ltd.

Presman, E., and I. Sonin (1975): "Equilibrium Points in a Game Related to the Best Choice Problem," *Theory of probability and its Applications*, 20, 770–781.

Reimer, M., and K. Sandmann (1995): "A discrete Time Approach for European and American Barrier Options," *Discussion paper B-272.*

Riedel, F. (2009): "Optimal Stopping with Multiple Priors," *Econometrica*, 77, 857–908.

——— (2010): "Optimal Stopping Under Ambiguity in Continuous Time," *IMW working paper 429.*

Routledge, B., and S. Zin (2009): "Model uncertainty and liquidity," *Rev. Econ. Dyn.*, 12.

Saffi, P. A. C., and K. Sigurdsson (2008): "Price Efficiency and Short Selling," *AFA 2008 New Orleans Meetings Paper*.

Savage, L. (1954): *The Foundations of Statistics*. Wiley, New York.

Scheinkman, J. A., and W. Xiong (2003): "Overconfidence and speculative Bubbles," *The Journal of Political Economy*, 111(6), 1183–1219.

Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57, 571–587.

Shleifer, A., and R. W. Vishny (1997): "The Limits of Arbitrage," *The Journal of Finance*, 52(1), 35–55.

Snell, L. (1952): "Applications of the Martingale Systems Theorem," *Transactions of the American Mathematical Society*, 73, 293–312.

Todd, P. (2009): "Searching for the Next Best Mate," Working Paper, Center for Adaptive Behavior and Cognition, Munich University.

Ui, T. (2009): "The Ambiguity Premium vs the Risk Premium under limited Market Participation," *Unpublished manuscript, Yokohama Natl. Univ.*

Uppal, R., and T. Wang (2003): "Model Misspecification and Underdiversification," *The Journal of Finance*, 58(6), 2465–2486.

Von Neumann, J., and O. Morgenstern (1953): *Theory of Games and Economic Behavior*. Princeton, Princeton University Press.