UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

AMBIGUITY AND ECONOMIC MODELS

DISSERTATION ZUR ERLANGUNG DES GRADES

DOCTOR RERUM POLITICARUM

AN DER

FAKULTÄT FÜR WIRTSCHAFTS- UND

SOZIALWISSENSCHAFTEN

DER RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG

VORGELEGT VON

BORIS ROLAND WIESENFARTH

OKTOBER 2015

Preface

The research described in this dissertation is submitted for the degree of Doctor of Philosophy at Heidelberg University (Dr. rer. pol.) and was carried out under the supervision of Professor Jürgen Eichberger in the Department of Economic Theory I at Heidelberg University between October 2010 and September 2015. Chapter 2 and Chapter 4 of this thesis are the outcome of collaborative work with Florian Kauffeldt *(Chapter 2)* and Daniel Heyen *(Chapter 4)*. The remaining chapters are based on entirely autonomous research.

When I started working on my dissertation, I was not yet aware of how exciting and challenging the implementation of this project was going to be. Now that the final version of this thesis is available, it is a pleasure to express my gratitude to those people that accompanied and supported me during my time at Heidelberg University.

First of all, I am deeply grateful to my supervisor Jürgen Eichberger, who supported me in many respects. Without his guidance, generous support, and insightful comments at all stages of this project, this thesis would not have been possible. Jürgen gave me the opportunity to develop my own research agenda and to conduct research in accordance with my own interests. I would also like to show my great appreciation to Jörg Oechssler who kindly agreed to act as a second supervisor of this thesis.

Furthermore, I would like to offer special thanks to my coauthors, Daniel Heyen and Florian Kauffeldt, who worked with me on some of the papers incorporated in this dissertation. Discussions with them have been both valuable and illuminating.

In addition, I am indebted on the participants of the 14^{th} annual meeting of the German Economic Association in Magdeburg, the 13^{th} Annual International Conference on Health Economics, Management and Policy in Athens, and the 11^{th} World Congress on Health Economics in Milan, as well as the participants of the workshop on Decision Making under Uncertainty at Mainz University.

I would also like to show my great appreciation to the participants of the weekly internal seminars at Heidelberg University for their insightful comments and suggestions. Particularly Jörg Oechssler, Christoph Vanberg, Jonas Hedlund, and Peter Dürsch provided advice and comments that were of enormous help to me.

Moreover, I would like to express my gratitude to all my colleagues at Heidelberg University. In particular, I am deeply grateful to Gabi Rauscher, our secretary, and Anna Larsson from the Academic Writing Support service, who kindly agreed to eliminate errors and provided comments at the crucial final stage of this dissertation. Furthermore, I would like to thank my colleagues, Christoph Brunner and Andreas Reischmann, as well as Freya Schadt for an excellent working atmosphere.

I wish to thank my parents, Roland and Claudia Wiesenfarth, for their love and support throughout my life. Last, but not least, I owe my deepest gratitude to my wife Sophea. Her loving care, moral support, and warm encouragement gave me strength and motivation. Finally, I want to thank my daughter Veronika for coming into the world during that time. The joy of being a father was an important additional source of motivation during the final stage of this dissertation.

Heidelberg, October 2015

Boris R. Wiesenfarth

Contents

List of Figures

List of Tables

For my family…

Chapter 1

Introduction

Throughout the last decades new theories of decision making under uncertainty have increasingly found their way into economic models and applications. One of the main objectives of this thesis is to investigate which additional insights one gains by introducing "Knightian uncertainty", or "ambiguity", into well-established economic models. By now, there is a large number of scientific articles from different areas of economics and related sciences discussing the implications of ambiguity for their respective field. In a sense, ambiguity has become highly topical and an interesting object of research for economists around the world.

In this spirit, I consider ambiguity related to a variety of applications ranging from Industrial Organization, Health Economics to Information Economics. Even though these applications stem from different areas of economics, it turns out that there is a common theme and methodology connecting them. Moreover, ambiguity might provide an additional source of explanation for a variety of observed deviations from standard expected utility theory, in cases where reliable information is absent, incomplete, or when decision-makers base decisions on unverifiable and contradictory information.

In the introductory chapter, I provide a brief overview on models of decision making under uncertainty relevant for this thesis. In this respect, my primary concern is to provide the basic background knowledge needed to understand the term "ambiguity" and its significance for economics. The main focus is thereby to sketch briefly the historical development of decision theory and resulting models by means of simple and manageable examples. The introduction is divided into two parts. The first part deals with static models of decision making and the second part considers extensions of the classical static models to dynamic settings. Readers, who dispose of sound knowledge of decision theory, might skip chapter 1 and proceed with chapter 2 right from the start.

The first field of application for ambiguity treated in this thesis is spatial competition between firms. This is done in chapter 2, which is based on the article Kauffeldt and Wiesenfarth [2014]. In this study, we analyze the impact of ambiguity and ambiguity attitude on product differentiation in a Hotelling duopoly game. The main contribution of this article is to investigate how partial probabilistic information on consumer demand shapes equilibrium product designs. Therefore, we suggest a general and tractable formal framework assuming that firms exhibit Choquet-expected utility preferences. More specifically, we assume that firm managers' beliefs are represented by neo-additive capacities. In this context, we highlight the importance of partial probabilistic information for observed product design behavior by shedding new light on a variety of real-world applications of Hotelling models under uncertainty treated in the literature. We find that, in many cases, their interpretations are not robust with respect to this novel aspect of uncertainty.

Chapter 3 of this thesis is based on the article Wiesenfarth [2015] and contemplates ambiguity in the context of preventive health care. Using a theoretical framework, I investigate how information on the efficacy of a preventive measure affects patients' preventive activities under Knightian uncertainty. Information is modeled by means of an imprecise signal, and patients' preferences are assumed to be of the Choquet-type with beliefs in the form of neo-additive capacities. It turns out that Knightian uncertainty can, depending on the underlying updating rule, provide an explanation for poor patient compliance as well as excessive preventive behavior. Moreover, I can demonstrate that information might reinforce extreme behavior under Knightian uncertainty, even if information is correctly communicated.

Chapter 4 is based on a note published in the Journal of Mathematical Economics by Heyen and Wiesenfarth [2015] and relates to a recent article by Çelen [2012]. The author generalizes the well-known Blackwell's theorem to MEU-preferences. We show that the notion of the value of information used in Çelen [2012] generates dynamically inconsistent behavior. The reason for this observation is that Çelen'sdefinition of the value of information is in conflict with the principle recursively defined utility. As a consequence of this finding, we propose an alternative, recursively defined value of information under Knightian uncertainty.

In chapter 5, I contemplate the approach and results of this thesis from a meta-perspective. The connecting element between the different chapters of this thesis can be found in the fact that I consider a variety of well-established economic models and investigate how the introduction of ambiguity alters the conclusions drawn from these models. Inherent to such research assignment is an underlying process of model selection and adjustment. As it turns out, this process follows clear rules and procedures, which I present by means of a very simple and tractable baseline model, the monopoly market with demand ambiguity. First, I treat the question of how to assess whether it is admissible and reasonable to introduce ambiguity into a particular economic model or not. Next, I assume that the first question can be answered positively. Due to the availability of a growing number of competing models of decision making under uncertainty, it is not easy for practitioners to see, which of these models is appropriate for a specific baseline model. For this reason, I investigate the monopoly problem under demand ambiguity for a variety of preference specifications and compare the conclusions drawn from each specification.

Chapter 6 provides a short conclusion of the main findings of this thesis.

1.1 Models of Decision Making Under Uncertainty

In contrast to a decision problem under certainty where one or several decision-makers have to make a decision knowing which out of several consequences will be triggered off given a certain action, a decision problem under uncertainty is characterized by a situation where this mechanism remains unclear. In order to illustrate this, contemplate the following abstract example. There is a farmer selling his organic vegetables on a local market. The farmer disposes of two different choices of action, he can sell his products in city a or in city b. Due to his long-lasting experience over the last twenty years, the farmer knows that organic food is more popular in city b than in city a. More precisely, he knows that if he sells his vegetables in city a, he will get 100 €, if he sells his products in city b, he will obtain 150 €. If the farmer prefers more money to less money, he will decide to sell his products in city b. This problem is a decision problem under certainty, the decision-maker knows how a certain action affects his future payoffs. The subsequent example is an illustration for a decision problem under uncertainty. An ice-cream seller with an ice cream cart can sell her ice-cream at two different locations c and d. Location c is inside a shopping mall, location d is in the center of the city's historical marketplace, where many tourists pass by if the weather is nice. Due to her experience, the ice-cream seller knows that if she decides to sell in the shopping mall, she will get 100 € if the weather is nice, let me denote this event with ω_1; and she will get 150 € if it rains, let me denote this event with ω_2. If she decides to sell her ice cream in the center of the historical marketplace, she will get 150 € if the weather is nice, and 100 € if it rains. This example can be illustrated in a simple diagram.

action	**no rain**	**rain**
mall	100 €	150 €
marketplace	150 €	100 €

TABLE 1.1: Decision Problem of the Ice-Cream Seller

In contrast to the farmer, the ice-cream seller's decision problem is more intricate. Which decision should she take? The answer depends on her information and assessment of the events "it rains" and "it doesn't rain". If she knew on the one hand with certainty that it was going to rain today, she would sell her ice-cream in the mall. On the other hand, if she knew with certainty that the weather was going to be fine, she would like to sell her ice-cream at the marketplace. Thus, the ice-cream seller's optimal answer depends on which of the two states of the world "rain" and "no rain" is going to materialize. Evidently, it is unrealistic to assume that the ice-cream seller knows every day with certainty which of these events is going to occur. Assume, for example, that she heard in the local weather forecast that the probability of rain is 20% today, and assume furthermore that she fully trusts this weather forecast. In this case, her expected value or expectation given the lottery[1] $L_1 = (100€, 0.8; 150€, 0.2)$ is $EV_1 = 0.8 \cdot 100$ €$+\ 0.2 \cdot 150$ €$= 110$ € if she goes to the mall. Given the lottery $L_2 = (150€, 0.8; 100€, 0.2)$, one obtains the expected value $EV_2 = 0.8 \cdot 150$ €$+0.2 \cdot 100$ €$= 140$ € if she goes to the marketplace. Thus, given this calculation of expected values based on her "belief" in the weather forecast, and the assumption that the ice-cream seller prefers more money to less money, she would decide to sell her ice cream at the local marketplace.

A closer look at the seller's decision problem shows that her choice of different actions is equivalent to a choice between two different lotteries L_1 and L_2, each one them defined on a common set of possible consequences $X = \{100, 150\} = \{x_1, x_2\}$. In fact, both levels of decision making are connected in a general way via the expected utility model (henceforth EU model) axiomatized by Von Neumann and Morgenstern [1944].

The authors show that a decision-maker's preference on the set of (finite) lotteries $\mathcal{L}$ satisfies the axioms of completeness, transitivity, continuity, and independence if and only if there exists a utility function $\mathcal{U} : \mathcal{L} \to \mathbb{R}$ defined on the set of lotteries $\mathcal{L}$ and a Bernoulli utility function $u : X \to \mathbb{R}$ defined on the set of consequences such that

$$\mathcal{U}(L) = \sum_{i=1}^{n} p_i u(x_i) \tag{1.1}$$

where $p_i = L(x_i)$ the probability of the outcome x_i. Thus, the above stated decision

[1] A (finite) lottery is a probability distribution defined on a (finite) state space Ω. The state space is a set comprising all possible states of the world. In this example, we have $\Omega = \{\omega_1, \omega_2\}$. Furthermore, I introduce the notation $L_i = (x_1, p_1; ..., x_n, p_n)$ where x_i denotes the outcome of the lottery in state ω_i and p_i denotes the probability of state ω_i for $i = 1, ..., n$.

problem of the ice-cream seller implicitly assumes that her Bernoulli utility function is given by the identity function $u(x) = x$.

One of the major challenges of the ice-cream decision problem is that the seller heavily depends on the "objectively" given rainfall probability provided by the local weather forecast. Assume for example that due to some reason, the seller has no access to the information provided by the weather forecast. Still, she has to make a decision. But on which probability should her decision be based on? The underlying problem is extensively discussed in the literature and dates back to Knight [1921], who differentiates between "calculable" and "incalculable risk". Another term for incalculable risk is "ambiguity". In the stylized ice-cream seller example, the risk calculation of the rainfall probability has already been performed by the meteorological service. In absence of this crucial information, the seller needs to rely on her individual observations and judgments to form a purely subjective belief. What is the difference between objective and subjective probabilities? A subjective probability is tied to a single individual and arises in situations of scarce information on the underlying randomization process. An objective probability is a probability that is based on reliable and and plausible data or background information on the randomization device. Imagine, for instance, that we observe the outcome of a pseudorandom number generator or of a coin toss with a fair coin. In both cases, we know the underlying probability distribution that generates the outcomes. If, in one of these cases, the ice-cream seller met another ice-cream seller disposing of the same reliable information on the randomization process and acting rational[2] like her, then both would agree on the same objective probability.

At this particular point, the problem of subjective probabilities is that we do not know whether they exist or not. A certain progress regarding this question has been achieved by Ramsey [1931], who suggests a procedure of subjective belief elicitation by observing an individual's willingness to pay for certain bets. By showing preference for certain bets, the decision-maker reveals which events he judges more likely than other events given his

[2] A decision-maker is considered rational when she consistently conforms to a certain set of laws or axioms.

or her current state of knowledge. Thus, the task associated to demonstrating whether a subjective belief exists or not is to derive a probability distribution directly from an individual's preference. Given Ramsey's result, it remains questionable which properties a preference relation needs to satisfy in order to allow for belief elicitation. De Finetti [1937] addresses this problem by providing a set of axioms on the decision-maker's preference such that a subjective probability exists. Moreover, the decision-maker selects an action by comparing expectations based on this subjective belief. The problem of De Finetti's approach is that the agent compares bundles by comparing expectations with respect to the implicitly given subjective belief, in contrast to Von Neumann and Morgenstern [1944] who obtain a utility function and represent preferences over a set of objective lotteries. Savage [1954] reconciles the advantages of Von Neumann and Morgenstern [1944] and De Finetti [1937] by axiomatizing the so-called subjective expected utility model (henceforth SEU-model). If a decision-maker's preference conforms to Savage's set of axioms, we can infer that a subjective belief as well as a representing utility function exist, and that the latter selects an action such that she maximizes her subjective expected utility. Instead of selecting between different actions from a choice set, a decision-maker in Savage's world selects among different (Savage-)acts, which are functions $g : \Omega \to X$ mapping from states to consequences. For each state $\omega \in \Omega$, the value $g(\omega)$ denotes the outcome the decision-maker obtains if ω materializes under the assumption that g was selected. Translated into the example of the ice-cream seller, Savage's concept of acts would imply that the seller goes for one of the functions g and h, where $g(\omega_1) = h(\omega_2) = 100$ and $g(\omega_2) = h(\omega_1) = 150$, instead of simply selecting between going to the "mall" or to the historical "marketplace".

Even today, Savage's subjective expected utility model is still considered as a benchmark model of decision making under uncertainty. Its sound axiomatic foundation and intuitive calculus made it popular for applications in all areas of economics. In the light of Savage's result, Knight's problem seems solved since the seller can determine her optimal action according to the SEU calculus. The problem with this approach is that the ice-cream seller might not conform to the axioms postulated in Savage's theory. Ellsberg [1961] shows with the famous Ellsberg paradox, that a subjective belief might not always exist, since decision-makers might display a preference for bets with purely objective probabilities. In the following, I illustrate the Ellsberg paradox within the scope of the example of the ice-cream seller. Assume that, after a few years, the ice-cream seller has learned to further differentiate how the weather conditions affect her payoffs. She knows that in the event of rain, she obtains 150 € if she goes to the mall, and 100 € if she sells at the historical market. In the event of a partly cloudy or cloudless skies, she obtains 100 € if she sells at the mall, and 180 € if she sells outside. In the event of a cloudy sky without rain, she obtains 120 € in the mall and in the historical market. Furthermore, she agreed with two different intermediaries on a weather-dependent contract providing her with the option to bring the ice-cream to one of them in the morning. Intermediary 1 pays her 100 € for the ice-cream in the event of sunny or cloudy weather, and 150 € in the event of rain. Intermediary 2 pays her 150 € in the event of sunny weather, and 100 € else. The new, extended decision problem of the ice-cream seller is summarized by means of Table 1.2.

action	**sunny**	**cloudy**	**rainy**
mall	100 €	120 €	150 €
marketplace	150 €	120 €	100 €
intermediary 1	100 €	100 €	150 €
intermediary 2	150 €	100 €	100 €

Table 1.2: Decision Problem for the Illustration of Ellsberg's Paradox

Assume that the seller listens to the local weather forecast and obtains the information that the probability of rain is 30% today, but she does not get any conclusive information on the probability of cloudy or sunny skies. I denote with p_{sunny}, p_{cloudy} and p_{rainy} the respective subjective probabilities. Assume furthermore that the seller prefers intermediary 1 towards intermediary 2 and that, if she had to select between the marketplace and the mall, she would prefer to sell at the marketplace. If the seller exhibits such a preference, she displays behavior that is inconsistent with Savage's SEU-theory. This can be demonstrated by means of the following calculations. Preferring intermediary 1 to intermediary 2 implies

$$p_{sunny}u(100€) + p_{cloudy}u(100€) + p_{rainy}u(150€)$$
$$>p_{sunny}u(150€) + p_{cloudy}u(100€) + p_{rainy}u(100€)$$

Simplifying this inequality yields the statement

$$p_{sunny} < p_{rainy}$$

Similarly, preference for the marketplace implies

$$p_{sunny}u(180€) + p_{cloudy}u(120€) + p_{rainy}u(100€)$$
$$>p_{sunny}u(100€) + p_{cloudy}u(120€) + p_{rainy}u(150€),$$

which is equivalent to the statement

$$p_{sunny} > p_{rainy}$$

This contradicts the existence of a subjective probability for sunny weather.

The Ellsberg paradox has paved the way for a multitude of new theories of decision making under uncertainty. Schmeidler [1989] provides an axiomatic foundation of the Choquet-expected utility theory (henceforth CEU-theory). Instead of probabilities that satisfy the properties of σ-additivity and finite additivity, the Choquet model implies that a decision-maker's belief is non-additive and represented by a more general mathematical structure called "capacity" or "charge". In the following, I give a definition of the term capacity (charge).

Definition 1.1 (Cf. Chateauneuf et al. [2007], p. 540)**.** Consider the measurable space (Ω, Σ).[3] A capacity (charge) is a set function $\nu : \Sigma \to \Omega$ mapping events from the algebra Σ to real numbers with two additional properties. The function is normalized, implying

[3]A measurable space is a pair (Ω, Σ) where Ω is the sample space or state space, and Σ is a σ-algebra of events defined on Ω.

$\nu(\emptyset) = 0$, $\nu(\Omega) = 1$ and monotonic with respect to the set inclusion relation $\subset$, implying $A \subset B \Leftrightarrow \nu(A) \leq \nu(B)$.

The definition of a capacity is rather abstract. Nevertheless, there are certain classes of capacities which entail insightful behavioral interpretations for economic models. In order to illustrate the term capacity and its properties, I relate this notion to the ice-cream seller. To begin with, I denote with R the event of rainfall, S denotes the event of sunny weather, and C denotes the event of cloudy skies. Furthermore, I denote with RS the event that rainfall or sunny skies occur, RC denotes the event of rainfall or cloudy skies, and SC denotes the event of sunny skies or cloudy skies. I define a capacity by $\nu(\emptyset) = 0$, $\nu(\Omega) = 1$, and $\nu(A) = \alpha$ for all other events where $\alpha \in [0, 1]$. This is an example for a Hurzwicz-capacity.[4]. The aforementioned capacity does not fulfill the property of finite additivity; the probability, as measured by the capacity ν, of the union of two disjoint events is not equal to the sum of the probabilities of the single events. For example R and S are disjoint events, formally $R \cup S = \emptyset$, but $\nu(RS) = \alpha \neq \nu(R) + \nu(S)$. Another example for a capacity is the so-called neo-additive or non-extreme outcome capacity that I define in the following.

Definition 1.2 (Cf. Eichberger et al. [2009], p. 359)**.** Let q be a probability measure on Ω. Then, for real numbers α and δ, one can define a neo-additive capacity ν by $\nu(\varnothing) = 0$, $\nu(\Omega) = 1$, $\nu(A) = \delta\alpha + (1 - \delta)q(A)$ where $A \in \Sigma \backslash \{\emptyset, \Omega\}$ is a nonempty and strict subset from the σ-algebra Σ.

In order to illustrate the notion of a neo-additive capacity, I relate this definition to the ice-cream seller. Assume for example that the ice-cream seller listens every day to the same weather forecast from the same provider. Due to the large amount of past observations, she considers this weather forecast as a very reliable source of information, and she is confident that the probabilities of rainy, cloudy, or sunny weather reaped from this forecast are very accurate. Suppose now that there is an alternative weather forecast from

[4]For the general definition of a Hurwicz-capacity see e.g. Chateauneuf et al. [2007], page 541.

a new provider available. The ice-cream seller disposes of no evidence proving the reliability of this forecast. One day, she listens to this alternative weather forecast because she missed out on the weather prediction she usually resorts to. Assume furthermore that the alternative weather prediction is detailed enough to provide her with probabilities of sunny, cloudy, and rainy weather. A natural question in this context is whether the ice-cream seller would trust the alternative weather forecast in the same way as her usual weather forecast. Neo-additive capacities capture the ice-cream seller's reliability concerns with the confidence parameter δ. If δ is equal to zero, the ice-cream seller fully trusts the probability distribution arising from the weather forecast. If δ is equal to zero, the seller dismisses this underlying probability completely. For intermediate values of δ the probability of an event $A \in \Sigma \backslash \{\emptyset, \Omega\}$ is given by a convex combination of the fixed probability α and $q(A)$. Suppose that both weather forecasts post the same probabilities for rainfall, sunny, and cloudy weather inducing the same "objective" probability distribution q, then the seller's concern with respect to the reliability of the alternative weather forecast may be expressed by assuming that δ_1 is smaller than δ_2 where δ_i for $i = 1, 2$ corresponds to the confidence parameter δ in the event of the usual weather forecast ($i = 1$) and the alternative forecast ($i = 2$). The parameter α captures the ice-cream seller's attitude towards ambiguity. Assume for example that the seller dismisses completely the alternative weather forecast. In this case, her belief is, by construction, represented by a Hurzwicz-capacity, namely the same Hurwicz-capacity which has been introduced as an illustration of Definition 1.1. The Hurwicz-capacity assigns to all events $A \in \Sigma \backslash \{\emptyset, \Omega\}$ the same probability α. Thus, the ice-cream seller would assign the same probability α to the event rain R, to the event sunshine S, and to the event rain or sunshine RS. The smaller α, the more likely it is that the ice-cream seller underestimates the probability of very likely events and overestimates the likelihood of events with very small probabilities. A general behavior like that might be labeled "pessimistic", since the ice-cream seller worries more about unlikely, small events and considers likely events not as likely as they might seem. If, on the other hand, the ice-cream seller assigned a probability α close to one to all events $A \in \Sigma \backslash \{\emptyset, \Omega\}$, this would imply that the seller consistently underestimates small events and overestimates a larger number of very likely events. With a similar reasoning, one can argue that such a behavior might be labeled "optimistic".

The Choquet-model presumes a different notion of acts than the SEU-model. In the same way as in the framework developed by Anscombe and Aumann [1963], an act is defined as a mapping $f : \Omega \to \mathcal{L}$ from the state space Ω to the set $\mathcal{L}$ of finite lotteries on X. Henceforth, the set of possible acts is defined as

$$\mathcal{F} = \{f | f : \Omega \to \mathcal{L}\}.$$

How does this novel definition of acts fit into the example of the ice-cream seller? Again, as in the Savage case, we can identify each of the seller's actions with a specific act. Consider, for example, the action "going to the mall". For each weather condition, or state of the world, Table 1.2 indicates the seller's payoffs. The previous definition of an act assumes that the entries of Table 1.2 specify lotteries. Instead of receiving a fixed outcome, the seller obtains for each action a lottery. An act is thus a complete plan, specifying for each state of the world, which lottery the seller is going to play. In the end, nature plays out this lottery and determines the seller's payoff. Of course, this is not an explanation yet of how we can fit the ice-cream seller's decision problem into this framework. This task can be accomplished by identifying each fixed monetary outcome with a lottery that pays out exactly the same amount of money with probability one. Now, everything is in place to specify an act f associated with the action "going to the mall"; f maps the state "sunny weather" to a lottery that pays out 100 € with probability one, the state "cloudy weather" to a lottery that pays out 120 € with probability one, and the state "rainy weather" to a lottery that pays out 150 € with probability one.

In the expected utility model, a decision-maker's utility over the lottery space $\mathcal{L}$ is represented by an expectation[5]. In the Choquet model, this expectation is replaced by a "generalized expectation" the Choquet integral, which is based on a broader notion of

[5] Cf. equation (1.1).

integration than the "usual" expectation. More specifically, a Choquet integral allows for integration with respect to non-additive probabilities.

Definition 1.3 (Cf. Denneberg [1994], p. 62)**.** Let (Ω, Σ) be a measurable space, $\nu : \Sigma \to \mathbb{R}_+$ a monotonic set function and $h : \Omega \to \mathbb{R}$ a Σ-measurable function.[6] Then the Choquet integral of h with respect ν is defined as

$$\int_\Omega h d\nu := \int_{-\infty}^{0} \nu(\{\omega | h(\omega) > x\}) - \nu(\Omega) dx + \int_{0}^{\infty} \nu(\{\omega | h(\omega) > x\}) dx \tag{1.2}$$

where the integrals on the right-hand side of (1.2) are improper Riemann integrals.

Remark 1.1. If ν is a capacity, it satisfies the normalization $\nu(\Omega) = 1$.

Remark 1.2 (Cf. Denneberg [1994], page 62 et seq.)**.** In many applications and decision problems in economics, one deals with situations where the function h takes only finitely many values on a partition of the state space Ω. Such functions are called step functions. Let h be a step function where h takes the values $d_1 > d_2 > ... > d_n$, $d_{n+1} := 0$ and let $(A_i)_{i=1}^n$ be a partition of Ω with $A_i \cap A_j = \emptyset$ for all $i \neq j$, $\cup_{i=1}^n A_i = \Omega$ and $h(\omega) = d_i$ if $\omega \in A_i$. The Choquet integral of a step function is given by

$$\int_\Omega h d\nu = \sum_{i=1}^{n} (d_i - d_{i+1}) \nu \left(\bigcup_{j=1}^{i} A_j \right).$$

Remark 1.3. For more details on integration with respect to non-additive measures see Denneberg [1994].

In the following, I condense the notation slightly by assuming that for an act f the expression $f(\omega)[x_i]$ denotes the probability of outcome x_i given the lottery $f(\omega)$. Schmeidler [1989] proposes a set of axioms on $\mathcal{F}$ that entail the following representation of CEU-preferences:

[6] Measurability implies that the event $\{\omega | h(\omega) > x\}$ is contained in the σ-algebra Σ for all $x \in \mathbb{R}$.

$$f \succsim g \quad \text{iff} \quad \int_\Omega \sum_{x_i} f(\omega)[x_i]u(x_i)d\nu \geq \int_\Omega \sum_{x_i} g(\omega)[x_i]u(x_i)d\nu \tag{1.3}$$

This representation involves double integration. The integrands are expected utilities with respect to the lotteries $f(\omega)$ and $g(\omega)$ respectively. Thus, for each state, the decision-maker is confronted with different lotteries and forms expected utilities given these lotteries. The probability of a state $\omega \in \Omega$, however, is condensed in the capacity ν. What the CEU decision-maker does, is forming an expectation or "average", based on non-additive probabilities with respect to all expected utilities that might occur for different realizations of a given act. In the end, the agent ranks all acts according to the ranking induced by these "averaged expectations".

The Choquet integral formed with respect to neo-additive capacities has a specific and intuitive representation. In the following, I introduce the concept of null events and simple functions, which are technical prerequisites for the aforementioned representation.

Definition 1.4 (Cf. Chateauneuf et al. [2007], p. 540 et seq.)**.** A set $A \in \Sigma$ is called null or a null-event with respect to the capacity ν if $\nu(A) = 0$.

Thus, a null-event is an event which carries zero probability where the term "probability" is generalized to capacities. For a given capacity, the set of null-events is henceforth denoted with $\mathcal{N}$. In the ice-cream seller example, the only null-event is the empty set $\emptyset$.

Definition 1.5 (Cf. Chateauneuf et al. [2007], p. 540-541)**.** A function $f : \Omega \to \mathbb{R}$ is called simple if it is Σ-measurable and finitely valued.

In cases where the underlying capacity is neo-additive, the Choquet integral has the following representation:

Lemma 1.6 (Lemma 3.1 from Chateauneuf et al. [2007], page 541)**.** *The Choquet-expected value of a simple function f with respect to a neo-additive capacity ν is given by*

$$\int_\Omega f d\nu := \delta\Big(\alpha \max\{x : f^{-1}(x) \notin \mathcal{N}\} + (1-\alpha)\min\{y : f^{-1}(y) \notin \mathcal{N}\}\Big)$$
$$+ (1-\delta)E_\pi[f]$$

Proof. Cf. Chateauneuf et al. [2007], page 542. □

Assume that the ice-cream seller's belief is represented by a neo-additive capacity. How does she determine her optimal action? As pointed out, she selects among four different acts $f_1, f_2, f_3, f_4 \in \mathcal{F}$, each act arising from one of her four actions available.

act	**action**
f_1	sell at the mall
f_2	sell at the marketplace
f_3	instruct intermediary 1
f_4	instruct intermediary 2

For each act, she contemplates the Choquet integral

$$\int_{\Omega} f_i d\nu$$

and selects the action that entails the highest Choquet expected value. The parameter δ captures the seller's confidence into the prior π arising from the weather forecast. If $\delta = 0$, the seller gives full weight to the expectation with respect to this π. In this case, the seller fully trusts the posted probabilities and acts as an expected utility maximizer. If $\delta = 0$, she gives full weight to a convex combination of extreme outcomes that occur with non-zero probability. What are these extreme outcomes? Since each act maps to lotteries that select an outcome with probability one, she considers for each act the best and worst-case monetary outcome associated with this act. In cases where the seller selects "selling at the mall", the best-case outcome is 150€, and the worst-case outcome is 100€; the seller considers the value $\alpha 150€ + (1 - \alpha)100€$. In cases where α equals zero, the seller gives full weight to the worst case and no weight to the best case. In cases where α equals one, the agent gives full weight to the best case and no weight to the worst case. For intermediate values of δ the seller assigns an overall weight of δ to the convex combination of extreme outcomes and of $1 - \delta$ to the expectation with respect to

π.

The MMEU model and the Choquet model are interrelated. More specifically, there is a certain class of capacities, called convex capacities, that allow for a representation of any Choquet expected utility as a MMEU with a specific set of priors, the so-called core of the capacity. In the following, I define the terms convexity and core.

Definition 1.7. A capacity ν is called convex, if for all events $A, B \subset \Omega$ the following inequality holds:

$$\nu(A) + \nu(B) \leq \nu(A \cup B) + \nu(A \cap B)$$

In the literature, cf. Schmeidler [1989], a convex capacity is frequently associated with agents that display uncertainty averse behavior.[7] A decision-maker whose belief is represented by a convex capacity underestimates the occurrence of smaller, single events.

Definition 1.8. Let ν be a capacity and Ω a finite state space. Then the core of the capacity is defined as

$$core(\nu) = \Big\{ q : q(A) \geq \nu(A),\ q \text{ probability},\ q(\Omega) = \nu(\Omega), A \in \Sigma \Big\}.$$

Schmeidler [1986] shows that for a convex capacity with nonempty core and any real-valued Bernoulli utility $u : X \to \mathbb{R}$ one has

$$\int_\Omega u \, d\nu = \min_{p \in core(\nu)} \int_\Omega u \, dp$$

Another prominent model of decision making under uncertainty is the Multiple prior model or MaxMin-expected utility model (henceforth MMEU-model) axiomatized by Gilboa and Schmeidler [1989]. Similar to the Choquet model, acts are defined as mappings $f : \Omega \to \mathcal{L}$, and preferences are defined on the set of acts $\mathcal{F}$. Gilboa and Schmeidler [1989] provide the subsequent representation for MMEU preferences:

[7] This view is criticized by Epstein [1999].

$$f \succsim g \quad \text{iff} \quad \min_{p \in Q} \int_{\Omega} \sum_{x_i} f(\omega)[x_i]u(x_i)dp \geq \min_{p \in Q} \int_{\Omega} \sum_{x_i} g(\omega)[x_i]u(x_i)dp \tag{1.4}$$

Q is a nonempty and closed set of priors defined on the state space Ω, and $u : X \to \mathbb{R}$ is a utility function.[8] Again, as in the case of the Choquet model, the decision-maker compares averages of expected utilities. One major difference to the Choquet model is that the averaging occurs with respect to additive probabilities. For each additive prior in the set Q, the agent forms an averaged expected utility. When evaluating an act, the decision-maker considers the lowest of all these expected utilities. In short, the agent compares worst-case averaged expected utilities. A decision-maker displaying such behavior is highly pessimistic. Among all probabilistic scenarios, the latter bases his comparison of acts solely on the worst case. A model making use of multiple priors and allowing for optimistic and pessimistic attitudes towards ambiguity is the so-called α-MEU model developed by Ghirardato et al. [2004]. In their framework, a decision-maker with a prior set Q considers for a given act $f \in \mathcal{F}$ a convex combination of best- and worst-case expected utilities.

$$\alpha \min_{p \in Q} \int_{\Omega} \sum_{x_i} f(\omega)[x_i]u(x_i)dp + (1-\alpha) \max_{p \in Q} \int_{\Omega} \sum_{x_i} f(\omega)[x_i]u(x_i)dp \tag{1.5}$$

Until now, the α-MEU model lacks a complete axiomatization. Ghirardato et al. [2004] argue that the α-MEU model provides a clear separation between ambiguity and a decision-maker's attitude towards ambiguity. The α-MEU model is criticized by Eichberger et al. [2011]. The authors show that a decision-maker that conforms to the axioms suggested in Ghirardato et al. [2004] can only display extreme attitudes towards ambiguity characterized by extreme optimism $\alpha = 0$, or extreme pessimism $\alpha = 1$.

A generalization of the MMEU model is given by the variational representation of preferences axiomatized by Maccheroni et al. [2006]. Given their set of axioms, Maccheroni

[8] Bewley [2002] proposes a model of decision making under uncertainty which allows for a similar representation as the MMEU model. In contrast to Gilboa and Schmeidler [1989], Bewley's model is based on the assumption of incomplete preferences.

et al. [2006] obtain the subsequent representation of preferences on $\mathcal{F}$:

$$f \succsim g \Leftrightarrow \min_{p \in Q} \left\{ \sum_{x_i} f(\omega)[x_i]u(x_i) + c(p) \right\} \geq \min_{p \in Q} \left\{ \sum_{x_i} g(\omega)[x_i]u(x_i) + c(p) \right\} \tag{1.6}$$

The function $c : \Delta \to [0, \infty]$ is a grounded[9], convex, and lower semicontinuous function. If the function c equals zero, the model reduces to the MMEU model. The rationale speaking for the introduction of variational preferences is again the objective to separate ambiguity and a decision-maker's attitude towards ambiguity.[10] A special case of the variational model are the so-called multiplier preferences introduced by Hansen and Sargent [2001]. In their model, the authors specify c as a multiple of the so-called relative entropy, which is a mathematical concept established by Kullback and Leibler [1951]. Given two probability distributions p and p^*, the relative entropy $R(p, p^*)$, or Kullback-Leibler divergence, constitutes a distance measure between p and p^* where q is usually referred to as reference probability.

Definition 1.9. If p and p^* can be represented by densities f and f^*, then the Kullback-Leibler divergence is defined as

$$R(p, p^*) = \int_{-\infty}^{\infty} \log \frac{f(x)}{f^*(x)} dx.$$

Multiplier preferences are based on the following functional form

$$\mathcal{U}^{MUPR}(f) = \min_{p \in Q} \left\{ \sum_{x_i} f(\omega)[x_i]u(x_i) + \gamma R(p, p^*) \right\} \tag{1.7}$$

where $f \in \mathcal{F}$ is a simple act, and $\gamma \geq 0$ is a parameter measuring the impact of the relative entropy. If γ equals zero, the objective reduces to the objective of the MMEU model. For large values of γ, the agent gives large weight to the Kullback-Leibler divergence. In this case, the agent bases her decision on a prior in the prior set Q that (nearly) minimizes the distance to the reference probability p^*. An axiomatic foundation of multiplier preferences

[9] Grounded means that the infimum value of c is zero, cf. Maccheroni et al. [2006], p. 1456.

[10] Maccheroni et al. [2006] state that attitude towards ambiguity is condensed in the function c.

is provided by Strzalecki [2011].

Another model of decision making under uncertainty is the Smooth ambiguity or KMM model introduced by Klibanoff et al. [2005]. Again, as in previous models of decision making under uncertainty, there is a state space Ω and a σ-algebra $\Sigma(\Omega)$ of possible events. Furthermore, the KMM model operates with second-order probabilities. In this respect, I denote with Δ the set of all probability distributions on Ω and $\Sigma(\Delta)$ is a σ-algebra of events defined on Δ. An event $A \in \Sigma(\Delta)$ consists thus of a set of probability distributions, each probability defined on the state space Ω. Let $\mu : \Sigma(\Delta) \to [0, 1]$ be a second-order probability distribution on Δ. Then, the value $\mu(A)$ for an event $A \in \Sigma(\Delta)$ denotes the probability that one of the probability distributions in A is the true underlying probability. Given the set of axioms in Klibanoff et al. [2005] a decision-maker's utility $\mathcal{U}$ on the set of simple acts $\mathcal{F}$ takes the subsequent functional form:

$$\mathcal{U}^{KMM}(f) := \int_{\Delta} \Phi \left(\int_{\Omega} \sum_{x_i} f(\omega)[x_i]u(x_i)dp \right) d\mu(p). \tag{1.8}$$

The function $u : X \to \mathbb{R}$ is a Bernoulli utility, and $\Phi : u(X) \to \mathbb{R}$ is a strictly increasing function. The KMM functional can be interpreted in the following way: the inner integral corresponds to an expected utility with respect to the additive subjective belief p. Similar to previously introduced models of decision making under uncertainty, the decision-maker forms an average of multiple expected utilities. In the context of the KMM model, this averaging occurs with respect to the second-order probability μ. In the general case, the function Φ distorts the inner expected utilities, unless Φ equals the identity. In this special case, the model reduces to the model of Anscombe and Aumann [1963]. A special feature of the KMM model is that the reduction of compound lotteries to simple lotteries is in general not possible.

The KMM model is criticized by Epstein [2010]. The author argues by means of thought experiments that the KMM's axiomatic foundation is problematic from a normative point of view and that separation between ambiguity and ambiguity attitude cannot be achieved by KMM. Furthermore, Epstein [2010] states that the benefits of the KMM model as compared to the MMEU model remain unclear.

To conclude, I want to mention briefly that there are more models of decision making under uncertainty, which I am not going to discuss hereafter, since they are not essential for the understanding of this thesis. Among these models is the vector expected utility model axiomatized by Siniscalchi [2009], a model using so-called source functions developed by Abdellaoui et al. [2011], the issue-preference model introduced by Ergin and Gul [2009], the confidence function model developed by Chateauneuf and Faro [2009], the model of uncertainty averse preferences introduced by Cerreia-Vioglio et al. [2011], and the expected uncertain utility model axiomatized by Gul and Pesendorfer [2014]. For a discussion of all of these models, except the model by Abdellaoui et al. [2011], I recommend the survey article by Machina and Siniscalchi [2014].

1.2 Dynamic Models of Decision Making Under Uncertainty

By now, there is a variety of articles extending the static models of decision making under uncertainty treated in the previous section to intertemporal settings. A major challenge in this context is the problem that these extensions might induce decision-makers to violate dynamic consistency. In the literature, one can find, depending on its context, different notions and definitions of dynamic consistency. Roughly speaking, the idea of dynamic consistency is tied to the view that any intertemporal decision problem can be represented by an event or decision tree. Ex-ante the decision-maker determines an optimal plan indicating which action he or she takes once arriving at a specific node of the tree. Now, dynamic consistency requires the decision-maker to stick to the original plan once a specific node at the tree is reached.[11] This means that, even though the decision-maker obtains the information that she arrived at a certain node she will not

[11] Cf. Machina [1989], page 1636 et seq.

revise her plan, she still finds her ex-ante plan optimal. Ghiradato [2002] provides a formal definition of dynamic consistency in a Savage framework. In order to further clarify the definition in Ghiradato [2002], I introduce the subsequent notation.

Notation 1.1. For two acts $f, g \in \mathcal{F}$ and an event $A \in \Sigma$ the act $f^A g$ denotes the act that equals g on the complement A^c of A and equals f on the event A. The preference relation $\succsim_A$ denotes the so-called conditional preference relation. It refers to the decision-maker's preference after he or she obtained the information that the event A has occurred.

Furthermore, Ghiradato [2002] introduces Savage's definition of null events, cf. Savage [1954].

Definition 1.10 (Null event)**.** An event A is called Savage-null with respect to $\succsim$ if and only if for all acts $f, g, h, h' \in \mathcal{F}$

$$f^{A^c}h \succsim g^{A^c}h' \quad \text{if and only if} \quad f \succsim g$$

Definition 1.10 says that no matter how we change the acts f and g on the event A, the original preference relation remains unaffected. In the following, I denote with Σ' the set of all Savage non-null events.

Definition 1.11 (Dynamic consistency, cf. Ghiradato [2002], page 87)**.** For all events $A \in \Sigma'$ and acts $f, g \in \mathcal{F}$, both the following conditions hold:

1. If $f \succsim_A g$ then $f^A g \succsim g$
2. If $f^A g \succsim g$ then $f \succsim_A g$

The first condition says that, if the decision-maker knows that A has occurred and he prefers f over g given this information, then he should also prefer the act $f^A g$, which is just a modification of f on the complement A^c, from an ex-ante perspective. The second condition implies the reverse logical direction. If the decision-maker prefers the modified act $f^A g$ from an ex-ante perspective, then he should prefer f over g when he knows that A has occurred.

Dynamic consistency is a prerequisite for the applicability of the backward induction principle. In the following, I give an example for an event tree and illustrate dynamic consistency as well as the backward induction principle under risk. Assume that there are three time points $t = 0, 1, 2$ and a decision-maker and that has two actions available A and B at time $t = 0$. Furthermore, there are two states of the world ω_1 and ω_2 that materialize at time $t = 1$. The agent knows the probability p_i of each state ω_i occurring. At time $t = 2$ the agent selects among two different actions C and D and obtains her payoffs immediately afterwards. The following event tree illustrates this decision problem where $p_i \in \mathbb{R}$ denote the decision-maker's payoffs.

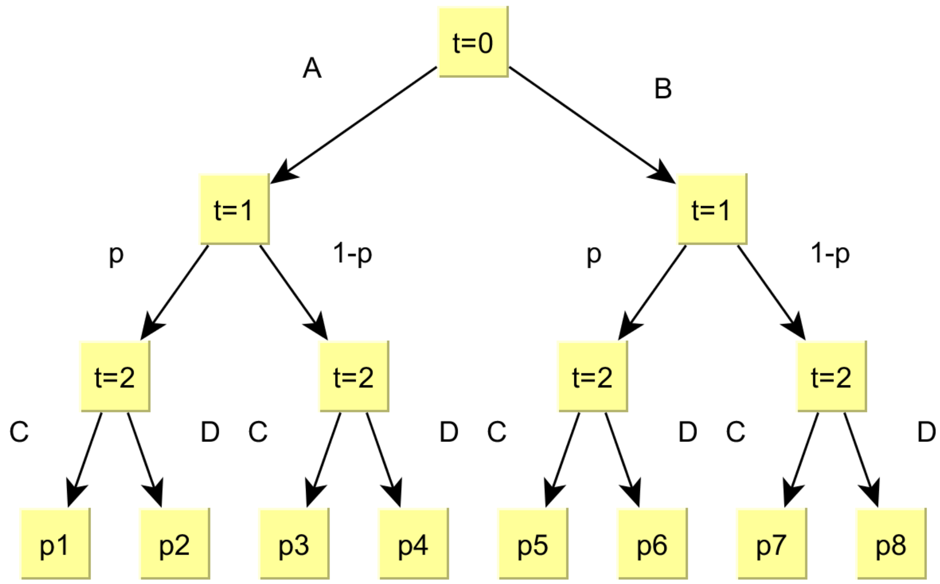


FIGURE 1.1: Decision Tree Under Risk

A complete plan for the decision-maker specifies what she does at each decision node, for instance (A, C) stands for selecting action A at time $t = 0$ and action C at time $t = 2$. The agent's optimal plan can be determined with the backward induction principle. Assume for the sake of simplicity that the decision-maker is a risk neutral expected utility maximizer, and thus that her utility function equals the identity. First, she determines the optimal action at time $t = 2$. Therefore, she compares the expected utilities induced by the plan (A, C) with the expected utility induced by the plan (A, D). Similarly, she compares the expected utility induced by the plan (B, C) with the expected utility induced by the plan (B, D). In each of these cases, she selects the action yielding the highest expected utility value. As a last step, she compares the highest expected utility values induced by action A and B and selects the plan that induces the highest overall expected utility value.

There is a variety of articles demonstrating the conflict of dynamic consistency with non-expected utility models. Machina [1989] provides examples under which decision-makers display dynamically inconsistent behavior when deviating from expected utility theory. Karni and Schmeidler [1991] find that dynamic consistency[12] and the independence axiom of the EU-theory by Von Neumann and Morgenstern [1944] are equivalent if decision-makers satisfy the reduction of compound lotteries axioms and consequentialism. Furthermore, the article by Epstein and Le Breton [1993] demonstrates that if preferences are based on beliefs[13] and beliefs are updated by Bayes' rule, the decision-maker holds a single Bayesian belief. This finding is problematic for non-expected utility models since it implies that one of the conditions postulated in Epstein and Le Breton [1993] need to be relaxed in order to justify beliefs implied by the models of decision making under uncertainty presented in the previous section. Ghiradato [2002] provides an axiomatization of a dynamic version of Savage's SEU model incorporating dynamic consistency and another prominent axiom of dynamic models of decision making under uncertainty, called consequentialism.

Definition 1.12 (Consequentialism, cf. Ghiradato [2002], page 88)**.** For any $A \in \mathcal{A}'$ and all $f, g \in \mathcal{F}$, $f(\omega) = g(\omega)$ for each $\omega \in A$ implies $f \sim_A g$.

Consequentialism requires that if two acts coincide on a non-null event A and the agent obtains the information that A has occurred then she must be indifferent between both

[12] The terms dynamic consistency and consequentialism are defined in the space of compound lotteries.

[13] The preference relation $\succsim$ is based on beliefs if and only if there exists a relation $\succsim_l$ defined on the algebra Σ of events such that for all events $A, B \in \Sigma$ the DM prefers to bet on event A if and only if $A \succsim_l B$, cf. Epstein and Le Breton [1993], page 2.

acts.

In the following, I give a short summary on dynamic extensions of decision-theoretic models featuring ambiguity. Epstein and Schneider [2003] provide a dynamic extension of the MEU model using a recursive structure. The authors maintain dynamic consistency by restricting the set of priors to so-called rectangular prior sets and assuming that each prior is updated by Bayes' rule individually. The definition of rectangularity is thereby tied to the agent's underlying information structure, the so called filtration. In the following, I discuss the terms filtration and rectangularity.

Definition 1.13. A filtration $\mathcal{F} = \{\mathcal{F}_t\}$ is a sequence of sub-σ algebras of Ω where $\mathcal{F}_s \subset \mathcal{F}_l$ for $s \leq l$.

Remark 1.4. The σ-algebra $\mathcal{F}_t$ represents the agent's information on the decision problem at time t. In most relevant intertemporal decision problems it is assumed that the state space Ω is finite. In these cases, there exists a finite partition of the state space Ω and that this partition generates the σ-algebra $\mathcal{F}_t$.[14] In this case, one can identify each sub-σ-algebra $\mathcal{F}_t$ with its generating partition. Moreover, denote with $\mathcal{F}_t(\omega)$ the component of the generating partition that contains ω. At time t, the decision-maker cannot differentiate between states in the same component $\mathcal{F}_t(\omega)$.

In order to define the term rectangularity, I introduce the following notation from Epstein and Schneider [2003], page 7.

Notation 1.2. Consider the measurable space (Ω, Σ) and a probability p defined on that space. Consider the sub-σ algebra $\mathcal{F}_t$ of the filtration $\mathcal{F}$. By Remark 1.4 there exists a finite partition of the state space

$$\Omega = \sum_{i=1}^{n} A_i^t$$

[14]A collection $\mathcal{M}$ of subsets of Ω generates the σ-algebra Σ if Σ is the smallest σ-algebra that contains $\mathcal{M}$ where the term "small" refers to set inclusion $\subseteq$.

such that this partition generates the sub-σ algebra $\mathcal{F}_t$. For each $\omega \in \Omega$ there exists a unique event A^t_ω in the partition such that $\omega \in A^t_\omega$. In this context, I define

$$p_t(\cdot) := p(\cdot|A^t_\omega)$$

as the conditional probability with respect to the sub-σ-algebra $\mathcal{F}_t$ and

$$p_t^{+1} = p_t|_{\mathcal{F}_{t+1}}$$

is the restriction of p_t on the sub-σ algebra $\mathcal{F}_{t+1}$ called one-step-ahead conditional probability. The one-step-ahead conditional can be interpreted in the following way: given the information $\mathcal{F}_t(\omega)$ at time t, the decision-maker might be able to exclude certain realizations of the state space. More specifically, she knows that the observed realization ω lies in the event A^t_ω. At time $t+1$ the decision-maker knows that the realization lies in the event A^{t+1}_ω, which is a subset of A^t_ω. This is because, by definition, the filtration at time $t+1$ is finer than the filtration at time t. More specifically, the partition generating $\mathcal{F}_{t+1}$ is finer than the partition generating $\mathcal{F}_t$. Thus, at time t, it is possible to restrict the partition generating $\mathcal{F}_{t+1}$ to all events that are subsets of A^t_ω. Now, the one-step-ahead conditional probability assigns a probability to these events given the information available at time t. Figure 1.2 illustrates graphically the remarks of this paragraph. The state space is represented by the ellipse. Given the information structure of the problem, the decision-maker knows at time t that the true ω lies either in the red part or in the blue part of the ellipse. At time $t+1$ the decision-maker knows that the true ω lies in one of the events 1 to 7. Assume, for instance, that the decision-maker knows at time t that the true ω lies in the red part of the ellipse. In this case, the decision-maker can already exclude events 5 to 7. Given the information at time t the decision-maker forms a prior defined on the events that constitute the remaining part of the ellipse. What is for instance the probability that ω lies in event 1. It is the one-step-ahead probability

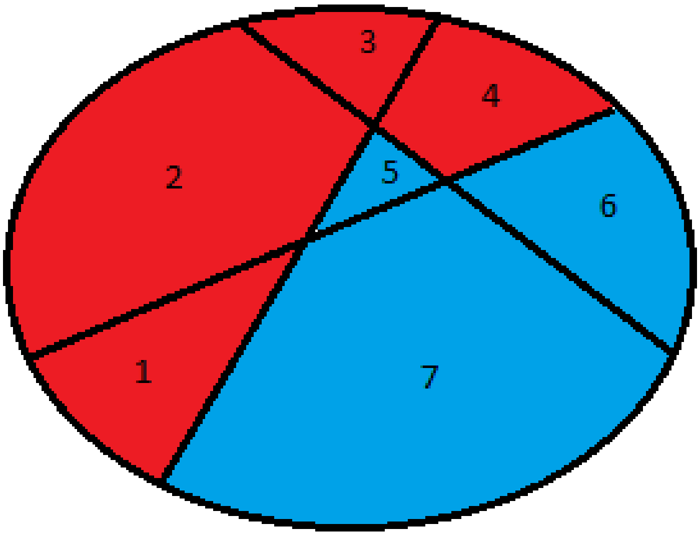


FIGURE 1.2: Example for Filtrations and One-Step-Ahead Conditionals

$p_t^{+1}(B)$ where B is the event $B = \{\omega \text{ lies in event } 1\}$. Assume now that the decision-maker's belief structure at time $t = 0$ is represented by a set of priors $\mathcal{P}$. In this case, the decision-maker successively updates each prior by Bayes' rule and obtains for each t the so-called set of Bayesian updates

$$\mathcal{P}_t(\omega) := \{p_t(\cdot) : p \in \mathcal{P}\}.$$

Similarly, the set of one-step-ahead conditionals at time t is defined as

$$\mathcal{P}_t^{+1}(\omega) := \left\{p_t^{+1}(\cdot) : p \in \mathcal{P}\right\}.$$

In the following, I define the term rectangularity of a prior set.

Definition 1.14 (Rectangularity, cf. Epstein and Schneider [2003], page 8)**.** The prior set $\mathcal{P}$ is called $\mathcal{F}_t$-rectangular if and only if

$$\mathcal{P}_t(\omega) = \left\{\int_\Omega p_{t+1}(\omega')dm : p_{t+1}(\omega') \in \mathcal{P}_{t+1}(\omega')\ \forall\, \omega', m \in \mathcal{P}_t^{+1}(\omega)\right\}$$

Subsequently, I give an intuitive interpretation of this definition. Consider a general probability measure p_t. Given the underlying filtration $\mathcal{F}$, we can deduce its Bayesian update p_{t+1} as well as its one-step-ahead conditional p_t^{+1}. Since Epstein and Schneider [2003] require dynamic consistency in their framework, agents with these preferences can determine optimal plans by using the previously explained backward induction principle. Working backwards, the decision-maker needs to reconstruct the probability p_t from its one-step-ahead conditional and its Bayesian update p_{t+1}. Formally,

$$p_t(\omega) = \int_\Omega p_{t+1} dp_t^{+1}(\omega) \tag{1.9}$$

In a single prior world, this reconstruction is always possible due to the law of iterated expectations, cf. Epstein and Schneider [2003], page 7. The decision-maker simply integrates the Bayesian-update with respect to the one-step-ahead conditional distribution to reconstruct the original distribution p_t. In a world of multiple priors, the aforementioned reconstruction process might fail, since the agent ignores which of the multiple one-step ahead conditionals and Bayesian updates she needs to combine to generate the specific distribution p_t. Therefore, when applying the backward induction principle, the agent might blend one-step-ahead conditionals with Bayesian updates that were not intended to be combined together from an ex-ante perspective. This inevitably leads to an enlarged prior set $\hat{\mathcal{P}}_t(\omega) \supset \mathcal{P}_t(\omega)$ when working backwards. Now, rectangularity makes sure that, given the filtration $\mathcal{F}$, the prior set $\mathcal{P}$ is "rich enough", such that the updated prior set $\mathcal{P}_t(\cdot)$ corresponds to the prior set $\hat{\mathcal{P}}_t(\cdot)$, which one would obtain when applying the backward induction principle.

By now, there are frameworks that go beyond the recursive multiple prior model. Wang [2003] axiomatizes updating rules for dynamic models of decision making under uncertainty that are non-Bayesian. As a special case, the author obtains the so-called "generalized Bayes rule" which corresponds exactly to the updating rule employed in the multiple prior framework of Epstein and Schneider [2003]. Hayashi [2005] provides an axiomatization of an intertemporal model of decision making under uncertainty that yields the recursive multiple prior framework of Epstein and Schneider [2003] and the model developed by Kreps and Porteus [1978] as special cases. As a consequence, the author refers to this model as "generalized recursive multiple priors utility". In contrast to Epstein and Schneider [2003], the model introduced by Hayashi [2005] allows the authors to disentangle ambiguity, risk aversion, and intertemporal substitution.[15]

An intertemporal version of the Choquet model called iterated Choquet expected utility is axiomatized by Nishimura and Ozaki [2003]. The question of dynamic consistency in the Choquet model is addressed by Eichberger et al. [2005], who consider the class of convex capacities and informational framework with an underlying fixed filtration. In their framework, dynamic consistency can be implied if beliefs are additive over the final stage of the underlying filtration.

A dynamically consistent and consequentialist extension of the Smooth ambiguity model is axiomatized by Klibanoff et al. [2009].

A dynamic version of the vector expected utility model is treated by Siniscalchi [2011].

In the previous sections, I sketched briefly the historical development of decision theory and introduced a variety of models of decision making under uncertainty. For a more detailed summary on the development of decision theory and models of decision making under uncertainty see, for instance, the surveys by Machina and Siniscalchi [2014] or Etner et al. [2012].

[15]See the abstract of Hayashi [2005].

Chapter 2

Spatial Competition Under Ambiguity

2.1 Introduction

Product development is one of the most influential processes for the success of an enterprise, see for instance Brown and Eisenhardt [1995]. Firms compete by creating products with new or different characteristics, amongst others, in order to enter new markets, to retain current customers or to attract new purchasers.

A well-known and widely studied model of product differentiation is the location-then-price duopoly game introduced by Hotelling [1929].[1] In his original framework, Hotelling discussed a model with two firms and uniformly distributed consumers along a compact interval facing linear transportation costs. At the first stage of the game, firms choose simultaneously their locations on this interval. At the second stage, firms face price competition.

[1] The "location" in Hotelling's game is typically interpreted as a position in a geographical or product type space. In this paper, we focus in the following on the latter interpretation.

A vast literature deals with a multitude of extensions of Hotelling's model.[2] In particular, d'Aspremont et al. [1979] show that, under linear cost functions, the existence of a subgame-perfect Nash equilibrium (*henceforth SPNE*) is not guaranteed. As a resort to this complication, d'Aspremont et al. [1979] replaced Hotelling's original assumption of linear transportation costs by quadratic ones. In the literature, Hotelling models with quadratic cost functions are frequently denoted by "AGT-models".[3]
Since in most real-world situations, firms are confronted with uncertain consumer preferences, a part of the more recent literature analyzes the impact of demand uncertainty on equilibrium product differentiation. Balvers and Szerb [1996] consider a Hotelling framework incorporating random shocks on the quality of each firm's product under the assumption that there is no price competition. Harter [1996] considers a Hotelling model with demand location uncertainty where firms enter the market sequentially. Similar to Harter [1996], Casado-Izaga [2000][4], Meagher and Zauner [2004], and Meagher and Zauner [2005] discuss extensions of Hotelling's model where demand uncertainty is introduced by enabling the midpoint of the consumer interval to be probabilistic. Meagher and Zauner [2005] generalize Casado-Izaga [2000] by parametrizing the length of the support. They find that equilibrium differentiation increases in the size of the support. Meagher and Zauner [2004] restrict this support to compact subsets of the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$ but allow for a broad class of density functions. Again, Meagher and Zauner *(henceforth MZ)* come to the conclusion that uncertainty constitutes a differentiation force, namely an increase in the variance of the underlying probability distribution over the midpoint leads to more pronounced equilibrium product differentiation.
All the contributions mentioned above imply that firms' beliefs are represented by a unique and common prior. However, in reality, this assumption may be violated for several reasons. First of all, the assumption of a unique common probability distribution for both firms is more restrictive than it may seem to be at first glance, especially in

[2] See e.g. Gabszewicz and Thisse [1992] for a survey.
[3] AGT stands for D'Asprémont, Gabszewicz and Thisse
[4] Casado-Izaga [2000] assume that consumers are uniformly distributed on the interval $[\Theta, \Theta+1]$ where Θ is drawn from a uniform distribution $[0, 1]$. Consequently, the midpoint of the consumer interval follows implicitly a uniform distribution on $[\frac{1}{2}, \frac{3}{2}]$.

situations where both firms are ex-ante completely uninformed or incapable to rely on past experiences or observable data. Furthermore, critiques in favor of a unique common probability distribution may argue that it is possible to apply the "principle of insufficient reason"[5] in case of missing information. However, Ellsberg [1961] indicates in his famous mind experiment that in situations of "ambiguity" where probabilities are unknown, or imperfectly known, a considerable share of individuals displays preferences which are incompatible with probabilistic beliefs. By now, several decision theoretic models of ambiguity have been developed. Prominent examples are the multiple prior model of Gilboa and Schmeidler [1989], the Choquet expected utility (*henceforth CEU*) model of Schmeidler [1989], and the smooth ambiguity model of Klibanoff et al. [2005].
Although ambiguity is prevalent in many real-world situations, there are almost no Hotelling models incorporating this type of uncertainty. To our knowledge, Król's [2012] article is the sole contribution on this topic. Król [2012] introduces complete ambiguity[6] into the framework of Meagher and Zauner [2004] and examines, amongst others, the influence of ambiguity attitude on firms' decisions if firms use the Arrow/Hurwicz α-maxmin criterion[7]. Król [2012] finds that uncertainty can be an agglomeration force if firms are sufficiently pessimistic.

The present paper studies the impact of confidence and pessimism on product differentiation. Inspired by the contributions of MZ and Król [2012], we develop a Hotelling model with demand location uncertainty by using the framework of Meagher and Zauner [2004] and Schmeidler's concept of CEU. More specifically, we assume that firms' beliefs are represented by a neo-additive capacity introduced by Chateauneuf et al. [2007]. Our framework provides additional analytical tools for understanding product differentiation under demand uncertainty. Besides firms' ambiguity attitude, we distinguish four different sources of ambiguity and determine their influence on firms' product design choices:

[5] The "principle of insufficient reason" or "principle of indifference", enunciated in the works of Pierre-Simon Laplace, see e.g. Laplace [1812], states that if decision-makers have no information about the frequency of occurrence of elementary events, and therefore no reason to believe that one elementary event will occur preferentially compared to another, they might consider them as equally likely.

[6] Complete ambiguity or ignorance refers to a situation where probabilistic information is absent.

[7] See Hurwicz [1951] and Arrow and Hurwicz [1972].

(i) the variance of firms' prior beliefs, (ii) the degree of ambiguity, (iii) the size of the support of the uncertainty and (iv) the magnitude of the parameter of consumers' quadratic cost functions. In particular, (ii) offers plausible possible explanations for real-life phenomena. In fact, the models of Meagher and Zauner [2004][8] and Król [2012] are special cases of the capacity model.

Our paper is organized as follows: In the following section, we give a detailed description of our model. As a second step, we derive firms' pure strategy subgame-perfect product design choices for the Hotelling game under ambiguity. Thereby, we assume that firms' beliefs are represented by neo-additive capacities. In section 4, we carry out a comparative static analysis with respect to all model parameters and study implications for equilibrium product characteristics and Choquet expected profits. Section 5 presents implications for possible applications of the Hotelling model under demand location uncertainty. In particular, we reexamine the examples mentioned in Król [2012]. Finally, section 6 concludes with a summary and a discussion of our findings.

[8]With a technical restriction. For more details see section 3, especially Remarks 3.1 and 2.3.

2.2 Basic Framework

Our framework is inspired by the modified AGT-model of Meagher and Zauner [2004]. There are two firms, $i = 1, 2$, interacting in a two-stage Hotelling duopoly game. Each firm produces a homogeneous commodity at constant marginal production costs which are normalized to zero. At the first stage of the game, firms select simultaneously their product characteristics x_i from the real line under the assumption that $x_1 \leq x_2$. At the second stage, firms face price competition setting prices $p_i \in \mathbb{R}_+$ simultaneously as well. Furthermore, there is a unit mass of consumers, each consumer being uniquely characterized by a specific taste, $s \in \mathbb{R}$, representing her ideal commodity. Consumer tastes are assumed to be uniformly distributed on an interval of the form $[M - \frac{1}{2}, M + \frac{1}{2}]$ where $M \in \mathbb{R}$. A customer whose taste is located at s and consumes product i, faces a disutility from not consuming her ideal product. Consumers' utility losses depend on the squared distance between s and the selected product design x_i, formally $t(s-x_i)^2$ where $t \in \mathbb{R}_{++}$.[9] In addition, customers need to pay the price p_i of product i. As a consequence, total costs are given by $p_i + t(s-x_i)^2$. Moreover, we assume that customers purchase one unit of the homogeneous good from the firm that brings about the lowest total costs. Implicitly, this guarantees that consumers' outside option is non-binding, or in other words, that there is no reservation price.

In the certainty model M and t are a fixed and exogenously given parameters known to both firms throughout the game. In the risk model of Meagher and Zauner [2004] M is unknown to both firms whereas the scaling parameter t is normalized to 1. In the model of Król [2012] firms face ambiguity with respect to (t, M) resolving ambiguity with the Arrow/Hurwicz α-maxmin criterion. Similar to these models, we presume that the realization $(\hat{t}, \hat{M})$ of (t, M) is revealed to both firms before the price competition.

Assumption 1. Uncertainty is resolved at the second stage of the game.

[9] The parameter t allows for an up- or downward distortion of this quadratic disutility.

As postulated in Assumption 1, the realization $(\hat{t}, \hat{M})$ is revealed to both firms after the product design competition but before the price game. The rationale behind this assumption lies in the fact that most firms are able to adjust prices more easily than product designs, see e.g. Meagher and Zauner [2004]. If for instance actual sales volumes differ from estimated sales volumes, firm managers are usually in the position to readjust retail prices accordingly.

In addition, we assume that firms dispose of some probabilistic information condensed in a common prior q. We refer to q as "reference probability distribution" or "reference prior". Similar to the risk case, one needs to make several assumptions concerning the reference probability q which are summarized in Assumption 2.

Assumption 2. The reference prior q on (t, M) satisfies the subsequent requirements:

(R1) The variance of M exist: $\mathbb{E}_q|M^2| < \infty$.

(R2) The expectation of M is normalized to zero: $\mathbb{E}_q[M] = 0$.

(R3) The distribution of M has no atoms.

(R4) The support of M is given by the symmetric interval $[-L, L] \subseteq \left[-\frac{1}{2}, \frac{1}{2}\right]$.

(R5) The support of t is given by the interval $[\underline{t}, \overline{t}]$ where $\underline{t} \in (0, 1]$ and $\overline{t} \geq 1$.

(R6) The expectation of t is normalized to 1: $\mathbb{E}_q[t] = 1$.

(R7) The random variables t and M are uncorrelated.

$\mathbb{E}_q$ denotes the expectation formed with respect to the prior q.

At the first stage of the game, the random variable M enters quadratically into each firm's objective function.[10] This observation provides a justification why firms' product design choices solely depend on the first and second moment of M. On these grounds, Assumption (R1) guarantees the existence of best response functions. Moreover, taking

[10] See equation (2.1) and Lemma 2.5.

(R1) and (R4) together, we can formulate the following lemma which proves to be very useful for the mathematical considerations in the comparative statics section.

Lemma 2.1. The Requirements (R_1) and (R_4) imply

$$\mathbb{E}_q[M] \in [-L, L] \qquad \text{and} \qquad Var_q(M) \in \left[0, L^2\right]$$

Proof. The proof of this lemma is contained in the appendix.

The Requirements (R2) and (R6) are introduced for reasons of symmetry and tractability. Requirement (R3) is purely technical in nature and can be replaced in order to allow for discrete distributions or mixtures of continuous and discrete distributions. (R4) makes sure that the support of M is a compact subset of the interval $[-\frac{1}{2}, \frac{1}{2}]$ restricting the size of uncertainty to be relatively small. In addition, it assures that the extreme intervals for possible realizations of the consumer distribution $[-L - \frac{1}{2}, -L + \frac{1}{2}]$ and $[L - \frac{1}{2}, L + \frac{1}{2}]$ always have a non-empty intersection. This is a necessary assumption due to the following reason: If the size of uncertainty is large enough, one encounters three possible cases.[11]

(1) The firm located left becomes a monopolist.

(2) Both firms share the market.

(3) The firm located to the right becomes a monopolist.

When the size of uncertainty is small, only the second case applies. In this paper, we intend to restrict the analysis to the duopoly case. Furthermore, (R4) and (R5) imply that the support of q is a subset of $[\underline{t}, \overline{t}] \times [-L, L]$. Lastly, (R7) ensures that we can disentangle the effects of t and M.

[11]See Meagher and Zauner [2005] for a detailed investigation of these additional cases for the risk case.

2.3 Introducing Ambiguity into the Game

We introduce ambiguity by assuming that firms' beliefs are represented by non-additive probabilities or capacities. A capacity is a normalized and monotonic set function.

Definition 2.1 (Compare Schmeidler [1989], page 578)**.** Let Ω be a finite or infinite non-empty set of states of the world and let Σ be an algebra of events defined on it. A capacity is a real-valued function $\nu : \Sigma \to \mathbb{R}$ such that

(1) $\nu(\varnothing) = 0$ and $\nu(\Omega) = 1$ (normalization)

(2) $E, F \in \Sigma$ and $E \subseteq F$ implies $\nu(E) \leq \nu(F)$ (monotonicity).

A capacity can be seen as a generalized probability measure that does not necessarily retain σ-additivity. The expectation with respect to a non-additive measure is formed by using a Choquet integral[12]. In the present paper our analysis relies on a distinct class of capacities, named neo-additive capacities, axiomatized by Chateauneuf et al. [2007].

Definition 2.2 (Compare Eichberger et al. [2009], page 359)**.** Let q be a probability distribution on $\Omega = [\underline{t}, \overline{t}] \times [-L, L]$ satisfying Assumption 2. Then, for real numbers α and δ, a neo-additive capacity ν is defined by $\nu(\varnothing) = 0$, $\nu(\Omega) = 1$, $\nu(A) = \delta\alpha + (1-\delta)q(A)$ where $A \in \Sigma$ is a nonempty and strict subset of Ω.

From our point of view, neo-additive capacities display several nice features. The parameter δ can be interpreted as a measure of ambiguity or of firms' confidence in the common reference prior q. Thus, one can contemplate our model as a setting where firms exhibit uncertainty with respect to their prior beliefs due to imprecise or unreliable information. Moreover, the parameter α describes firms' attitude towards ambiguity. The higher α, the more pessimistic firm managers are. As a result, neo-additive capacities allow for a clear separation between the degree of ambiguity and firms' ambiguity attitude which is, as we

[12] See Choquet [1955].

want to argue in this paper, essential for many economic applications. Consequently, we assume that the neo-additive capacity represents firms' ex-ante uncertainty.

Assumption 3. Each firm's belief on (t, M) is represented by a neo-additive capacity ν.

The rationale speaking for the introduction of neo-additive capacities lies in in the fact that firms might not completely trust the information available at the time of making their product choice. There are multiple reasons why this might be the case, e.g. firms introducing newly innovated products into the market might dispose of data on similar products that are already established in the market but have no data on the new good. It seems plausible that firms use this data to predict the market outcome, still firms cannot account for short-term trends in consumer tastes. Furthermore, data reliability is closely tied to the comparability of the reference product with the newly innovated product. The more heterogeneous both products are, the less plausible it seems to rely on available data on the reference product. Neo-additive capacities allow for a model of partial information where firms have a certain stock of data available whose reliability might be questionable up to a certain degree. Interpreted in this way, the model developed by Król [2012] refers to a situation where firms have ex-ante no information about the distribution of consumer tastes or completely distrust information available at the time of making their product design choices. Moreover, neo-additive capacities allow for an additional interpretative component with respect to a multitude of possible real-world applications of Hotelling models under uncertainty by adding an additional explanatory source for increasing or decreasing product differentiation under ambiguity.

2.4 Solving the Game

In this section, we determine equilibrium product differentiation under ambiguity by backward induction. In a first step, we solve the price subgame at the second stage where the midpoint M of the consumer distribution and the cost parameter t are fixed and known to both firms.

2.5 Price Subgame

According to Assumption 1, the realization $(\hat{t}, \hat{M})$ is known to both firms at the second stage. Equilibrium prices are zero if firms do not differentiate their products. Otherwise, firms' equilibrium prices depend on the distance between firms' averaged product design $\bar{x} := \frac{x_1+x_2}{2}$ and the realized midpoint $\hat{M}$. There is an interior equilibrium where each firm charges a positive price:

Lemma 2.2. If $x_1 \leq x_2$ and $(\hat{M} - \bar{x}) \in [-\frac{3}{2}, \frac{3}{2}]$, firms charge the subsequent equilibrium prices:

$$p_1^* = \frac{2}{3}\hat{t}\Delta_x\left(\bar{x} - \hat{M} + \frac{3}{2}\right) \quad \text{and} \quad p_2^* = \frac{2}{3}\hat{t}\Delta_x\left(-\bar{x} + \hat{M} + \frac{3}{2}\right)$$

Proof. See Anderson et al. [1997], page 107 and Meagher and Zauner [2004], page 203.

Apart from the interior equilibrium, there are two more boundary equilibria where one of the two firms becomes a monopolist:

Lemma 2.3 (Boundary price equilibria)**.** If $x_1 \leq x_2$ and $(\hat{M} - \bar{x}) \notin [-\frac{3}{2}, \frac{3}{2}]$, firms charge the subsequent equilibrium prices:

$$p_1^* = 2\hat{t}\Delta_x\left(\bar{x} - \hat{M} - \tfrac{1}{2}\right) \quad \text{and} \quad p_2^* = 0 \quad \text{if} \quad (\hat{M} - \bar{x}) < -\tfrac{3}{2}$$

or

$$p_2^* = 2\hat{t}\Delta_x\left(\hat{M} - \bar{x} - \tfrac{1}{2}\right) \quad \text{and} \quad p_1^* = 0 \quad \text{if} \quad (\hat{M} - \bar{x}) > \tfrac{3}{2}.$$

Proof. See Anderson et al. [1997], page 107 and Meagher and Zauner [2004], page 203.

2.6 Product Design Competition

As shown in the previous section, one obtains for a fixed pair (x_1, x_2) of product characteristics a unique equilibrium for the price subgame. By making use of the equilibrium prices from Lemma 2.2 and 2.3, we obtain firms' second stage profits for the realization $(\hat{t}, \hat{M})$ depending on firms' product characteristics:

$$\Pi_i(x_i, x_j, \hat{t}, \hat{M}) = \begin{cases} \hat{t}\,\Delta_x \left[1 + 2\,(-1)^i(\hat{M} - \bar{x})\right] & \text{for} \quad (-1)^i\,(\hat{M} - \bar{x}) > \frac{3}{2} \\ \hat{t}\,\Delta_x \left[3(-1)^i + 2(\hat{M} - \bar{x})\right]^2 /18 & \text{for} \quad (\hat{M} - \bar{x}) \in [-\frac{3}{2}, \frac{3}{2}] \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

where $\bar{x} := \frac{x_1 + x_2}{2}$, $\Delta_x := x_2 - x_1$ and $j := 3 - i$.

In the following, we elaborate on each firm's objective function at the first stage of the game. In order to do so, we rely on the fact that the second piece of (2.1) is monotonic in $(\hat{t}, \hat{M})$ as specified in Lemma 2.4 below.

Lemma 2.4. If the condition $(\hat{M} - \bar{x}) \in [-\frac{3}{2}, \frac{3}{2}]$ is met, firm i's profit function $\Pi_i(x_1, x_2, \hat{t}, \hat{M})$ is strictly increasing in $\hat{t}$, strictly decreasing in $\hat{M}$ for firm 1, and strictly increasing for firm 2, provided that $x_1 < x_2$.

Proof. The proof of the lemma is contained in the appendix.

At the first stage of the game, the distribution of (t, M) is unknown. In accordance with Assumption 3 and Definition 2.2, firms consider the Choquet expected value of their first stage profits. We denote firms' objectives as $\mathbb{CEU}[\Pi_i(x_i, x_j, \hat{t}, \hat{M})]$. Note that the Choquet-expected value is formed with respect to a neo-additive capacity. Following Lemma 3.1 in Chateauneuf et al. [2007], page 541, we obtain the representation (2.2) of firm i's Choquet expected profit at the first stage of the game.

$$\begin{aligned}\mathbb{CEU}[\Pi_i(x_1,x_2,t,M)] = \int \Pi_i(x_i,x_j,\hat{t},\hat{M})d\nu &= (1-\delta)\mathbb{E}_q[\Pi_i(x_i,x_j,t,M)] \\ &+ \delta[(1-\alpha)\max\{\Pi_i(x_i,x_j,\hat{t},\hat{M}) : (\hat{t},\hat{M}) \in \mathrm{supp}(t,M)\} \\ &+ \alpha\min\{\Pi_i(x_i,x_j,\hat{t},\hat{M}) : (\hat{t},\hat{M}) \in \mathrm{supp}(t,M)\}]\end{aligned} \tag{2.2}$$

Remark 2.1. These Choquet expected profits allow for a nice interpretation, namely that they generalize Hotelling models treated in the literature before. For $\delta = 0$ and a constant scaling factor $t = 1$, we obtain the model of Meagher and Zauner [2004] with a normalized mean of M. In case of $\delta = 1$ and $\bar{t} = 1$, the framework boils down to the model of Król [2012]. Thus, we can consider these specifications as extreme cases of the capacity model.

As a next step, we consider the second part of equation (2.2). Making use of Lemma 2.4, we obtain for $(\hat{M} - \bar{x}) \in [-\frac{3}{2}, \frac{3}{2}]$ the following explicit functional relationships:

$$\begin{aligned}\max\{\Pi_1(x_i,x_j,\hat{t},\hat{M}) : (\hat{t},\hat{M}) \in \mathrm{supp}(t,M)\} &= \Pi_1(x_1,x_2,\bar{t},-L) \\ \min\{\Pi_1(x_i,x_j,\hat{t},\hat{M}) : (\hat{t},\hat{M}) \in \mathrm{supp}(t,M)\}] &= \Pi_1(x_1,x_2,\underline{t},L) \\ \max\{\Pi_2(x_i,x_j,\hat{t},\hat{M}) : (\hat{t},\hat{M}) \in \mathrm{supp}(t,M)\} &= \Pi_2(x_1,x_2,\bar{t},L) \\ \min\{\Pi_2(x_i,x_j,\hat{t},\hat{M}) : (\hat{t},\hat{M}) \in \mathrm{supp}(t,M)\}] &= \Pi_2(x_1,x_2,\underline{t},-L)\end{aligned} \tag{2.3}$$

Remark 2.2. One can interpret these results as follows. Firm 1's best-case scenario occurs when the realized midpoint $\hat{M}$ of the consumer interval equals the lower support boundary $-L$. This is true, since we assume, w.l.o.g., that firm 1 is the firm whose product characteristic is located left of firm 2's product characteristic. Therefore, it is more convenient for firm 1 if the consumer distribution is located closer to its own product design. Similarly, firm 1's worst-case scenario occurs when the midpoint of the consumer interval takes as realization the upper support boundary L. For firm 2 the reverse result holds.

The first term of firm i's Choquet expected profit equals the "usual" expectation of its profit function with respect to the reference prior $\mathbb{E}_q[\Pi_i(x_1,x_1,t,M)]$. In order to elaborate on this part, we need the following Lemma which can be considered as an analogue to the global competition lemma in Meagher and Zauner [2004].[13]

Lemma 2.5 (Global competition)**.** Under Assumptions 1,2, and 3, one has at any pure strategy SPNE for the Hotelling game with ambiguous demand location uncertainty that the support $[-L, L]$ of M is contained in $[\bar{x}-\frac{3}{2}, \bar{x}+\frac{3}{2}]$, formally $[-L, L] \subset \left[\bar{x}-\frac{3}{2}, \bar{x}+\frac{3}{2}\right]$.

Proof. The proof of the lemma is contained in the appendix.

Lemma 2.5 proves very useful when it comes to determining firms' subgame-perfect product design choices. In fact, due to Lemma 2.3, one could expect that there are equilibria where, for some realizations of uncertainty, one or the other firm can monopolize the market. However, according to Lemma 2.5, firm i's objective function at the first stage of the game is given by the Choquet expected value of the second piece of (2.1).
The global competition lemma implies that $\mathbb{E}_q[\Pi_i(x_i, x_j, t, M)]$ depends solely on the the mean vector $\mathbb{E}_q[(t, M)] = (\mu_t, \mu_M)$ and the variance-covariance matrix

$$Cov_q(t, M) = \begin{pmatrix} \sigma_t^2 & 0 \\ 0 & \sigma_M^2 \end{pmatrix}.$$

The following lemma provides an explicit mathematical form for $\mathbb{E}_q[\Pi_i(x_i, x_j, t, M)]$.

Lemma 2.6. If $x_1 \leq x_2$ w.l.o.g., then, under Assumptions 1,2, and 3, at any pure strategy SPNE for the Hotelling game under uncertainty, firms choose product characteristics,

[13] For the Hotelling model under certainty, Anderson et al. [1997] point out a similar property in footnote 8.

(x_1^*, x_2^*), such that firm i's expected profit w.r.t. the reference prior q is

$$\mathbb{E}_q[\Pi_i(x_i^*, x_j^*, t, M)] = \mu_t \int_{-L}^{L} (-1)^j \,\frac{2}{9}\,(x_j^* - x_i^*)\,\left(\bar{x}^* - \left(M + \frac{3}{2}(-1)^i\right)\right)^2 f_q(M)dM$$

$$= \frac{(-1)^j}{18}\,\mu_t\,(x_j^* - x_i^*)\Big\{(2\bar{x}^* - 3(-1)^i)^2 - 4\mu_M(2\bar{x}^* - 3(-1)^i) + 4(\mu_M + \sigma_M^2)\Big\} \tag{2.4}$$

where $\bar{x}^* = x_i^* + x_j^*$.

Proof. The proof of the lemma is contained in the appendix.

Next, after specifying firms' first-stage objective functions, we derive subgame-perfect product designs. Firm i's best reply, $R_i^*(\hat{x}_j)$, given the product characteristic choice of firm j, $\hat{x}_j$, is

$$R_i^*(\hat{x}_j) := \underset{x_i \in \mathbb{R}}{\arg\max}\Big\{(1-\delta)\mathbb{E}_q[\Pi_i(x_i, \hat{x}_j, t, M)] + \delta\big[(1-\alpha)\Pi_i(x_i, \hat{x}_j, \bar{t}, -L) + \alpha\Pi_i(x_i, \hat{x}_j, \underline{t}, L)\big]\Big\}$$

Solving for firms' mutual best replies, one obtain firms' subgame-perfect equilibrium differentiation as stated in the following proposition.

Proposition 2.1 (Equilibrium under ambiguity)**.** Under Assumptions 1,2, and 3, there is a unique pure strategy SPNE for the Hotelling game under ambiguity. Firms' equilibrium locations are given by

$$x_1^* = \frac{\delta\left(-(\alpha-1)(2L+3)^2\bar{t} + \alpha(3-2L)^2\underline{t} - 4\sigma^2 - 9\right) + 4\sigma^2 + 9}{4(\delta((\alpha-1)(2L+3)\bar{t} + \alpha(2L-3)\underline{t} + 3) - 3)}$$

$$x_2^* = \frac{\delta\left((\alpha-1)(2L+3)^2\bar{t} - \alpha(3-2L)^2\underline{t} + 4\sigma^2 + 9\right) - 4\sigma^2 - 9}{4(\delta((\alpha-1)(2L+3)\bar{t} + \alpha(2L-3)\underline{t} + 3) - 3)}$$

The equilibrium differentiation, $\Delta_x^* := x_2^* - x_1^*$, is

$$\Delta_x^* = \frac{\delta\left((\alpha-1)(2L+3)^2\overline{t} - \alpha(3-2L)^2\underline{t} + 4\sigma^2 + 9\right) - 4\sigma^2 - 9}{2(\delta((\alpha-1)(2L+3)\overline{t} + \alpha(2L-3)\underline{t} + 3) - 3)},$$

and firms' Choquet expected equilibrium profits are given by

$$\mathbb{CEU}[\Pi_i] = -\frac{\left(\delta\left(-(\alpha-1)(2L+3)^2\overline{t} + \alpha(3-2L)^2\underline{t} - 4\sigma^2 - 9\right) + 4\sigma^2 + 9\right)^2}{36(\delta((\alpha-1)(2L+3)\overline{t} + \alpha(2L-3)\underline{t} + 3) - 3)}$$

where $\Pi_i^* := \Pi_i(x_1^*, x_2^*, t, M)$.

Proof. Firms' equilibrium product designs and the proof of this proposition are contained in the appendix.

Remark 2.3. It is worthwhile to highlight and discuss some special cases of this equilibrium. Setting $\delta = 1$ and $\overline{t} = 1$, which corresponds to a situation under complete ambiguity, or without any confidence into the reference prior q, one obtains the equilibrium of Król [2012] in its full generality. Setting $\delta = 0$ and $\underline{t} = \overline{t} = 1$, we obtain the equilibrium in Meagher and Zauner [2004] with the slight difference that we impose a probability with zero mean. The normalization $\mathbb{E}_q[M] = 0$ ensures symmetry and is, in our view, not a strong restriction. We can interpret this assumption in the following way: Both firms determine the expected midpoint of the consumer interval and align all possible product designs symmetrically around this mean. If the mean is nonzero, firms can transform the set of all product characteristics to be centered around zero. After determining their product characteristic choices in the normalized setting, firms may retransform their product characteristic decision into the non-normalized product space and obtain the optimal product design. For consumer distributions with nonzero mean there are no solutions in closed-form for firms' subgame-perfect product characteristic choices. Nevertheless, it is plausible to argue that both firms shift their subgame-perfect locations into the direction of this mean.

2.7 Comparative Statics

The aforementioned Hotelling model under ambiguity yields interesting comparative static results. In this section, we discuss and interpret basic properties of firms' product design choices with respect to changes in the underlying model parameters. Similar to Król [2012], the following proposition examines ceteris paribus *(henceforth c.p.)* variations in the global ambiguity attitude α.

Proposition 2.2 (Variation in firms' ambiguity attitude α)**.** Under the Assumptions 1,2, and 3, one can observe at any SPNE of the Hotelling game under ambiguity the subsequent effects on optimal product designs:

$$\frac{\partial x_1^*}{\partial \alpha} \geq 0 \qquad \text{and} \qquad \frac{\partial x_2^*}{\partial \alpha} \leq 0$$

Proof. The proof of the proposition is contained in the appendix.

The results of Proposition 2.2 are related to the findings in Król [2012] stating that a higher degree of pessimism leads to lower product differentiation. This finding extends to our model, with the difference that the magnitude of the effect is weakened the more confidence firms have in the reference prior q. In case of full confidence, or absence of ambiguity, firms' attitude towards ambiguity becomes irrelevant for their product differentiation choices. To give some intuition: For a high degree of pessimism α, each firm gives a larger weight on the maxmin criterion than on the maxmax criterion. Therefore, the worst-case scenario becomes increasingly important. The worst-case of firm 1 is that the expectation of M equals L. As the expectation moves to the right and firm 1 considers this expectation as relevant, firm 1 has an incentive to select a product characteristic located on the right hand side of its initial characteristics. Similarly, for firm 2, the worst-case scenario corresponds to left boundary of the support $-L$. Since firm 2 gives increasingly more weight to this worst-case, there is an incentive for the latter to relocate to the left. All in all, equilibrium differentiation decreases.

To sum up these findings, we conclude that, contrary to the risk models of MZ, ambiguity is not per se a differentiation force. What matters is ambiguity attitude of both firms. We call this attitude the degree of global optimism or pessimism, since we consider a market where both firms exhibit the same ambiguity attitude. Hence, attitude towards ambiguity becomes a global characteristic of the market and could be interpreted as 'market sentiment'.

As a next step, we examine c.p. variations in the variance of the reference prior σ^2.

Proposition 2.3 (Variation in the variance σ^2)**.** If $0 \leq \delta < 1$ and the Assumptions 1, 2, and 3 hold, one has at any SPNE for the Hotelling game under ambiguity that optimal product designs react in the following way to an increase in σ^2:

$$\frac{\partial x_1^*}{\partial \sigma^2} < 0 \qquad \text{and} \qquad \frac{\partial x_2^*}{\partial \sigma^2} > 0$$

Proof. The proof of the proposition is contained in the appendix.

Uncertainty, as measured by the variance of the underlying distribution, constitutes a differentiation force. The intuition here[14] is that, in the Hotelling game, firms are confronted with two countervailing effects. If a firm selects, at given prices, a product characteristic that is more far away from the realized midpoint $\hat{M}$ than the characteristic selected by its competitor, it looses market share (*demand effect*). At the same time, however, one can observe that increasing product differentiation weakens price competition and leads to higher equilibrium prices (*price effect*). Due to the assumption of quadratic cost functions, the price effect dominates the demand effect. If a firm faces demand location uncertainty, the negative effect of loosing market shares in some realizations of uncertainty is not so dramatic as in the certainty case since there are other realizations of M where the latter's product design is better located than before. Consequently, an increasing variance of the underlying probability distribution strengthens the dominance of the price effect. Therefore, equilibrium differentiation is even more excessive than under

[14]Compare Meagher and Zauner [2004].

certainty. Of course, the same interpretation applies for the capacity model as long as $0 \leq \delta < 1$ with the sole difference that the effect of a c.p. increase in σ^2 is weaker the less confident firms are in the reference prior q.
The following proposition examines c.p. variations in the lower and upper support boundary of the transportation cost parameter.

Proposition 2.4 (Variations in the magnitude of the upper and lower support boundaries of t)**.** If $0 < \alpha \leq 1$ and $0 < \delta \leq 1$, then, under the Assumptions 1, 2, and 3, one has at any SPNE for the Hotelling game under ambiguity that

$$\frac{\partial x_1^*}{\partial \underline{t}} > 0 \qquad \text{and} \qquad \frac{\partial x_2^*}{\partial \underline{t}} < 0.$$

Similarly, for $0 \leq \alpha < 1$ and $0 < \delta \leq 1$, one obtains

$$\frac{\partial x_1^*}{\partial \bar{t}} < 0 \qquad \text{and} \qquad \frac{\partial x_2^*}{\partial \bar{t}} > 0.$$

Proof. The proof of the proposition is contained in the appendix.

The first part of Proposition 2.4 is quite similar to the respective statement in Król [2012]. Variations in the support of the transportation cost parameter can be interpreted as fluctuations in the magnitude of uncertainty around t. As $\underline{t}$ approaches one, the overall size of uncertainty with respect to t decreases. A ceteris paribus increase in $\underline{t}$ solely affects the pessimistic part of firms' first-stage profit functions. This deceases firms' equilibrium product differentiation. The following considerations explain why this is the case. Comparing the Hotelling model with a standard symmetric Bertrand competition, we observe the following important difference. In the standard Bertrand scenario, firms offer homogeneous products. The only Nash equilibrium in pure strategies is that firms set prices equal to marginal costs, implying zero profits for both firms. In a Bertrand world with heterogeneous products this finding is no longer true. By introducing transportation costs, the Hotelling framework adds an additional distinctive feature to a homogeneous and symmetric Bertrand competition rendering products per se more heterogeneous. It is therefore intuitive that a higher transportation cost weakens competition between firms. In the Hotelling model there are two countervailing incentives at work that determine firms' product design choices. One is that firms want to locate in the center of the Hotelling interval in order to obtain a higher market share. This is because firms' market share depends on the so-called indifferent consumer ξ.[15] All consumers located left of ξ strictly prefer to purchase the good from the firm located left. On the other hand, consumers located right of ξ strictly prefer to purchase the good from the other firm. If the firm located left c.p. relocates to the right, then the indifferent consumer also shifts to the right. In this case, the market share of this firm increases and, as a consequence, also its profits. A similar argument holds for the rival firm. If the firm located at the right c.p. relocates to the left, then its market share increases, and hence also its profit. To sum up, the firm located left has an incentive to relocate to the right and the firm located to the right has an incentive to relocate to the left.

The second incentive is that firms want to differentiate their products more in order to weaken price competition. If product differentiation gets lower, price competition gets stronger since both product become increasingly homogeneous. Therefore, in the limit, the only distinguishing feature of a product boils down to its price. Now, if price competition is weakened by a higher transportation cost, it is plausible that firms have an incentive to reduce product differentiation in order to obtain a higher market share. To summarize the results. Increasing uncertainty with respect to the transportation cost parameter t entail a higher degree of product differentiation.

The following proposition explores a c.p. increase in firms' confidence level δ.

[15] The indifferent consumer ξ can be obtained by equating total costs $p_1 - t(\xi - x_1)^2 = p_2 - t(\xi - x_2)^2$ and solving this expression for ξ.

Proposition 2.5 (Variation in the confidence level δ)**.** Under the Assumptions 1, 2, and 3, one has at any SPNE for the Hotelling game under ambiguity that

$$\frac{\partial x_1^*}{\partial \delta} = -\frac{\partial x_2^*}{\partial \delta} = \begin{cases} < 0 & \text{for} \quad 0 \leq \alpha < \alpha^* \\ = 0 & \text{for} \quad \alpha = \alpha^* \\ > 0 & \text{for} \quad 1 \geq \alpha > \alpha^* \end{cases}$$

where $\alpha^* = \alpha^*(\delta, \underline{t}, \overline{t}, \sigma^2, L)$ is a cutoff-value defined by

$$\alpha^* = \frac{(2L+3)(3L-2\sigma^2)}{(2L+3)\overline{t}(3L-2\sigma^2) - (2L-3)\underline{t}(3L+2\sigma^2)}.$$

Taking these results together we obtain for Δ^*

$$\frac{\partial \Delta^*}{\partial \delta} = \begin{cases} > 0 & \text{for} \quad 0 \leq \alpha < \alpha^* \\ = 0 & \text{for} \quad \alpha = \alpha^* \\ < 0 & \text{for} \quad 1 \geq \alpha > \alpha^*. \end{cases}$$

Proof. The proof of the proposition is contained in the appendix.

The findings of Proposition 2.5 can be summarized in the following way: If firms' attitude to ambiguity exhibits sufficiently strong optimism, one can conclude that a lower confidence into the reference prior increases equilibrium differentiation. Adverse results hold for sufficiently pessimistic firms. Furthermore, there is an intermediate value of global pessimism α^* such that firms' equilibrium differentiation remains unchanged no matter which global confidence level firms might assign to the reference probability distribution of the midpoint M.

Proposition 2.6 (Variation in the size of the support L)**.** If $0 < \delta \leq 1$ and Assumptions 1, 2, and 3 hold, on has at any SPNE for the Hotelling game under ambiguity that

$$\frac{\partial x_1^*}{\partial L} = -\frac{\partial x_2^*}{\partial L} = \begin{cases} < 0 & \text{for} \quad 0 \leq \alpha < \hat{\alpha} \\ = 0 & \text{for} \quad \alpha = \hat{\alpha} \\ > 0 & \text{for} \quad 1 \geq \alpha > \hat{\alpha} \end{cases}$$

where $\hat{\alpha} \in [0, 1]$ is a cutoff-value with $\hat{\alpha} = \hat{\alpha}(\delta, \underline{t}, \bar{t}, \sigma^2)$. Taking these results together we obtain for Δ^*

$$\frac{\partial \Delta^*}{\partial L} = \begin{cases} > 0 & \text{for} \quad 0 \leq \alpha < \hat{\alpha} \\ = 0 & \text{for} \quad \alpha = \hat{\alpha} \\ < 0 & \text{for} \quad 1 \geq \alpha > \hat{\alpha}. \end{cases}$$

Proof. The proof of the proposition is contained in the appendix.

An increase in the support fosters decreasing product differentiation if firms are sufficiently pessimistic. For an intermediate value of pessimism firms do not relocate. If firms are sufficiently optimistic, an increase in L yields higher equilibrium differentiation.

Size of the Support and Degree of Ambiguity

The degree of ambiguity, or of firms' confidence in the reference prior, plays a central role in this paper. For this reason, we discuss in the following its meaning in conjunction with the support of the uncertainty. As Proposition 2.6 shows, our model replicates similar comparative static results as in Król [2012] by varying the length L of the support of the midpoint M. Even though similar product differentiation choices might be generated by variations in the size of the support L, as compared to variations in the confidence level δ, it indispensable to notice meaningful differences between the two sources of ambiguity. First of all, variations in L and δ might go in similar directions, but the magnitude of both effects is different. In fact, both effects are interrelated. An increase in the support has a stronger impact on equilibrium differentiation if firms' confidence in the reference prior is low. In case of full confidence, changes in the support do not affect firms' product design decisions. Secondly, there is a clear difference between both sources of uncertainty concerning economic applications. The support of M consists of all possible midpoint realizations of the consumer interval. Before firms perform their design choices, they anticipate all possible demand realizations and summarize them in the support interval $[-L, L]$. An increase in the support interval would imply that firms allow ex-ante for a larger variety of feasible demand realizations. In our view, it is plausible to argue that, in many economic applications, the size of the support L is an exogenously fixed variable. What would it actually mean if L was an endogenous variable? It would mean that firms adjust their views on possible demand realizations in the light of higher or lower uncertainties by including or excluding certain market demand scenarios. Furthermore, this would imply that firms were ex-ante wrongly informed or had not precise information about lower and upper bounds of market demand in face of uncertainty. We do not want to argue that such a scenario is completely implausible, our point is that the interpretation of support variations is closely tied to firms' wrong perception of possible demand realizations.

In contrast to the previous interpretation, c.p. variations in the confidence level δ depart on the assumption of an exogenously fixed support length. Firms know possible upper and lower bounds of demand and consider demand uncertainty defined on a fixed support. The reference prior q might reflect firms' ex-ante information about the market environment, e.g. firms might have observable data or can pursue market research to estimate an underlying probability distribution for market demand. Under the assumption that firm managers are sufficiently pessimistic, increasing product differentiation might have different reasons. One explanation could be that firms become more optimistic, meaning that due to a change in the market environment firms adjust their ambiguity attitudes to account for this new situation. On the other hand, it might be the case that firms obtain more reliable data on market outcomes, therefore increasing their confidence in the data available but do not readjust their attitude towards ambiguity. In such a scenario, a higher confidence into the reference prior weakens the impact of pessimism on product differentiation choices.

2.8 Examples and Applications

In this section, we apply our model to a variety of real-life examples. At first, we discuss sports betting regarding horse racing and football games. The second application refers to financial markets, or to be more precise to the mutual funds market. Furthermore, we want to mention that similar cases were already discussed in Król [2012]. The purpose of this section consists of providing the reader with additional insight into the mechanics of the capacity approach. In particular, we want discuss implications of confidence and pessimism for the interpretation of these examples. One reason why the aforementioned applications are so apt to be discussed in a Hotelling framework, is given by the fact that in these markets exists a relatively clear measure of firms' product differentiation. We will discuss these measures in the respective subsections. Moreover, consumer preferences are often fluctuating depending on partially unobserved factors, e.g. individual subjective evaluations. Due to firms' imperfect probabilistic information regarding market demand, it is plausible that ambiguity is prevalent in those markets.

Sports Betting

In case of sports betting, the odd of a bookmaker represents product characteristic. Furthermore, the preferences of a bettor over odds are determined by subjective probability estimations of the particular sporting event. Since bookmakers usually want to make a profit regardless of the result of the sporting event, one can assume that they are worst-case-oriented.[16] If bookmakers were rather optimistic, they would constantly offer odds

[16] For more details see Król [2012].

exceeding the expectation of the underlying distribution and eventually run the risk of bankruptcy.

Horse racing

Smith et al. [2009] examine horse racing data from the UK. The authors provide evidence for an increased fluctuation and divergence of betting exchange prices shortly before the race. This can be interpreted as a higher degree of ambiguity with respect to bettors' preferences. At the same time, bookmakers' odds are getting increasingly similar.[17] Supposing that horse racing bookmakers exhibit a sufficiently high degree of pessimism, our model provides a possible explanation for this observation. Recall that, given that firms are sufficiently pessimistic, an increase of ambiguity leads to decreasing product differentiation. The intuition here lies in the fact that pessimistic firms place more weight on worst-case scenarios. Moreover, the worst-case scenario for the firm whose product characteristic is located on the left hand side corresponds to the realization $M = L$. Similarly, the worst-case scenario for the firm on the right hand side corresponds to the realization $M = -L$. If firms become more pessimistic, the firm whose product design is located left selects a product characteristic right from its initial characteristic. Hence, the higher firms' pessimism, the lower equilibrium product differentiation. The strength of this effect increases with increasing ambiguity, δ, since firms' confidence in their prior belief determines the influence of the worst-case scenario on their decision process.

Football games

Bookmakers' odds on football games exhibit an interesting feature. Typically, whenever a rather strong team plays against a rather weak team there is little divergence between bookmakers' odds in favor of a victory of the strong team. In contrast, the odds for a victory of the weak team are more volatile. In fact, odds become less volatile when

[17] As pointed out by Król [2012], one can verify that the differences between bookmakers' odds are decreasing in the corresponding time period by using price comparison websites.

the perceived relative strength of both teams is fairly similar. This observation can be verified by comparing bookmakers' match day odds. In the examples below, we analyzed the odds of ten bookmakers.

Remark 2.4. Both examples are games from match day 22 on February 23, 2014 of the German Bundesliga.[18] The first example refers to a game of Bayern Munich, representing the strong team, versus Hannover 96 representing the weaker team. Stated odds are to be considered as multiplication factors of the placed bet in case of winning the bet. For instance, suppose one puts a bet of €1 in favor of a victory of Hannover 96 at bet365. In case Hannover 96 wins, the bettor receives €10. The second game, SC Freiburg versus FC Augsburg, is more balanced in terms of relative strength. Estimators used in our examples are

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{for the mean and } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \text{ for the variance.}$$

Bookmaker	**Odd Hannover 96**	**Odd Bayern Munich**
bet365	10	1.25
Sportingbet	11	1.222
Tipico	13	1.25
bwin	9.5	1.22
Interwetten	9	1.27
Bet-at-home	11.5	1.24
Betsson	12.5	1.19
mybet	14	1.22
Betvictor	10.5	1.25
Unibet	12	1.25
Estimated Variances	**2.5667**	**0.0005**

Table 2.1: Example for the Constellation of a Strong Team Versus a Weak Team

If bookmakers agreed on a unique prior over the outcome of the game, this phenomenon would be inexplicable. Again, the explanation might lie in bookmakers' confidence in their prior beliefs. Assume bookmakers' choose their odds such that a bettor will always

[18]Data was collected online on February 19th, 2014 at 3:30 pm from the respective websites of the bookmakers.

Bookmaker	Odd SC Freiburg	Odd FC Augsburg
bet365	3.1	2.25
Sportingbet	3	2.3
Tipico	3.1	2.35
bwin	2.85	2.3
Interwetten	2.75	2.4
Bet-at-home	3	2.3
Betsson	3.1	2.27
mybet	3.2	2.3
Betvictor	3.125	2.3
Unibet	2.95	2.35
Estimated Variances	**0.0189**	**0.017**

Table 2.2: Example for the Constellation of Two Balanced Teams

loose a fraction of his money if she bets on both teams, then odds on one team become a function of the odds of the other team.[19] In the situation described above, it is obviously very likely that the strong team wins. Hence, bookmakers' can be confident that the bulk of bettors will bet on the very strong team. This leads to two effects. Firstly, in order to avoid bankruptcy, bookmakers' need to choose odds close to one for a win of the strong team. Secondly, since bookmakers' face little ambiguity over bettors' preferences, they can differentiate their odds for the weak team. This result is in line with Proposition 2.5.

Mutual Funds

Król [2012] provides the example of the managed mutual funds' market. In this context, one can interpret a position in the product space as a portfolio's position ranging from safe investments to risky portfolios. Król [2012] shows, based on data about the daily returns of the fifteen most popular actively managed US mutual funds, that, after the financial crisis 2008, fund managers tend to differentiate their products less. Król [2012] argues that, before the crisis, financial firms' did not consider the post-crisis range of investor behavior as possible.[20] For this reason, the crisis forced firms to revise their

[19] See Król [2012].

[20] In particular, the shift of consumer preferences toward safe investments due to decreasing stock prices during the crisis.

beliefs. Furthermore, the author interprets conservative stress test simulations following the crisis as a signal sent out to the competitors that a firm uses a worst-case-based approach for decision making. This is exactly the point where we want to add to the debate. For instance, consider government-imposed stress tests after the crisis. If one interprets such stress tests as signals, then the strategy of a firm is independent of its type. Since each type sends the same signal, no new information is revealed to the other firms. In our view, it is debatable whether stress test simulations induced a shift in fund managers' ambiguity attitude towards a more pessimistic preference approach, or whether exactly those fund managers knew more clearly that investors would prefer more secure assets after the crisis. If so, a possible explanation for lower post-crisis product differentiation is that firms were less uncertain about investor preferences. In our view, it is not implausible that fund managers' ambiguity attitudes remained relatively stable even though government stress tests were imposed. Furthermore, due to market research and historical data[21] it is likely that fund managers know the whole range of possible individual investor behaviors.[22] However, investor preferences are highly fluctuating since they depend on investors' subjective evaluations of the fund's performance which itself is based on numerous observed and unobserved factors as recent stock market developments or individual future expectations. At the end and shortly after the financial crisis, firms' were highly confident in terms of investor preferences since it was self-evident that, post-crisis, the majority of investors would prefer assets which were rather safe. Again, this finding is in line with our model.

[21]Financial firms' can rely on past data of various historical economic crises including stock market crashes (e.g. the Great Depression in the 1930s), bubbles (e.g. dot-com bubble in 2000), and financial crises (e.g. Asian financial crisis in 1997).

[22]This would induce that variations in the support of the midpoint of the consumer distribution cannot account for the observation of decreasing product differentiation.

2.9 Conclusion

We present an extension of Hotelling's model incorporating ambiguity in the form of demand location uncertainty as well as uncertainty with respect to the intensity of transportation costs. Ambiguity is introduced by representing firms' beliefs with neo-additive capacities. We analyze firms' optimal product characteristic choices and find a unique SPNE in pure strategies for the Hotelling game under ambiguity.

Our model incorporates a variety of different sources of uncertainty. First of all, there is the variance σ^2 of the reference probability q. As in the standard risk case of Meagher and Zauner [2004], a higher variance implies that firms increase product differentiation. Thus, if the measure of uncertainty is given by the variance of the underlying reference probability, it can be considered as differentiation force.

Secondly, there is the length of the support interval of M. The larger the support of M, the larger the number of demand realizations that firms consider as possible market outcomes. Hence, the length of the support interval might be interpreted as an additional measure of uncertainty. As our results show, the effects on an increasing support are strongly related to firms' ambiguity attitude α and the degree of ambiguity δ. If firms are rather pessimistic, a larger support results in lower equilibrium differentiation, if firms are rather optimistic, a larger support engenders opposite results. All in all, uncertainty as measured by the support length can be - depending on parameters - a differentiation or agglomeration force.

A third measure of uncertainty is given by the confidence parameter δ reflecting firms uncertainty on observables. Interpreted in this way, rising uncertainty is tied to lower data reliability yielding lower confidence levels in the reference probability q. Again, similar to the case of support variations, this can trigger off opposing effects. When firms are pessimistic enough, equilibrium differentiation is going down, when firms a rather optimistic product differentiation is going to increase. One can also argue the other way round. For a given confidence level, increasing pessimism yields lower equilibrium differentiation, whereas an increase in optimism increases equilibrium differentiation.

Finally the last source of uncertainty lies in the support $[\underline{t}, \overline{t}]$ of the transportation cost parameter t. As the lower support boundary $\underline{t}$ decreases, firms' equilibrium differentiation remains the same in case of full optimism and full confidence and increases in all other cases. Similarly, as the upper support boundary $\overline{t}$ increases, firm' equilibrium differentiation remains the same in case of full pessimism and full confidence and increases in all other cases. Thus, excluding these boundary cases, we can say that the size of uncertainty with respect to the transportation cost parameter constitutes an differentiation force.

As we can see from the preceding line of arguments, one should be very cautious when it comes to drawing conclusions from real-world applications of Hotelling models under uncertainty. In our view, it is indispensable to clearly identify the driving factors of an observed increase or decrease in product differentiation since the interpretation and conclusions from observed firm behavior might change in the light of different sources of uncertainty. In particular, it seems worthwhile for policymakers to disentangle the effect of confidence and ambiguity attitude on product differentiation, since it might really matter for official regulatory procedures whether observed product differentiation choices are to be attributed to perceived changes in data-reliability or whether firms feature more or less optimistic behavioral patterns.

2.10 Mathematical Proofs

Proof of Lemma 2.1. The support of M is restricted to the interval $[-L, L] \subset \left[-\frac{1}{2}, \frac{1}{2}\right]$. The mean and the variance of M exists. For the mean we can perform the following line of estimates:

$$\mathbb{E}_q[M] = \int\limits_{\mathbb{R}} M d\mathbb{P} \leq \int\limits_{\mathbb{R}} L d\mathbb{P} = L \int\limits_{\mathbb{R}} 1 d\mathbb{P} = L$$

and

$$\mathbb{E}_q[M] = \int\limits_{\mathbb{R}} M d\mathbb{P} \geq \int\limits_{\mathbb{R}} -L d\mathbb{P} = -L \int\limits_{\mathbb{R}} 1 d\mathbb{P} = -L$$

Similarly, for the second moment of M we obtain

$$E[M^2] = \int_{\mathbb{R}} M^2 d\mathbb{P} \leq \int_{\mathbb{R}} L^2 d\mathbb{P} = L^2 \quad \text{and} \quad \mathbb{E}_q[M^2] = \int_{\mathbb{R}} M^2 d\mathbb{P} \geq 0$$

and for the variance σ^2 we conclude

$$\sigma_M^2 = \mathbb{E}_q[M^2] - \mathbb{E}_q[M]^2 \leq \mathbb{E}_q[M]^2 \leq L^2 \quad \text{and} \quad \sigma_M^2 \geq 0.$$

□

Proof of Lemma 2.4. Lemma 2.5 implies that firms' second-stage profits at the realization $(\hat{t}, \hat{M})$ equal the second piece of (2.1):

$$\begin{aligned} \Pi_1 &= \frac{1}{18}\hat{t}(x_2 - x_1)[-3 + 2(\hat{M} - \bar{x})]^2 \\ \Pi_2 &= \frac{1}{18}\hat{t}(x_2 - x_1)[3 + 2(\hat{M} - \bar{x})]^2 \end{aligned}$$

Both profit functions are continuously differentiable with respect to $\hat{t}$ and $\hat{M}$. Differentiation with respect to $\hat{t}$ yields

$$\begin{aligned} \frac{\partial \Pi_1}{\partial \hat{t}} &= \frac{2}{9}(x_2 - x_1)\left[\frac{x_1 + x_2}{2} - \hat{M} + \frac{3}{2}\right]^2 > 0 \\ \frac{\partial \Pi_2}{\partial \hat{t}} &= -\frac{2}{9}(x_1 - x_2)\left[\frac{x_1 + x_2}{2} - \hat{M} - \frac{3}{2}\right]^2 > 0. \end{aligned}$$

Differentiation with respect to $\hat{M}$ yields

$$\begin{aligned} \frac{\partial \Pi_1^*}{\partial \hat{M}} &= \underbrace{-\frac{4}{9}\hat{t}(x_2 - x_1)}_{<0} \underbrace{\left[\frac{x_1 + x_2}{2} - \hat{M} + \frac{3}{2}\right]}_{>0} < 0 \\ \frac{\partial \Pi_2^*}{\partial \hat{M}} &= \underbrace{\frac{4}{9}\hat{t}(x_1 - x_2)}_{<0} \underbrace{\left[\frac{x_1 + x_2}{2} - \hat{M} - \frac{3}{2}\right]}_{<0} > 0. \end{aligned}$$

□

Proof of Lemma 2.5. The proof of the lemma follows exactly the same line of arguments as in the proof of Lemma 3.1 in Król [2012], page 602 with a slight modification in case 3. There are three different cases to be considered.

1. Case 1 refers to a situation where either firm 1 or firm 2 can monopolize the market for certain realizations of the midpoint M. If firm 1 can monopolize the market for certain realizations of M, we can conclude that firm 1 will monopolize the market if $\hat{M} = -L$, since w.lo.g. firm 1 is the firm left of firm 2. Similarly, we can conclude that firm 2 can monopolize the market for $\hat{M} = L$. This is finding is impossible. If firm 1 monopolizes the market for the realization $\hat{M} = -L$, we have by Lemma 3.1, equation (2.3) that $\frac{x_1+x_2}{2} - \frac{3}{2} > -L$. If firm 2 monopolizes the market, we have by (2.3) that $\frac{x_1+x_2}{2} + \frac{3}{2} < L$. Thus, we must have that $L + \frac{x_1+x_2}{2} > \frac{3}{2}$ and $L - \frac{x_1+x_2}{2} > \frac{3}{2}$ holds at the same time implying $\left|\frac{x_1+x_2}{2}\right| < L - \frac{3}{2}$. This is a contradiction since L is assumed to be smaller than $\frac{1}{2}$.

2. Case 2 describes a scenario where one of the two firms can monopolize the market for each realization $\hat{M}$ of uncertainty. If firm j is a monopolist, the other firm can deviate from its original location in order to obtain a positive market share and therefore make strictly positive profits. Król [2012] suggests the location $x_{-j} = -x_j$.

3. Case 3 refers to a situation where, w.l.o.g., firm 1 can monopolize the market for some realizations of uncertainty, in particular the realization $\hat{M} = -L$ and for the remaining realizations, in particular the realization $\hat{M} = L$, there exists a competitive equilibrium. Consider now the profit function of firm 2 in case of a competitive equilibrium[23] :

$$\frac{\partial \Pi_2}{\partial x_2}(x_1, x_2, L, t) = \frac{t\ (2\,L - 3\,x_2 + x_1 + 3)\ (2\,L - x_2 - x_1 + 3)}{18}$$

We want to show that

$$\frac{\partial \Pi_2}{\partial x_2}(x_1, x_2, L, t) < 0.$$

[23] We consider the profit function of firm 2, Król [2012] considers the profit function of firm 1.

We determine the sign of both brackets. Consider the expression in within the second bracket first. We have

$$2L + 3 - x_1 - x_2 > 0 \quad \Leftrightarrow \quad 2L + 3 > x_1 + x_2 \quad \Leftrightarrow \quad L + \frac{3}{2} > \bar{x}$$

The last condition corresponds to the requirement for a competitive solution in cases where the midpoint $M = L$ realizes. Therefore it must be, by assumption, positive. The second bracket is negative. The monopolistic outcome for the midpoint realization $M = -L$ requires $L + \bar{x} > \frac{3}{2}$. Solving this inequality for x_2, we obtain $x_2 > 3 - 2L - x_1$. By using this inequality, we can conduct an estimation for the expression in the first bracket:

$$3 + 2L + x_1 - 3x_2 < 8L - 6 + 4x_1 < 8L - 8 < 0$$

The last inequality follows from the fact that $L < \frac{1}{2}$ and $x_1 < 0$. Thus, we proved that

$$\frac{\partial \Pi_2}{\partial x_2}(x_1, x_2, L, t) < 0.$$

This finding shows that firm 2 has an incentive to move leftwards in order to reduce both firms' product differentiation and that a strict competitive solution does not exist under the above stated parameter restrictions. Since we consider a symmetric scenario, a similar argument holds for a scenario where firm 2 becomes a monopolist. For the remaining cases $\hat{M} = L$ and $\hat{M} = -L$ there is a competitive solution.

□

Proof of Lemma 2.6. The first part of firms' Choquet expected profit is

$$E_q[\Pi_i(x_1, x_2, t, M)] = \int_{-L}^{L} (-1)^j \frac{2}{9} t (x_j - x_i) \left(\frac{x_i + x_j}{2} - \left(M + \frac{3}{2}(-1)^i \right) \right)^2 f(M) dM.$$

This expectation is of the form

$$\mathbb{E}_q[g_i(t)\, h_i(M)]$$

with real-valued Borel-measurable functions g_i and h_i for $i = 1, 2$. We define

$$g_i(t) = t \qquad \text{and} \qquad h_i(M) = (-1)^j \, \frac{2}{9}\, t\,(x_j - x_i) \left(\frac{x_i + x_j}{2} - \left(M + \frac{3}{2}(-1)^i \right) \right)^2 .$$

By $(R7)$, t and M are uncorrelated. By Lemma 5.20 in Meintrup and Schäffler [2006], page 131, we obtain that $g_i(t)$ and $h_i(M)$ are uncorrelated as well. Thus, we can conclude

$$\mathbb{E}_q[\Pi_i^*(x_1, x_2, t, M)] = E_q[g_i(t)\, h_i(M)] = \mathbb{E}_q[g_i(t)]\, \mathbb{E}_q[h_i(M)] = \mu_t\, \mathbb{E}_q[h_i(M)].$$

In the following, we can rely on the results in Meagher and Zauner [2004], page 205, since $\mathbb{E}_q[h_i(M)]$ is equal to firm i's expected profit function in the risk case. Thus,

$$\mathbb{E}_q[\Pi_i(x_1, x_2, t, M)] = t_\mu \int_{-L}^{L} (-1)^j \, \frac{2}{9}\, (x_j - x_i) \, \left(\frac{x_i + x_j}{2} - \left(M + \frac{3}{2}(-1)^i \right) \right)^2 \, f(M) dM$$

$$= \frac{(-1)^j}{18}\, t_\mu\, (x_j - x_i) \{ (x_i + x_j - 3(-1)^i)^2$$

$$- 4\mu_M (x_i + x_j - 3(-1)^i) + 4(\mu_M + \sigma_M^2) \}$$

□

Proof of Proposition 2.1. We derive expected CEU profits at the first stage of the game. We obtain for firm 1:

$$\begin{aligned}
&\mathbb{CEU}[\Pi_1(x_1, x_2, \alpha, \delta, \underline{t}, \overline{t}, \sigma^2, L)] \\
&:= \delta \left(\frac{2\,(1-\alpha)\,\overline{t}\,\,(x_2 - x_1)\,\left(L + \frac{x_2+x_1}{2} + \frac{3}{2}\right)^2}{9} + \frac{2\,\alpha\,\underline{t}\,\,(x_2 - x_1)\,\left(-L + \frac{x_2+x_1}{2} + \frac{3}{2}\right)^2}{9} \right) \\
&+ \frac{(1-\delta)\,\,(x_2 - x_1)\,\left((x_2 + x_1 + 3)^2 + 4\,\sigma^2\right)}{18}.
\end{aligned} \tag{A.1}$$

Similarly, we obtain for firm 2

$$\begin{aligned}
&\mathbb{CEU}[\Pi_2(x_1, x_2, \alpha, \delta, \underline{t}, \overline{t}, \sigma^2, L)] \\
&:= \delta \left(\frac{2\,\alpha\,\underline{t}\,\,(x_2 - x_1)\,\left(L + \frac{x_2+x_1}{2} - \frac{3}{2}\right)^2}{9} + \frac{2\,(1-\alpha)\,\overline{t}\,\,(x_2 - x_1)\,\left(-L + \frac{x_2+x_1}{2} - \frac{3}{2}\right)^2}{9} \right) \\
&+ \frac{(1-\delta)\,\,(x_2 - x_1)\,\left((x_2 + x_1 - 3)^2 + 4\,\sigma\right)}{18}.
\end{aligned} \tag{A.2}$$

Taking the derivative of (A.1) with respect to x_1 yields

$$\begin{aligned}
&\frac{\partial \mathbb{CEU}[\Pi_1(x_1, x_2, \alpha, \delta, \underline{t}, \overline{t}, \sigma^2, L)]}{\partial x_1} := \\
&- \frac{2\delta\,\,(1-\alpha)\,\,\overline{t}\left(L + \frac{x_2+x_1}{2} + \frac{3}{2}\right)^2}{9} + \frac{2\delta\,\,(1-\alpha)\,\,\overline{t}\,\,(x_2 - x_1)\,\left(L + \frac{x_2+x_1}{2} + \frac{3}{2}\right)}{9} \\
&+ \frac{2\delta\,\alpha\,\underline{t}\,\,(x_2 - x_1)\,\left(-L + \frac{x_2+x_1}{2} + \frac{3}{2}\right)}{9} - \frac{2\delta\,\alpha\,\underline{t}\left(-L + \frac{x_2+x_1}{2} + \frac{3}{2}\right)^2}{9} \\
&- \frac{(1-\delta)\,\left((x_2 + x_1 + 3)^2 + 4\,\sigma\right)}{18} + \frac{(1-\delta)\,\,(x_2 - x_1)\,\,(x_2 + x_1 + 3)}{9}.
\end{aligned} \tag{A.3}$$

Similarly, we take the derivative of (A.2) with respect to x_2

$$\begin{aligned}
&\frac{\partial \mathbb{CEU}[\Pi_2(x_1,x_2,\alpha,\delta,\underline{t},\overline{t},\sigma^2,L)]}{\partial x_2} := \\
&\frac{2\delta\,\alpha\,\underline{t}\left(L+\frac{x_2+x_1}{2}-\frac{3}{2}\right)^2}{9}+\frac{2\delta\,\alpha\,\underline{t}\;(x_2-x_1)\;\left(L+\frac{x_2+x_1}{2}-\frac{3}{2}\right)}{9} \\
&+\frac{2\delta\;(1-\alpha)\;\overline{t}\;(x_2-x_1)\;\left(-L+\frac{x_2+x_1}{2}-\frac{3}{2}\right)}{9}+\frac{2\delta\;(1-\alpha)\;\overline{t}\left(-L+\frac{x_2+x_1}{2}-\frac{3}{2}\right)^2}{9} \\
&+\frac{(1-\delta)\;\left((x_2+x_1-3)^2+4\,\sigma\right)}{18}+\frac{(1-\delta)\;(x_2-x_1)\;(x_2+x_1-3)}{9}
\end{aligned} \tag{A.4}$$

Now, we solve the following system of equations:

$$\begin{aligned}
\frac{\partial \mathbb{CEU}[\Pi_1(x_1,x_2,\alpha,\delta,\underline{t},\overline{t},\sigma^2,L)]}{\partial x_2} &= 0 \\
\frac{\partial \mathbb{CEU}[\Pi_2(x_1,x_2,\alpha,\delta,\underline{t},\overline{t},\sigma^2,L)]}{\partial x_2} &= 0
\end{aligned} \tag{A.5}$$

and obtain three solution pairs. The first solution pair (x_1^*, x_2^*) is given by:

$$x_1^* = \frac{\delta\left(-(\alpha-1)(2L+3)^2\overline{t}+\alpha(3-2L)^2\underline{t}-4\sigma^2-9\right)+4\sigma^2+9}{4(\delta((\alpha-1)(2L+3)\overline{t}+\alpha(2L-3)\underline{t}+3)-3)}$$

$$x_2^* = \frac{\delta\left((\alpha-1)(2L+3)^2\overline{t}-\alpha(3-2L)^2\underline{t}+4\sigma^2+9\right)-4\sigma^2-9}{4(\delta((\alpha-1)(2L+3)\overline{t}+\alpha(2L-3)\underline{t}+3)-3)}$$

The second pair of solutions (x_1^{**}, x_2^{**}) is given by:

$$\begin{aligned}
x_1^{**} = \Bigg(&-\Big(\delta(\delta((\alpha-1)^2(2L+3)^2\overline{t}^2+2(\alpha-1)\overline{t}(2L(6\alpha L\underline{t}-L+3)-9\alpha\underline{t}+9) \\
&+\alpha\underline{t}(4L(\alpha(L-3)\underline{t}+L+3)+9(\alpha\underline{t}-2))+4\sigma^2(-\alpha\overline{t}+\alpha\underline{t}+\overline{t}-1)+9) \\
&+4(\alpha-1)\overline{t}((L-3)L+\sigma^2)-2\alpha\underline{t}(2L(L+3) \\
&+2\sigma^2-9)+2(-9(\alpha-1)\overline{t}+4\sigma^2-9))-4\sigma^2+9\Big)^{\frac{1}{2}} \\
&+\delta(-(\alpha-1)(2L+3)\overline{t}+\alpha(3-2L)\underline{t}-3)+3\Bigg) \\
&\cdot\Big(2(\delta((\alpha-1)\overline{t}-\alpha\underline{t}+1)-1)\Big)^{-1}
\end{aligned}$$

and

$$
\begin{aligned}
x_2^{**} = -\Bigg(& \Big(\delta(\delta((\alpha-1)^2(2L+3)^2\bar{t}^2 + 2(\alpha-1)\bar{t}(2L(6\alpha L\underline{t} - L + 3) - 9\alpha\underline{t} + 9) \\
& + \alpha\underline{t}(4L(\alpha(L-3)\underline{t} + L + 3) + 9(\alpha\underline{t} - 2)) + 4\sigma^2(-\alpha\bar{t} + \alpha\underline{t} + \bar{t} - 1) + 9) \\
& + 4(\alpha-1)\bar{t}((L-3)L + \sigma^2) - 2\alpha\underline{t}(2L(L+3) + 2\sigma^2 - 9) \\
& + 2(-9(\alpha-1)\bar{t} + 4\sigma^2 - 9)) - 4\sigma^2 + 9\Big)^{\frac{1}{2}} \\
& - \delta((\alpha-1)(2L+3)\bar{t} + \alpha(2L-3)\underline{t} + 3) + 3\Bigg) \\
& \cdot \Big(2(\delta((\alpha-1)\bar{t} - \alpha\underline{t} + 1) - 1)\Big)^{-1}
\end{aligned}
$$

Finally, the last pair of solutions (x_1^{***}, x_2^{***}) is given by:

$$
\begin{aligned}
x_1^{***} = \Bigg(& \Big(\delta(\delta((\alpha-1)^2(2L+3)^2\bar{t}^2 + 2(\alpha-1)\bar{t}(2L(6\alpha L\underline{t} - L + 3) - 9\alpha\underline{t} + 9) \\
& + \alpha\underline{t}(4L(\alpha(L-3)\underline{t} + L + 3) + 9(\alpha\underline{t} - 2)) + 4\sigma^2(-\alpha\bar{t} + \alpha\underline{t} \\
& + \bar{t} - 1) + 9) + 4(\alpha-1)\bar{t}((L-3)L + \sigma^2) - 2\alpha\underline{t}(2L(L+3) + 2\sigma^2 - 9) \\
& + 2(-9(\alpha-1)\bar{t} + 4\sigma^2 - 9)) - 4\sigma^2 + 9\Big)^{\frac{1}{2}} \\
& + \delta(-(\alpha-1)(2L+3)\bar{t} + \alpha(3-2L)\underline{t} - 3) + 3\Bigg) \\
& \cdot \Big(2(\delta((\alpha-1)\bar{t} - \alpha\underline{t} + 1) - 1)\Big)^{-1}
\end{aligned}
$$

and

$$
\begin{aligned}
x_2^{***} = \Bigg(& \Big(\delta(\delta((\alpha-1)^2(2L+3)^2\overline{t}^2 + 2(\alpha-1)\overline{t}(2L(6\alpha L\underline{t} - L + 3) \\
& - 9\alpha\underline{t} + 9) + \alpha\underline{t}(4L(\alpha(L-3)\underline{t} + L + 3) + 9(\alpha\underline{t} - 2)) + 4\sigma^2(-\alpha\overline{t} + \alpha\underline{t} \\
& + \overline{t} - 1) + 9) + 4(\alpha-1)\overline{t}((L-3)L + \sigma^2) - 2\alpha\underline{t}(2L(L+3) + 2\sigma^2 - 9) \\
& + 2(-9(\alpha-1)\overline{t} + 4\sigma^2 - 9)) - 4\sigma^2 + 9\Big)^{\frac{1}{2}} \\
& + \delta((\alpha-1)(2L+3)\overline{t} + \alpha(2L-3)\underline{t} + 3) - 3\Bigg) \\
& \cdot \Big(2(\delta((\alpha-1)\overline{t} - \alpha\underline{t} + 1) - 1)\Big)^{-1}
\end{aligned}
$$

The first pair of solutions (x_1^*, x_2^*) satisfies the global competition condition according to Lemma 2.5. We demonstrate this by using Wolfram Mathematica version 10.0.0.0. You can find the code at the end of the proof section. The problem is analyzed in sections 5 to 7 of the code. Mathematica returns the value "true" for the first pair of solutions. The solution pairs (x_1^{**}, x_2^{**}) and (x_1^{***}, x_2^{***}) do not fulfill the global competition condition

$$L - \frac{3}{2} < \bar{x} < -L + \frac{3}{2}.$$

This is examined in sections 8 and 9 of our code. Therefore, we define, in a first step, the means

$$\overline{x}_2 = \frac{x_1^{**} + x_2^{**}}{2} \qquad \text{and} \qquad \overline{x}_3 = \frac{x_1^{***} + x_2^{***}}{2}.$$

Using numerical optimization techniques, we obtain that the range of $\overline{x}_2$ is given by $[1, 2]$. Similarly, the range of $\overline{x}_3$ is given by $[-2, -1]$. Moreover, $\overline{x}_2$ attains its minimum value 1 for $L = \frac{1}{2}$. This implies that $\overline{x}_2 \geq 1$. However, the global competition condition would require that $\overline{x}_2 < -\frac{1}{2} + \frac{3}{2} = 1$. This is a contradiction. Similarly, $\overline{x}_3$ attains its maximum value -1 for $L = \frac{1}{2}$. As a consequence, we can infer that $\overline{x}_3 \leq -1$. In order to meet the requirements of Lemma 2.5, $\overline{x}_3$ also needs to satisfy $\overline{x}_3 > \frac{1}{2} - \frac{3}{2} = -1$. This excludes (x_1^{***}, x_2^{***}) as a feasible solution.

As a next step, we show that the first pair of solutions is indeed a maximizer for both firms. The second order derivative of evaluated at (x_1^*, x_2^*) yields

$$
\begin{aligned}
\frac{\partial^2 \mathbb{CEU}[\Pi_i(x_1^*, x_2^*, \cdot)]}{\partial x_i^2} := \Big(&\delta(\delta(3(\alpha-1)^2(2L+3)^2\bar{t}^2 \\
&+ 2(\alpha-1)\bar{t}(2L(10\alpha L\underline{t} - L + 9) - 27\alpha\underline{t} + 27) \\
&+ \alpha\underline{t}(3\alpha(3-2L)^2\underline{t} + 4L(L+9) - 54) \\
&+ 4\sigma^2(-\alpha\bar{t} + \alpha\underline{t} + \bar{t} - 1) + 27) + 4(\alpha-1)\bar{t}((L-9)L \\
&+ \sigma^2) - 2\alpha\underline{t}(2L(L+9) + 2\sigma^2 - 27) - 54((\alpha-1)\bar{t} + 1) \\
&+ 8\sigma^2) - 4\sigma^2 + 27\Big) \\
&\cdot \Big(18(\delta((\alpha-1)(2L+3)\bar{t} + \alpha(2L-3)\underline{t} + 3) - 3)\Big)^{-1}
\end{aligned}
$$

for both firms. First, I examine the sign of the denominator. It is

$$
\begin{aligned}
&18(\delta((\alpha-1)(2L+3)\bar{t} + \alpha(2L-3)\underline{t} + 3) - 3) \\
&= 36\alpha\delta L\bar{t} + 36\alpha\delta L\underline{t} + 54\alpha\delta\bar{t} - 54\alpha\delta\underline{t} - 36\delta L\bar{t} - 54\delta\bar{t} + 54\delta - 54 \\
&\leq 36\delta L\bar{t} + 18\alpha\delta\underline{t} + 54\delta\bar{t} - 54\alpha\delta\underline{t} - 36\delta L\bar{t} - 54\delta\bar{t} + 54 - 54 \\
&= -36\alpha\delta\underline{t} \\
&\leq 0
\end{aligned}
$$

Hence, the denominator is negative. Subsequently, we show that the numerator is non-negative. Taking the derivative of the numerator with respect to $\bar{t}$ yields

$$
\begin{aligned}
&- 2(1-\alpha)\delta(\delta(3(\alpha-1)(2L+3)^2\bar{t} + 2L(L(10\alpha\underline{t} - 1) + 9) \\
&- 27\alpha\underline{t} - 2\sigma^2 + 27) + 2((L-9)L + \sigma^2) - 27)
\end{aligned}
$$

Given the parameter restrictions of the model, this expression is non-negative. We verify this in sections 13 and 14 of the Mathematica code. Hence, the numerator becomes smaller as we insert the minimum value 1 for $\bar{t}$. Doing so, we obtain after several steps of algebra

$$
\begin{aligned}
h =&\Big(\alpha^2\delta^2\Big(4L^2(\underline{t}+3)(3\underline{t}+1)-36L\Big(\underline{t}^2-1\Big)+27(\underline{t}-1)^2\Big)\\
&-2\alpha\delta\Big(2L^2(9\delta\underline{t}+7\delta+\underline{t}-1)\\
&+18L(\delta(-\underline{t})+\delta+\underline{t}+1)-(\underline{t}-1)(2(\delta-1)\sigma^2+27)\Big)\\
&+4\delta(L((4\delta-1)L+9)+\sigma^2)-4\sigma^2+27\Big)
\end{aligned}
$$

What remains to be demonstrated is that this expression is non-negative. Using Mathematica, we check whether h can be negative under the restrictions $0 \leq \alpha \leq 1$, $0 \leq \delta \leq 1$, $0 \leq \underline{t} \leq 1$, $0 \leq L \leq \frac{1}{2}$ and $0 \leq \sigma^2 \leq \frac{1}{4}$, see sections 15 and 16 of the code. Mathematica returns the value "false".

We obtain the equilibrium profits by inserting the equilibrium locations x_i^* for $i = 1, 2$ into (A.1) and (A.2). After several steps of algebra, we obtain

$$
\mathbb{CEU}[\Pi_i] = -\frac{\left(\delta\left(-(\alpha-1)(2L+3)^2\overline{t}+\alpha(3-2L)^2\underline{t}-4\sigma^2-9\right)+4\sigma^2+9\right)^2}{36(\delta((\alpha-1)(2L+3)\overline{t}+\alpha(2L-3)\underline{t}+3)-3)}.
$$

The competitive differentiation is given by

$$
\begin{aligned}
\Delta_x^* &= x_2^* - x_1^* = 2x_2^*\\
&= \frac{\delta\left((\alpha-1)(2L+3)^2\overline{t}-\alpha(3-2L)^2\underline{t}+4\sigma^2+9\right)-4\sigma^2-9}{2(\delta((\alpha-1)(2L+3)\overline{t}+\alpha(2L-3)\underline{t}+3)-3)}.
\end{aligned}
$$

□

Before starting with the proofs of the comparative static analysis, we want to point out that for many of the estimations performed in the subsequent five proofs, we make use of the following intrinsic parameter restrictions:

- upper and lower support boundaries for M: $0 < L \leq \frac{1}{2}$
- upper and lower bound of the confidence parameter: $0 \leq \delta \leq 1$

- upper and lower bound of ambiguity attitude: $0 \leq \alpha \leq 1$
- upper and lower bound of the variance of M: $0 \leq \sigma^2 \leq L^2 \leq \frac{1}{4}$
- upper and lower bound of the transportation cost parameter: $0 < \underline{t} \leq 1 \leq \overline{t}$

Proof of Proposition 2.2. The derivative of x_1^* with respect to α is given by

$$\frac{\partial x_1^*}{\partial \alpha} = -\frac{\delta(2L-3)\underline{t}(2\delta L(2L+3)\overline{t} - (\delta-1)(3L+2\sigma^2)) + (\delta-1)\delta(2L+3)\overline{t}(3L-2\sigma^2)}{2(\delta((\alpha-1)(2L+3)\overline{t} + \alpha(2L-3)\underline{t}+3)-3)^2}$$

The denominator is positive. Therefore, the sign of the derivative is determined by its numerator. We analyze the sign of this expression in two steps. The first part of the numerator is

$$g_1 := -\delta(2L-3)\underline{t}(2\delta L(2L+3)\overline{t} + (1-\delta)(3L+2\sigma^2))$$

Due to the fact that $L < \frac{1}{2}$, one can infer that $2L-3 < 0$. Hence, one obtains $g_1 > 0$. The second part of the numerator is

$$g_2 := (1-\delta)\delta(2L+3)\overline{t}(3L-2\sigma^2)$$

Since $\sigma^2 < L^2 < L$, one can infer that

$$3L - 2\sigma^2 > 3L - 2L = L > 0.$$

Therefore, one has $g_2 > 0$ as well. This proves that $\frac{\partial x_1^*}{\partial \alpha} > 0$ and $\frac{\partial x_2^*}{\partial \alpha} = -\frac{\partial x_1^*}{\partial \alpha} < 0$. $\square$

Proof of Proposition 2.3. The derivative of x_1^* with respect to σ^2 is given by

$$\frac{\partial x_1^*}{\partial \sigma^2} = \frac{1-\delta}{\delta((\alpha-1)(2L+3)\overline{t} + \alpha(2L-3)\underline{t}+3)-3}$$

The numerator is non-negative since $1-\delta \geq 0$ for $0 \leq \delta \geq 1$. It is strictly positive for $0 \leq \delta < 1$. For the denominator, observe that $\delta(\alpha-1)(2L+3)\bar{t} \leq 0$ and $\delta(\alpha(2L-3)\underline{t}+3) \leq 0$, since $L < \frac{1}{2}$. Hence, the denominator is smaller or equal -3 and therefore negative. Thus, $\frac{\partial x_1^*}{\partial \sigma^2} \leq 0$ and $\frac{\partial x_2^*}{\partial \sigma^2} = -\frac{\partial x_1^*}{\partial \sigma^2} \geq 0$. For $\delta = 1$ both x_1^* and x_2^* are independent of σ^2. Therefore $\frac{\partial x_2^*}{\partial \sigma^2} = \frac{\partial x_1^*}{\partial \sigma^2} = 0$. □

Proof of Proposition 2.4. We have

$$\frac{\partial x_1^*}{\partial \underline{t}} = \frac{\alpha\delta\left(2(\alpha-1)\delta L\left(4L^2-9\right)\bar{t} + (\delta-1)(2L-3)(3L+2\sigma^2)\right)}{2(\delta((\alpha-1)(2L+3)\bar{t} + \alpha(2L-3)\underline{t}+3)-3)^2}$$

It is obvious that the denominator is positive. Turning to the numerator, we can see that $4L^2 - 9 \leq 0$ since $L < \frac{1}{2}$. Therefore, we can conclude that

$$2\alpha\delta(\alpha-1)\delta L\left(4L^2-9\right)\bar{t} \geq 0.$$

Similarly, since $2L - 3 < 0$ and $\alpha - 1 \leq 0$, one can infer that

$$\alpha\delta(\delta-1)(2L-3)(3L+2\sigma^2) \geq 0.$$

Consequently, the numerator is positive and $\frac{\partial x_1^*}{\partial \underline{t}} > 0$. Since $x_2^* = -x_1^*$, it follows that $\frac{\partial x_2^*}{\partial \underline{t}} = -\frac{\partial x_1^*}{\partial \underline{t}} < 0$. The derivative of x_1^* with respect to $\bar{t}$ is given by

$$\frac{\partial x_1^*}{\partial \bar{t}} = -\frac{(\alpha-1)\delta(2L+3)(2\alpha\delta L(2L-3)\underline{t} + (\delta-1)(3L-2\sigma^2))}{2(\delta((\alpha-1)(2L+3)\bar{t} + \alpha(2L-3)\underline{t}+3)-3)^2}$$

Clearly, the denominator is positive. Turning to the numerator, observe that the factor $-(\alpha-1)\delta(2L+3)$ is positive. Moreover, since $L < \frac{1}{2}$, one can infer

$$2\alpha\delta L(2L-3)\underline{t} \leq 0.$$

As a next step, we can show that

$$3L - 2\sigma^2 \geq 3L - 2L^2 > 3L - 2L = L > 0.$$

This implies $(\delta - 1)(3L - 2\sigma^2)) \leq 0$. In total, the numerator is negative. Therefore, one has $\frac{\partial x_1^*}{\partial \underline{t}} < 0$ and $\frac{\partial x_2^*}{\partial \underline{t}} > 0$. □

Proof of Proposition 2.5. The derivative of x_1^* with respect to δ is given by

$$\frac{\partial x_1^*}{\partial \delta} = \frac{(\alpha - 1)(2L + 3)\bar{t}(3L - 2\sigma^2) - \alpha(2L - 3)\underline{t}(3L + 2\sigma^2)}{2(\delta((\alpha - 1)(2L + 3)\bar{t} + \alpha(2L - 3)\underline{t} + 3) - 3)^2}$$

It is straightforward to see that the denominator is positive. The first part of the numerator is given by

$$g_3 := (\alpha - 1)(2L + 3)\bar{t}(3L - 2\sigma^2).$$

Since

$$3L - 2\sigma^2 \geq 3L - 2L = L > 0,$$

one can infer that $g_3 \leq 0$. Defining

$$g_4 := -\alpha(2L - 3)\underline{t}(3L + 2\sigma^2),$$

one obtains by $2L - 3 < 0$ that $g_4 \geq 0$. As a consequence, one can infer that $\frac{\partial x_1^*}{\partial \delta} > 0$ if $g_3 + g_4 > 0$ and $\frac{\partial x_1^*}{\partial \delta} < 0$ if $g_3 + g_4 < 0$. Moreover, one has $\frac{\partial x_1^*}{\partial \delta} = 0$ if $g_3 + g_4 = 0$. Solving the equation $-g_3 = g_4$ for α, one obtains the unique solution

$$\alpha^* := \frac{(2L + 3)(3L - 2\sigma^2)}{(2L + 3)\bar{t}(3L - 2\sigma^2) - (2L - 3)\underline{t}(3L + 2\sigma^2)}$$

Besides, one can see that $g_3 + g_4 > 0$ whenever $\alpha > \alpha^*$ and $g_3 + g_4 < 0$ whenever $\alpha < \alpha^*$. This establishes that the numerator has, for every parameter constellation, a unique zero α^* where $\frac{\partial x_1^*}{\partial \delta} < 0$ for all $0 \leq \alpha < \alpha^*$, $\frac{\partial x_1^*}{\partial \delta} = 0$ for $\alpha = \alpha^*$ and $\frac{\partial x_1^*}{\partial \delta} > 0$ for all $1 \geq \alpha > \alpha^*$.

Since $x_2^* = -x_1^*$, we obtain the postulated result for x_2^* without reexamining the respective derivative. □

Proof of Proposition 2.6. The derivative of ${x_1}^*$ with respect to L is given by

$$\begin{aligned}\frac{\partial x_1^*}{\partial L} =& -\delta((\alpha-1)\bar{t}((\delta-1)(12L-4\sigma^2+9)-24\alpha\delta L\underline{t})\\ &+\alpha\underline{t}(-\alpha\delta(3-2L)^2\underline{t}-(\delta-1)(12L+4\sigma^2-9))+(\alpha-1)^2\delta(2L+3)^2\bar{t}^2)\\ &\cdot(2(\delta((\alpha-1)(2L+3)\bar{t}+\alpha(2L-3)\underline{t}+3)-3)^2)^{-1}\end{aligned}$$

As we can see, the denominator is positive. Therefore the sign of the derivative solely depends on the numerator. Since $\delta \geq 0$ it is sufficient to consider the sign of numerator divided by δ. We denote this expression with $(*)$. Inserting $\alpha = 0$ into expression $(*)$ yields

$$\begin{aligned}&\bar{t}\left((\delta-1)(12L-4\sigma^2+9)-\delta(2L+3)^2\bar{t}\right)\\ &\leq \delta[-12L+4\sigma^2-9]\\ &\leq \delta[-12L-8]\\ &= -4\delta(3L+2)\\ &< 0\end{aligned}$$

This shows that the derivative is strictly negative for $\alpha = 0$. Similarly, inserting $\alpha = 1$ into $(*)$, we obtain

$$\underline{t}\left((\delta-1)(12L+4\sigma^2-9)+\delta(3-2L)^2\underline{t}\right) \tag{A.9}$$

We establish that expression (A.9) is strictly positive. It is

$$12L+4\sigma^2-9 \leq 6+1-9 = -2 < 0.$$

As a consequence, we obtain

$$(\delta - 1)(12L + 4\sigma^2 - 9) \geq 0.$$

This shows that the numerator is positive. Now, we demonstrated that $\frac{\partial x_1^*}{\partial L} < 0$ for $\alpha = 0$, and $\frac{\partial x_1^*}{\partial L} > 0$ for $\alpha = 1$. The derivative is continuous. By the intermediate value theorem for continuous functions, we obtain that there is $\hat{\alpha} \in (0, 1)$ such that $\frac{\partial x_1^*}{\partial L} = 0$ for $\alpha = \hat{\alpha}$. What remains to be shown is that $\hat{\alpha}$ is unique. In this case, we know that x_1^* is strictly decreasing in L for values of α smaller that $\hat{\alpha}$, constant for $\alpha = \hat{\alpha}$, and increasing for $1 \geq \alpha > \hat{\alpha}$. Solving expression $(*)$ for α, we know that we find at least one zero, the zero $\hat{\alpha}$ in the interval $[0, 1]$. Since $(*)$ is a quadratic function in α, we can conclude that it has one more root $\hat{\hat{\alpha}}$. This root cannot be located in the interval $[0, 1]$ as well. This we show by making use of a proof by contradiction. Assume, w.l.o.g., that $\hat{\hat{\alpha}}$ was in the interval $[0, 1]$ as well and that $\hat{\alpha} < \hat{\hat{\alpha}}$. We can distinguish two cases. Case 1 is that the quadratic function has a global maximum, and case 2 is that the quadratic function has a global minimum. Since we can find both roots in the interval $[0, 1]$, the global maximum, or alternatively the global minimum, are also located in this interval. Assume now that we have a quadratic function with a global maximum. In this case, we have that $(*)$ is smaller zero for $\alpha < \hat{\alpha}$, equal to zero for $\alpha \in \{\hat{\alpha}, \hat{\hat{\alpha}}\}$, and smaller zero for $\alpha \in (\hat{\hat{\alpha}}, 1]$. The last statement contradicts that $(*)$ is larger zero for $\alpha = 1$ what we already showed above. For a global minimum a similar line of arguments holds. Since both roots are located in the interval $[0, 1]$, we can deduce that the minimum is located in this interval as well. In this case we can conclude that $(*)$ is larger than zero for $\alpha < \hat{\alpha}$, equal to zero for $\alpha \in \{\hat{\alpha}, \hat{\hat{\alpha}}\}$, and again larger zero for $\alpha \in (\hat{\hat{\alpha}}, 1]$. The first statement contradicts that $(*)$ is smaller zero for $\alpha = 0$. To sum up, we have only one root in $[0, 1]$. $\square$

```
"1. Define Objectives for Firm 1 and Firm 2";

f1[x1_, x2_, alpha_, delta_, tlow_, thigh_, sigma_, L_] :=
  delta * ((2 * thigh * (1 - alpha) * (x2 - x1) * (L + (x2 + x1) / 2 + 3 / 2) ^2) / 9 +
      (2 * alpha * tlow * (x2 - x1) * (-L + (x2 + x1) / 2 + 3 / 2) ^2) / 9) +
   ((1 - delta) * (x2 - x1) * ((x2 + x1 + 3) ^2 + 4 * sigma)) / 18;


f2[x1_, x2_, alpha_, delta_, tlow_, thigh_, sigma_, L_] :=
  delta * ((2 * alpha * tlow * (x2 - x1) * (L + (x2 + x1) / 2 - 3 / 2) ^2) / 9 +
      (2 * (1 - alpha) * thigh * (x2 - x1) * (-L + (x2 + x1) / 2 - 3 / 2) ^2) / 9) +
   ((1 - delta) * (x2 - x1) * ((x2 + x1 - 3) ^2 + 4 * sigma)) / 18;


"2. Introduce Parameter Restrictions";

assumptions = And[L == 1 / 2, 0 ≤ alpha ≤ 1,
   0 < tlow <= 1, 0 <= delta ≤ 1, 0 <= sigma ≤ L^2, thigh ≥ 1];

"3. Define the Midpoint Between Both Firms";

mean = (x1 + x2) / 2;

"4. Solving for Mutual Best Responses";

solutions =
  FullSimplify[Solve[{D[f1[x1, x2, alpha, delta, tlow, thigh, sigma, L], x1] == 0,
     D[f2[x1, x2, alpha, delta, tlow, thigh, sigma, L], x2] == 0}, {x1, x2}]];

"5. Store Solutions in a Table";

TableForm[Table[{solutions[[i, 1, 2]], solutions[[i, 2, 2]]},
   {i, Length[solutions]}], TableHeadings → {{"1", "2", "3"}, {"x1", "x2"}},
  TableAlignments → Center, TableSpacing → {3, 4}];

"6. Verify Whether Solution
   Satisfies the Global Competition Condition";
```

```
Table[{i, FullSimplify[(And[-3 / 2 < -L - mean, L - mean < 3 / 2] /. solutions[[i]]),
   assumptions]}, {i, Length[solutions]}]
```

```
{{1, True}, {2, 2 + 2 delta (-1 + thigh - alpha thigh + alpha tlow) >
    √(9 - 4 sigma + delta (-18 - 23 (-1 + alpha) thigh + 11 alpha tlow +
          4 sigma (2 + (-1 + alpha) thigh - alpha tlow) + delta (9 + 16 (-1 + alpha)^2
              thigh^2 + 4 sigma (-1 + thigh - alpha thigh + alpha tlow) + alpha tlow
              (-11 + 4 alpha tlow) - (-1 + alpha) thigh (-23 + 12 alpha tlow)))) &&
   2 delta (1 + (-1 + alpha) thigh - alpha tlow) <
    2 +
     √(9 - 4 sigma + delta (-18 - 23 (-1 + alpha) thigh + 11 alpha tlow +
           4 sigma (2 + (-1 + alpha) thigh - alpha tlow) + delta (9 + 16 (-1 + alpha)^2
               thigh^2 + 4 sigma (-1 + thigh - alpha thigh + alpha tlow) + alpha tlow
               (-11 + 4 alpha tlow) - (-1 + alpha) thigh (-23 + 12 alpha tlow))))},
 {3, 2 delta (1 + (-1 + alpha) thigh - alpha tlow) <
    2 +
     √(9 - 4 sigma + delta (-18 - 23 (-1 + alpha) thigh + 11 alpha tlow +
           4 sigma (2 + (-1 + alpha) thigh - alpha tlow) + delta (9 + 16 (-1 + alpha)^2
               thigh^2 + 4 sigma (-1 + thigh - alpha thigh + alpha tlow) + alpha tlow
               (-11 + 4 alpha tlow) - (-1 + alpha) thigh (-23 + 12 alpha tlow)))) &&
   2 delta (1 + (-1 + alpha) thigh - alpha tlow) + √(9 - 4 sigma +
        delta (-18 - 23 (-1 + alpha) thigh + 11 alpha tlow + 4 sigma
            (2 + (-1 + alpha) thigh - alpha tlow) + delta (9 + 16 (-1 + alpha)^2 thigh^2 +
              4 sigma (-1 + thigh - alpha thigh + alpha tlow) + alpha tlow (-11 +
                 4 alpha tlow) - (-1 + alpha) thigh (-23 + 12 alpha tlow)))) < 2}}
```

"7. Verify That the First Pair of Solutions Satisfies the Global Competition Condition";

```
x1st = (9 + 4 sigma +
      delta (-9 - 4 sigma - (-1 + alpha) (3 + 2 L)^2 thigh + alpha (3 - 2 L)^2 tlow)) /
   (4 (-3 + delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow)));

x2st = -x1st;

Simplify[x2st > 3 / 2 - L, assumptions]
```

```
3 delta + 4 sigma > 3 + 4 delta (sigma + alpha tlow)
```

```
Reduce[{3 delta + 4 sigma > 3 + 4 delta (sigma + alpha tlow), assumptions},
 {alpha, delta, sigma, tlow}]
```

```
False
```

"8. Define the Mean for the Second and Third Pair of Solutions";

```
mean2 =
  ((3 + delta (-3 - (-1 + alpha) (3 + 2 L) thigh + alpha (3 - 2 L) tlow) - √(9 - 4 sigma +
            delta (4 (-1 + alpha) ((-3 + L) L + sigma) thigh + 2 (-9 + 4 sigma -
                  9 (-1 + alpha) thigh) - 2 alpha (-9 + 2 L (3 + L) + 2 sigma) tlow +
               delta (9 + (-1 + alpha)^2 (3 + 2 L)^2 thigh^2 + 4 sigma
                    (-1 + thigh - alpha thigh + alpha tlow) + alpha tlow
                    (9 (-2 + alpha tlow) + 4 L (3 + L + alpha (-3 + L) tlow)) + 2 (-1 +
                      alpha) thigh (9 - 9 alpha tlow + 2 L (3 - L + 6 alpha L tlow)))))) /
       (2 (-1 + delta (1 + (-1 + alpha) thigh - alpha tlow))) -
      ((3 - delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow) +
          √(9 - 4 sigma + delta (4 (-1 + alpha) ((-3 + L) L + sigma)
                  thigh + 2 (-9 + 4 sigma - 9 (-1 + alpha) thigh) - 2 alpha
                  (-9 + 2 L (3 + L) + 2 sigma) tlow + delta (9 + (-1 + alpha)^2 (3 + 2 L)^2
                     thigh^2 + 4 sigma (-1 + thigh - alpha thigh + alpha tlow) + alpha
                     tlow (9 (-2 + alpha tlow) + 4 L (3 + L + alpha (-3 + L) tlow)) +
                    2 (-1 + alpha) thigh (9 - 9 alpha tlow + 2 L
                         (3 - L + 6 alpha L tlow)))))) /
        (2 (-1 + delta (1 + (-1 + alpha) thigh - alpha tlow)))) / 2;

mean3 =
  ((3 + delta (-3 - (-1 + alpha) (3 + 2 L) thigh + alpha (3 - 2 L) tlow) + √(9 - 4 sigma +
            delta (4 (-1 + alpha) ((-3 + L) L + sigma) thigh + 2 (-9 + 4 sigma -
                  9 (-1 + alpha) thigh) - 2 alpha (-9 + 2 L (3 + L) + 2 sigma) tlow +
               delta (9 + (-1 + alpha)^2 (3 + 2 L)^2 thigh^2 + 4 sigma
                    (-1 + thigh - alpha thigh + alpha tlow) + alpha tlow
                    (9 (-2 + alpha tlow) + 4 L (3 + L + alpha (-3 + L) tlow)) + 2 (-1 +
                      alpha) thigh (9 - 9 alpha tlow + 2 L (3 - L + 6 alpha L tlow)))))) /
       (2 (-1 + delta (1 + (-1 + alpha) thigh - alpha tlow))) +
      (-3 + delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow) +
         √(9 - 4 sigma + delta (4 (-1 + alpha) ((-3 + L) L + sigma) thigh +
               2 (-9 + 4 sigma - 9 (-1 + alpha) thigh) - 2 alpha (-9 + 2 L (3 + L) +
                  2 sigma) tlow + delta (9 + (-1 + alpha)^2 (3 + 2 L)^2 thigh^2 +
                  4 sigma (-1 + thigh - alpha thigh + alpha tlow) + alpha tlow
                   (9 (-2 + alpha tlow) + 4 L (3 + L + alpha (-3 + L) tlow)) + 2 (-1 +
                      alpha) thigh (9 - 9 alpha tlow + 2 L (3 - L + 6 alpha L tlow)))))) /
       (2 (-1 + delta (1 + (-1 + alpha) thigh - alpha tlow))) / 2;
```

"9. Determine the Mean's Range for the Second and Third Pair of Solutions";

```
NMinimize[{mean2, 0 <= L ≤ 1 / 2, 0 ≤ alpha ≤ 1, 0 <= tlow <= 1, 0 <= delta ≤ 1,
  0 <= sigma ≤ L^2, thigh ≥ 1}, {alpha, delta, tlow, thigh, sigma, L}]
{1., {alpha → 1., delta → 1., tlow → 0.878054, thigh → 1., sigma → 0.139562, L → 0.5}}

NMaximize[{mean2, 0 <= L ≤ 1 / 2, 0 ≤ alpha ≤ 1, 0 <= tlow <= 1, 0 <= delta ≤ 1,
  0 <= sigma ≤ L^2, thigh ≥ 1}, {alpha, delta, tlow, thigh, sigma, L}]
{2., {alpha → 2.6438 × 10^-9, delta → 1.,
  tlow → 0.0272573, thigh → 2.23364, sigma → 0.234253, L → 0.5}}
```

```
NMinimize[{mean3, 0 <= L ≤ 1 / 2, 0 ≤ alpha ≤ 1, 0 <= tlow <= 1, 0 <= delta ≤ 1,
  0 <= sigma ≤ L^2, thigh ≥ 1}, {alpha, delta, tlow, thigh, sigma, L}]
```

$\{-2., \{\text{alpha} \to 2.6438 \times 10^{-9}, \text{delta} \to 1., \text{tlow} \to 0.0272573, \text{thigh} \to 2.23364, \text{sigma} \to 0.234253, \text{L} \to 0.5\}\}$

```
NMaximize[{mean3, 0 <= L ≤ 1 / 2, 0 ≤ alpha ≤ 1, 0 <= tlow <= 1, 0 <= delta ≤ 1,
  0 <= sigma ≤ L^2, thigh ≥ 1}, {alpha, delta, tlow, thigh, sigma, L}]
```

$\{-1., \{\text{alpha} \to 1., \text{delta} \to 1., \text{tlow} \to 0.365717, \text{thigh} \to 1.79008, \text{sigma} \to 1.80912 \times 10^{-22}, \text{L} \to 0.5\}\}$

"10. Second-Order Derivative Firm 1";

```
FullSimplify[ D[f1[x1, x2, alpha, delta, tlow, thigh, sigma, L], {x1, 2}]]
```

$\frac{1}{9} (-6 - 3\,x1 - x2 + \text{delta}\,(6 + 3\,x1 + \text{alpha tlow}\,(-6 + 4\,L - 3\,x1 - x2) + x2 + (-1 + \text{alpha})\,\text{thigh}\,(6 + 4\,L + 3\,x1 + x2)))$

```
Secondorderderivative1[x1_, x2_, alpha_, delta_, tlow_, thigh_, sigma_, L_] :=
  1/9 (- 6 - 3 x1 - x2 + delta (6 + 3 x1 + alpha tlow (- 6 + 4 L - 3 x1 - x2) +
        x2 + (-1 + alpha) thigh (6 + 4 L + 3 x1 + x2)));
```

"11. Second-Order Derivative Firm 2";

```
FullSimplify[D[f2[x1, x2, alpha, delta, tlow, thigh, sigma, L], {x2, 2}]]
```

$\frac{1}{9} (-6 + x1 + 3\,x2 + \text{delta}\,(6 - x1 + (-1 + \text{alpha})\,\text{thigh}\,(6 + 4\,L - x1 - 3\,x2) - 3\,x2 + \text{alpha tlow}\,(-6 + 4\,L + x1 + 3\,x2)))$

```
Secondorderderivative2[x1_, x2_, alpha_, delta_, tlow_, thigh_, sigma_, L_] :=
 1/9 (- 6 + x1 + 3 x2 + delta (6 - x1 + (-1 + alpha) thigh (6 + 4 L - x1 - 3 x2) -
       3 x2 + alpha tlow (- 6 + 4 L + x1 + 3 x2)))
```

"12. Second-Order Derivatives Evaluated at Equilibrium Candidate Positions";

```
FullSimplify[Secondorderderivative1[(9 + 4 sigma +
      delta (-9 - 4 sigma - (-1 + alpha) (3 + 2 L)^2 thigh + alpha (3 - 2 L)^2 tlow)) /
    (4 (-3 + delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow))), (-9 -
      4 sigma + delta (9 + 4 sigma + (-1 + alpha) (3 + 2 L)^2 thigh - alpha (3 - 2 L)^2 tlow)) /
    (4 (-3 + delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow))),
  alpha, delta, tlow, thigh, sigma, L]]
```

```
(27 - 4 sigma + delta
    (8 sigma + 4 (-1 + alpha) ((-9 + L) L + sigma) thigh - 54 (1 + (-1 + alpha) thigh) -
      2 alpha (-27 + 2 L (9 + L) + 2 sigma) tlow + delta (27 + 3 (-1 + alpha)^2
          (3 + 2 L)^2 thigh^2 + 4 sigma (-1 + thigh - alpha thigh + alpha tlow) +
         alpha tlow (-54 + 4 L (9 + L) + 3 alpha (3 - 2 L)^2 tlow) +
         2 (-1 + alpha) thigh (27 - 27 alpha tlow + 2 L (9 - L + 10 alpha L tlow))))) /
 (18 (-3 + delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow)))
```

```
FullSimplify[Secondorderderivative2[(9 + 4 sigma +
      delta (-9 - 4 sigma - (-1 + alpha) (3 + 2 L)^2 thigh + alpha (3 - 2 L)^2 tlow)) /
    (4 (-3 + delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow))), (-9 -
      4 sigma + delta (9 + 4 sigma + (-1 + alpha) (3 + 2 L)^2 thigh - alpha (3 - 2 L)^2 tlow)) /
    (4 (-3 + delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow))),
  alpha, delta, tlow, thigh, sigma, L]]
```

```
(27 - 4 sigma + delta
    (8 sigma + 4 (-1 + alpha) ((-9 + L) L + sigma) thigh - 54 (1 + (-1 + alpha) thigh) -
      2 alpha (-27 + 2 L (9 + L) + 2 sigma) tlow + delta (27 + 3 (-1 + alpha)^2
          (3 + 2 L)^2 thigh^2 + 4 sigma (-1 + thigh - alpha thigh + alpha tlow) +
         alpha tlow (-54 + 4 L (9 + L) + 3 alpha (3 - 2 L)^2 tlow) +
         2 (-1 + alpha) thigh (27 - 27 alpha tlow + 2 L (9 - L + 10 alpha L tlow))))) /
 (18 (-3 + delta (3 + (-1 + alpha) (3 + 2 L) thigh + alpha (-3 + 2 L) tlow)))
```

"13. Take the Derivative of the Numerator With Respect to Thigh";

```
FullSimplify[
 D[(27 - 4 sigma + delta (8 sigma + 4 (-1 + alpha) ((-9 + L) L + sigma) thigh - 54
         (1 + (-1 + alpha) thigh) - 2 alpha (-27 + 2 L (9 + L) + 2 sigma) tlow +
       delta (27 + 3 (-1 + alpha)^2 (3 + 2 L)^2 thigh^2 +
          4 sigma (-1 + thigh - alpha thigh + alpha tlow) +
          alpha tlow (-54 + 4 L (9 + L) + 3 alpha (3 - 2 L)^2 tlow) + 2 (-1 + alpha)
           thigh (27 - 27 alpha tlow + 2 L (9 - L + 10 alpha L tlow))))), thigh]]
```

```
2 (-1 + alpha) delta
 (-27 + 2 ((-9 + L) L + sigma) + delta (27 - 2 sigma + 3 (-1 + alpha) (3 + 2 L)^2 thigh -
      27 alpha tlow + 2 L (9 + L (-1 + 10 alpha tlow))))
```

"14. Check Whether the Derivative Can be Negative";

```
assumptions = And[0 <= L ≤ 1 / 2, 0 ≤ alpha ≤ 1,
   0 < tlow ≤ 1, 0 <= delta ≤ 1, 0 <= sigma ≤ L^2, thigh >= 1];
```

```
Reduce[
 {2 (-1 + alpha) delta (-27 + 2 ((-9 + L) L + sigma) + delta (27 - 2 sigma + 3 (-1 + alpha)
        (3 + 2 L)^2 thigh - 27 alpha tlow + 2 L (9 + L (-1 + 10 alpha tlow)))) < 0,
  assumptions}, {alpha, delta, sigma, L, tlow, thigh}]
False
```

"15. Evaluate the Numerator of the Second-Order Derivative at Thigh=1";

```
Num[alpha_, delta_, tlow_, thigh_, sigma_, L_] :=
  (27 - 4 sigma + delta (8 sigma + 4 (-1 + alpha) ((-9 + L) L + sigma) thigh -
      54 (1 + (-1 + alpha) thigh) - 2 alpha (-27 + 2 L (9 + L) + 2 sigma) tlow +
      delta (27 + 3 (-1 + alpha)^2 (3 + 2 L)^2 thigh^2 + 4 sigma (-1 + thigh - alpha thigh +
           alpha tlow) + alpha tlow (-54 + 4 L (9 + L) + 3 alpha (3 - 2 L)^2 tlow) +
         2 (-1 + alpha) thigh (27 - 27 alpha tlow + 2 L (9 - L + 10 alpha L tlow)))));

FullSimplify[Num[alpha, delta, tlow, 1, sigma, L]]

27 - 4 sigma + delta (4 (L (9 + (-1 + 4 delta) L) + sigma) +
    alpha^2 delta (27 (-1 + tlow)^2 + 4 L^2 (3 + tlow) (1 + 3 tlow) - 36 L (-1 + tlow^2)) -
    2 alpha (- (27 + 2 (-1 + delta) sigma) (-1 + tlow) +
      18 L (1 + delta + tlow - delta tlow) + 2 L^2 (-1 + tlow + delta (7 + 9 tlow))))

27 - 4 sigma + delta (4 (L (9 + (-1 + 4 delta) L) + sigma) +
    alpha^2 delta (27 (-1 + tlow)^2 + 4 L^2 (3 + tlow) (1 + 3 tlow) - 36 L (-1 + tlow^2)) -
    2 alpha (- (27 + 2 (-1 + delta) sigma) (-1 + tlow) +
      18 L (1 + delta + tlow - delta tlow) + 2 L^2 (-1 + tlow + delta (7 + 9 tlow))))


27 - 4 sigma + delta (4 (L (9 + (-1 + 4 delta) L) + sigma) +
    alpha^2 delta (27 (-1 + tlow)^2 + 4 L^2 (3 + tlow) (1 + 3 tlow) - 36 L (-1 + tlow^2)) -
    2 alpha ((-27 - 2 (-1 + delta) sigma) (-1 + tlow) +
      18 L (1 + delta + tlow - delta tlow) + 2 L^2 (-1 + tlow + delta (7 + 9 tlow))))
```

```
"16. Can The Numerator
   Evaluated at Thigh=1 Become Negative?";

Reduce[{27 - 4 sigma + delta (4 (L (9 + (-1 + 4 delta) L) + sigma) +
        alpha^2 delta (27 (-1 + tlow)^2 + 4 L^2 (3 + tlow) (1 + 3 tlow) - 36 L (-1 + tlow^2)) -
        2 alpha (- (27 + 2 (-1 + delta) sigma) (-1 + tlow) +
           18 L (1 + delta + tlow - delta tlow) + 2 L^2 (-1 + tlow + delta (7 + 9 tlow)))) <
    0, 0 <= L ≤ 1 / 2, 0 ≤ alpha ≤ 1, 0 <= tlow ≤ 1, 0 <= delta ≤ 1,
  0 <= sigma ≤ L^2}, {alpha, delta,
  tlow,
  sigma,
  L}]

False
```

Chapter 3

Primary Prevention Under Ambiguity

3.1 Introduction

A substantial part of everyday medical decision-making is concerned with preventive care. The fundamental idea of prevention lies in the assumption that patients' behavior and commitments to preventive health care measures may actively influence their prospects of certain future health states. The literature differentiates between different concepts of prevention.[1] First of all, there is primary prevention, referring to situations before the incidence of disease. A potentially healthy agent can exert a distinct amount of effort that itself influences his or her probability of contracting an illness in the future. There are a multitude of preventive measures that fall into this category. Preventing obesity by doing regular exercise or following nutritional guidelines from health experts, such as limiting the daily amount of carbohydrates consumed, may significantly reduce the risk of acquiring diabetes or cardiovascular diseases.[2] Another example for primary preventive

[1] See for instance Kenkel [2000] or Etner and Jeleva [2013].

[2] For a detailed survey on the cost-effectiveness of primary preventive programs on diabetes and cardiovascular diseases read the report by Korczak et al. [2011].

measures are safety guidelines and schooling for workers exposed to certain health risks at the workplace; you may think of workers exposed to dangerous substances, machinery, or surroundings. In contrast to that, secondary and tertiary prevention is confined to scenarios after the occurrence of disease. Secondary prevention refers to measures such as cancer screenings where patients are unaware of their current health status. The term tertiary prevention applies after the disease has been diagnosed. One can think of preventive measures that facilitate patients' physical recovery, reduce their risk of relapse, or ameliorate their general state of health. Throughout this paper, I am going to focus on primary prevention and primary preventive programs.

An important observation related to primary preventive activities is that patients are indeed aware that the commitment to a specific preventive measure reduces their risk of contracting an illness. Imagine, for instance, that patients were asked whether they believe that regular exercise reduces their risk of contracting cardiovascular diseases; one would expect a large majority of patients to answer positively. If patients were, on the other hand, asked to quantify the impact of preventive effort on their individual disease probabilities, they would probably fail to give an accurate answer. Suppose, for instance, that patients were asked how strongly their risk of contracting a cardiovascular disease would decrease if they were engaged in one additional hour of sporting activities every week. In this case, one would expect that patients are either unable to provide an estimation, or come up with an estimation that they don't feel very confident with. This absence of knowledge can be explained by several possible reasons. First of all, patients need to incorporate and evaluate the impact of imprecisely known factors, such as their genetic predispositions to certain diseases, or lifestyle related factors[3], to form a probabilistic judgment. Secondly, even in the rather unrealistic situation that patients have direct access to recent scientific studies on the matter, the provision of these might be only of little help, since the findings of each survey are based on aggregate data for a certain sample of participants. Knowing this, patients might find it difficult to contrast aggregate results

[3]Examples for lifestyle related factors might be for instance nutrition, exposition to environmental risks, like pollution or hazardous substances, as well as stress.

with individual factors. Thirdly, patients might be confronted with a situation where scientific evidence on the respective programs is not available. This is, for instance, the case when patients participate in newly developed preventive programs where reliable data on its effectiveness is still absent. And finally, there might be confounding scientific or non-scientific evidence on the efficacy of a certain preventive regime. As an example for this, one might think of contradictory dietary recommendations, newly developed "wonder diets" promoted by some representatives of the pharmaceutical industry, or contradictory information arising from patients' eligibility for a second medical opinion. Consequently, patients ignore the true underlying relationship between preventive effort and disease probabilities. More importantly, they might consider a multitude of functional relationships between effort and disease probability possible. This is illustrated by means of the following stylized figure.

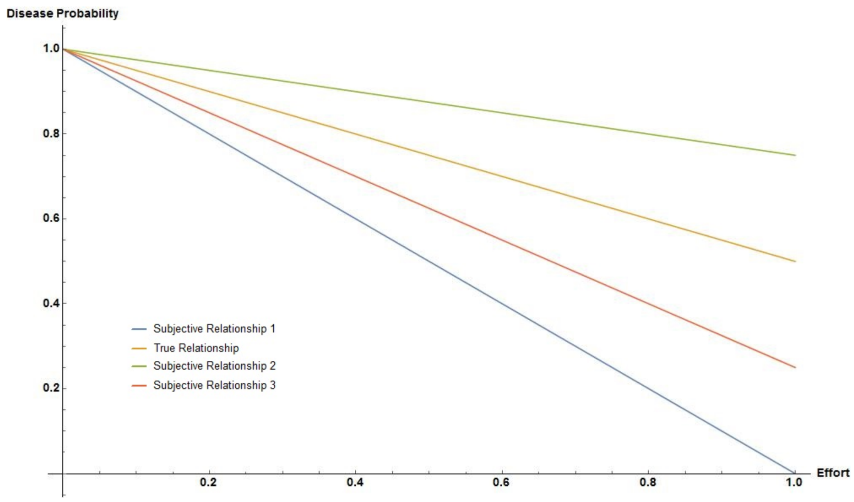


Figure 3.1: Preventive Relationships and the True Underlying Relationship

A key objective of any information campaign or health counseling on primary prevention is to ensure that patients are better informed about the effectiveness of a preventive regime. This objective is based on the premise that better informed agents are more likely to make better decisions in terms of their preventive activities. In my view, this assumption is highly problematic, since it is not clearly understood how patients process additional information in the light of imprecise a priori knowledge. In this paper, imprecise knowledge is modeled by "Knightian Uncertainty", or "Ambiguity", see Knight

[1921]. The idea that ambiguity is relevant for decision making on health matters is supported by a relatively new strand of empirical literature. Han et al. [2011] highlight the relevance of uncertainty for health care. The authors point out different sources and varieties of uncertainty in medical care and emphasize that, until now, clinical practice does not differentiate between different varieties of uncertainty as risk and ambiguity. One example why ambiguity might arise is confounding information about scientific evidence. Therefore, ambiguity might matter for communication schemes between physicians and patients. Han et al. [2007] consider the issue of conflicting information more concretely by empirically investigating the impact of contradictory mammography recommendations on women's behavior. The study finds that a higher degree of ambiguity yields a diminished uptake of mammography and lowers intentions for future mammography screenings. Similarly, Han et al. [2006] investigate the importance of ambiguity with respect to cancer preventability and cancer prevention recommendations. The authors find a positive correlation between ambiguity and perceived cancer risk or cancer worry. Furthermore, the study suggests that perceived ambiguity has a strong negative effect on cancer preventability. Politi et al. [2007] treat the question of how to communicate uncertainties about medical interventions to patients and state that further research is needed in order to fully understand how patients respond to risk and ambiguity.

This paper intends to study, on a theoretical basis, how additional information on a primary preventive regime impacts on patients' preventive activities when patients' prior knowledge is characterized by ambiguity. Note that the way information is processed under risk significantly differs from the way information is updated under ambiguity. As a consequence, it is not clearly understood how additional information on a preventive regime and patients' effort levels are interrelated under Knightian Uncertainty.

This paper intends to fill this gap by studying a model of physician counseling where patients with imprecise prior knowledge seek information from a physician on the relationship between effort and disease probabilities. Patients' imprecise a priori knowledge is thereby modeled by Knightian Uncertainty. More specifically, I assume that patients' preferences are of the Choquet-expected utility type. Beliefs are defined on a set of strictly ordered preventive relationships and represented by a neo-additive capacity, see Chateauneuf et al. [2007], which is a non-additive probability which allows researchers to model the magnitude of the imprecision of patients' beliefs as well as their attitudes towards ambiguity. The physician provides a random signal to patients. After receiving the signal, patients update their prior belief and exert effort. The optimal level of preventive activities is thereby determined according to a setting motivated by the self-insurance and self-protection model developed by Ehrlich and Becker [1972]. The following diagram illustrates the basic modeling framework.

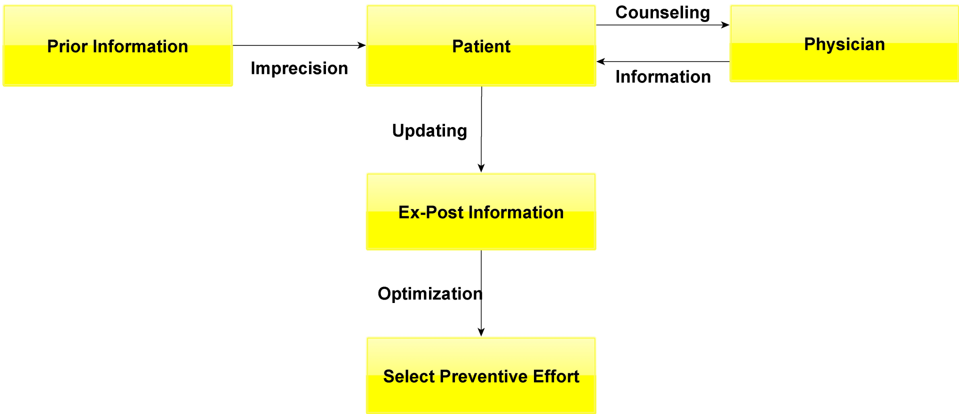


FIGURE 3.2: Basic Model Framework

Ehrlich and Becker [1972] consider a model with two states of the world, a "good" and a "bad" state, and a decision-maker that encounters a monetary loss in the bad state. Furthermore, the decision-maker can, by exerting a certain amount of effort, either reduce the amount of loss in the bad state (this is called "self-insurance"), or reduce the underlying probability of the bad state (this is called "self-protection"). To my knowledge, the following modifications and extensions of Ehrlich and Becker's model have been discussed in the health domain, and more specifically in the context of prevention. Eeckhoudt et al. [1998] relate Ehrlich and Becker's model to medical prevention. In particular, the authors introduce utilities depending on patients' health state and find that tertiary and secondary prevention are , whereas primary and tertiary prevention are complements. Eeckhoudt et al. [2001] investigate the link between primary and secondary prevention.

The paper demonstrates that policy-makers might reduce investment in primary preventive measures as soon as diagnostic tests become available. Hence, primary prevention and secondary prevention can be considered as substitutes. Zweifel et al. [2009] analyze the relationship between moral hazard, insurance, and prevention. Etner and Jeleva [2013] study the relationship between risk perception, prevention and diagnostic tests using the recursive rank-dependent utility model developed by Cohen et al. [2008]. The authors suggest a comprehensive framework incorporating primary and tertiary preventive activities, assuming that patients know the relationship between effort and the objective probability of disease, but might under- or overestimate this probability.

The idea to introduce Knightian uncertainty into Ehrlich and Becker's model has already been addressed by a number of papers. Snow [2011] incorporates ambiguity aversion by using the so-called KMM or Smooth Ambiguity model developed by Klibanoff et al. [2005]. The author introduces a model with two states of the world and concludes that, if decision-makers are risk-averse and ambiguity-averse at the same time, optimal self-insurance and self-protection increase with greater ambiguity aversion. Furthermore, the author states that "higher self-protection and self-insurance levels induce mean-preserving contractions in the distribution of expected utility that are valuable to ambiguity-averse decision-makers", see Snow [2010], page 39. Huang [2012] considers a self-insurance and self-protection model under ambiguity with KMM-preferences. The novelty of Huang's contribution consists in contemplating non-monetary costs of effort, higher order risk-preferences, and ambiguous target distributions assuming a wealth distribution defined on a compact support. The author concludes that ambiguity aversion entails higher effort levels whenever the individual can shift the initial wealth distribution towards a "preferred target distribution". Alary et al. [2013] examine a generalized version of the model in Snow [2011] with more than two states of the world and, using a willingness to pay approach, derive conditions under which ambiguity aversion increases the incentive to insure and self-insure but decreases the incentive to self-protect.[4] Robert and Therond [2014] use a theoretical approach by linking Yaari's dual approach, compare Yaari [1987],

[4] See Alary et al. [2013], page 18.

and ambiguity aversion with optimal prevention. The authors consider the so-called class of concave distortion risk measures and conclude that the willingness to pay for a risk reduction is always higher for a more ambiguity-averse decision-maker but not necessarily for a more risk-averse decision-maker.[5] Moreover, ambiguity-averse decision-makers exert less preventive effort when ambiguity refers to a less risky distribution.[5] Berger [2014] considers a two-period model with a recursive KMM approach,[6] where a decision-maker invests in prevention in the first period in order to improve the final wealth distribution or to influence the probability of being in an ambiguous state of the world in the second period.[7] The author concludes that the effect of ambiguity on self-protection cannot be signed. In order to give a sufficient condition for ambiguity to increase the demand for self-protection, Berger [2014] introduces a concept called "ambiguity prudence attitude", or "decreasing absolute ambiguity aversion", linking the problem of self-protection under ambiguity to the concept of prudence, which was until now only considered in the risk case.[8]

In contrast to the existing literature on self-protection, I study preventive behavior in the health domain, and more specifically in the context of primary prevention. The aim of this research is to analyze how learning affects self-protection when patients ignore the relationship between effort and disease probabilities. Until now, there is, to my knowledge, no article combining learning, Ehrlich and Becker's notion of self-protection, and primary prevention when decision-makers face Knightian Uncertainty.

This paper is organized as follows: In the next section, I give a detailed description of the model setup. The third section analyzes the impact of optimism and confidence on optimal self-protection. Section 4 examines the impact of additional information on preventive activities. Finally, section 5 describes my conclusions.

[5]Compare Robert and Therond [2014], page 11.

[6]See Klibanoff et al. [2009].

[7]Compare Berger [2014], page 4.

[8]Eeckhoudt and Gollier [2005] link prevention to prudence and find that prudence tends to reduce prevention.

3.2 Model

To begin with, suppose that there are two time points $t = 0, 1$ and a patient whose state-contingent Bernoulli utility $u : X \to \mathbb{R}$ is defined on the set $X = \mathbb{R} \times \Omega$. The state space $\Omega \subset \mathbb{R}$ consists of two states of the world, $\Omega = \{h_1, h_2\}$ where w.l.o.g. $h_1 \leq h_2$. The state h_1 refers to a situation where the patient contracts an illness, while h_2 refers to a situation where the patient remains healthy at $t = 1$. Each element $x \in X$ is a combination $x = (w, h_i)$ where w denotes the patient's wealth level and h_i denotes the patient's health status at time $t = 1$. As a next step, I make the following assumptions with respect to patients' utility function.[9]

Assumption 4. Patients' Bernoulli utility satisfies the following conditions:

(A_1) u is twice continuously differentiable

(A_2) u is strictly increasing in wealth, in formal terms $u_w > 0$

(A_3) u is concave with respect to w, formally $u_{ww} \leq 0$

Assumption (A_1) is a purely technical assumption. (A_2) asserts that patients prefer more money to less money in both health states. Requirement (A_3) is an assumption on risk-preferences and presumes that patients are either risk-averse or risk-neutral.
Patients' wealth in the bad state of the world is denoted by $W_1(V)$, and by $W_2(V)$ in the good state. Wealth is effort-dependent, with the following requirements.

Assumption 5. Patients' wealth functions satisfy the following conditions:

(W_1) Both wealth functions $W_i(V)$ are twice continuously differentiable in V.

(W_2) Wealth is decreasing in effort $W_i'(V) < 0$ for $i = 1, 2$.

[9]Please note that u_w denotes the partial derivative of u with respect to the wealth level w. Similarly, u_{ww} denotes the second order partial derivative of u with respect to w.

Requirement (W_1) is purely technical. Assumption (W_2) reflects the fact that effort is costly. Imagine for instance the case where a patient follows a regular workout program to prevent obesity. A higher level of effort can be interpreted as additional time spent in sporting activities every week. This is costly due to time spent, entry fees for health centers, additional transportation costs, expenses for sporting equipment, etc.

As a next step, I make the assumption that there is a twice continuously differentiable and convex function $\pi^{real} : [0,1] \to [0,1]$. For each effort level $V \in [0,1]$ the value $\pi^{real}(V)$ denotes the objective probability that the patient contracts an illness at time $t = 1$. Similarly, $1 - \pi^{real}(V)$ denotes the objective probability that the patient remains healthy at time $t = 1$. Hence, π^{real} describes the true underlying relationship between effort and disease probabilities. In this framework, π^{real} is unknown to patients for the reasons pointed out in the introductory section of this paper. The term "prevention" implies that effort alters the likelihood of an event. Therefore, I assume that preventive effort alters the true underlying patient-specific disease probability. More specifically, I assume that there is scientific evidence demonstrating that there is a positive relationship between effort and the probability of the "bad state", and that patients know that this relationship is positive, but cannot exactly quantify the impact of effort on that probability.[10] Formally, patients know that the function π^{real} is decreasing in V. Patients' a priori knowledge is modeled by a set of preventive relationships

$$\Phi \subset \Big\{\pi |\ \pi : [0,1] \to [0,1]\Big\}$$

with the following requirements:

Assumption 6. Patients' preventive relationships $\pi_i \in \Phi$ satisfy the following conditions:

(P_1) Φ is finite and contains exactly n elements

(P_2) π_i is twice continuously differentiable for $i = 1, ..., n$.

[10] Without this assumption, the term prevention would not be justified, since it is not clear whether the true underlying disease probability is positively affected by preventive activities.

(P_3) π_i is strictly decreasing $\pi_i'(V) < 0$

Requirement (P_1) says that patients consider n disease probability possible for a given effort level $V \in [0, 1]$. These are denoted by $\pi_i(V)$ for $i \in \{1, ..., n\}$. Assumption (P_2) is purely technical. (P_3) says that patients believe that a higher level of effort translates into lower disease probabilities.

Remark 3.1. Henceforth, I implicitly assume that Assumptions 1,2, and 3 are satisfied throughout this paper.

In the introductory section, I argued that patients face Knightian uncertainty with respect to the true relationship between effort and disease probabilities. Hence, patients consider each $\pi_i \in \Phi$ as a possible realization of a random variable with unknown distribution. This can be illustrated graphically by means of the so-called Machina triangle for a fixed level of effort V and the special case of three realizations where

$$\Phi(V) := \Big\{\pi_1(V), \pi_2(V), \pi_3(V)\Big\}.$$

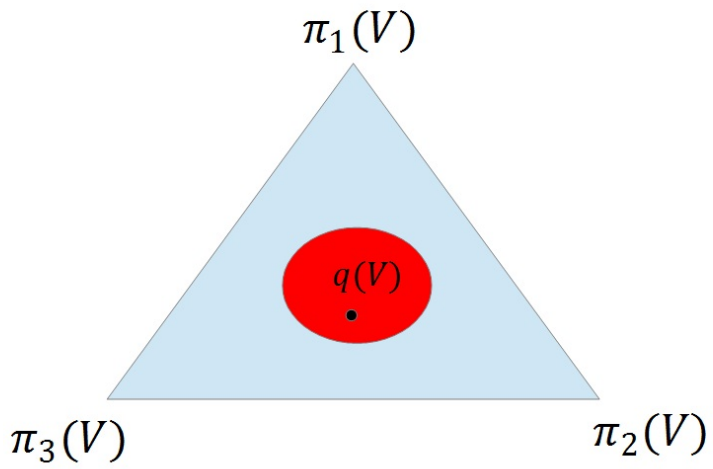


FIGURE 3.3: Patients' Beliefs in Case of Three Preventive Relationships

Each point in the triangle represents a specific belief $q(V)$. Formally $q(V)$ is an element of the simplex

$$\Delta(\Phi(V)) := \left\{ q(V) = (q_1(V), q_2(V), q_3(V)) \;\middle|\; q_i(V) \geq 0 \ \text{ and } \ \sum_{i=1}^{3} q_i(V) = 1 \right\}$$

where $q_i(V)$ denotes the probability that $\pi_i(V)$ is the true underlying disease probability. Under Knightian uncertainty, patients hold a subset of such beliefs. In the general case with n possible preventive relationships, a belief takes the form $q(V) = (q_1(V), ..., q_n(V))$. Throughout this paper, I assume that the belief $q(V)$ is independent of V. This implies that there is a belief $q = (q_1, ..., q_n) \in \Delta(\Phi)$ such that $q = q(V)$ for all $V \in [0, 1]$. This assumption is less restrictive than it may seem at first glance. Patients' effort is a choice variable. If there were instances such that $q(V_1) \neq q(V_2)$ for some $V_1, V_2 \in [0, 1]$, one could infer that patients' beliefs regarding which of the mechanisms $\pi_i \in \Phi$ describes the true relationship between effort and disease probability depend on their choice of effort. Such behavior seems implausible under the assumption that π^{real} is mostly based on external factors which are either fixed, such as patients' genetics, age, or gender, or not part of the preventive program, such as patients' job situation or place of residence.

One could argue that, even under Knightian Uncertainty, patients might still conform to Savage's [1954] subjective expected utility model.[11] Thus, a prevention model with purely subjective beliefs would be sufficient to account for the phenomenon of imprecise probabilistic knowledge. Contrary to this line of argument, Ellsberg's [1961] findings suggest that such an approach is severely problematic since a substantial share of decision-makers facing ambiguity does indeed display preferences that contradict the existence of a well-defined subjective belief. A decision-theoretic model which allows for deviations from SEU and accounts for Ellsberg's paradox is the so-called Choquet-expected utility model, pioneered by Schmeidler [1989]. Patients conforming to the Choquet model make decisions by maximizing a Choquet integral, which can be considered as a generalized

[11] An additional assumption to be made is that patients are able to reduce compound lotteries to simple lotteries. See for instance Segal [1990], page 353 for a formal description of the reduction of compound lotteries axiom.

expectation for non-additive probability measures. In this paper, patients' beliefs are represented by a special class of capacities, termed neo-additive capacities.[12] The following definition of a neo-additive capacity is adopted from Eichberger et al. [2009], page 359:

Definition 3.1. Let $q = (q_1, ..., q_n)$ be a probability measure on (Φ, Σ) where Σ denotes a σ-algebra of events on Φ. Then, for real numbers α and δ one can define a neo-additive capacity ν by $\nu(\varnothing) = 0$, $\nu(\Phi) = 1$, $\nu(A) = \delta\alpha + (1-\delta)q(A)$ where $A \in \Sigma$ is a nonempty and strict subset of Φ.

Subsequently, I presume that patients hold a neo-additive belief ν defined on Φ. In this case, each patient's objective function can be specified by evaluating a Choquet-integral[13] with respect to a neo-additive capacity. A functional representation[14] of such an integral is given by

$$\int_{\Phi} f d\nu = (1-\delta)E_q[f] + \delta\Big(\alpha \max\{x : f^{-1}(x) \notin \mathcal{N}\} + (1-\alpha)\min\{x : f^{-1}(x) \notin \mathcal{N}\}\Big) \tag{3.1}$$

where $f : \Phi \to \mathbb{R}$ is a simple function,[15] $\mathcal{N} = \{A \in \Sigma : \nu(A) = 0\}$ denotes the collection of null-events of the capacity ν, $E_q[f]$ denotes the expectation of f with respect to the probability distribution q, $\max\{x : f^{-1}(x) \notin \mathcal{N}\}$ is the set of those states of the world that induce the highest possible value of f, or the best case of f, under the assumption that each one of these states is not the realization of a null-event. Similarly, $\min\{x : f^{-1}(x) \notin \mathcal{N}\}$ denotes the set of those states of the world that induce the lowest possible value of f, or the worst case of f, under the assumption that each one of these states is not the realization of a null-event. This representation of the Choquet integral has an intuitive interpretation. A decision-maker with a neo-additive belief compares the expectation $E_q[f]$ with a combination of extreme outcomes, namely a convex combination of the best and worst case. The parameter δ is called the confidence parameter, and

[12]See Chateauneuf et al. [2007] for an axiomatization of neo-additive capacities.
[13]For more details, see Choquet [1955].
[14]See Lemma 3.1 by Chateauneuf et al. [2007], page 541.
[15]A simple function is measurable, real-valued function with a finite range, see Chateauneuf et al. [2007], page 540.

measures how strongly the decision-maker incorporates extreme cases into his evaluation. The parameter α is called the optimism parameter, and captures the magnitude to which decision-makers incorporate the worst case into the extreme-outcome part of their evaluation.

In order to derive a patient's objective function, one needs to answer the question of how the simple function f is defined in the context of health prevention. If patients knew the underlying preventive relationship π^{real}, the optimization problem would be given by

$$\max_{V\in[0,1]} \pi^{real}(V)u(W_1(V),h_1) + (1-\pi^{real}(V)u(W_2(V),h_2).$$

This is the standard expected utility case with effort-dependent disease probabilities and wealth functions. Under Knightian uncertainty, π^{real} is unknown. Hence, patients ignore whether they maximize an expected utility with respect to the relationship $\pi_i \in \Phi$, or whether they maximize an expected utility with respect to $\pi_j \in \Phi$ where $i, j \in \{1, ..., n\}$ and $i \neq j$. This observation provides information on how the simple function f needs to be defined. In the general case, f is a mapping from a state space $\tilde{\Omega}$ to a finite set of consequences $X \subset \mathbb{R}$. This means f assigns to every possible state of the world a resulting outcome for the decision-maker. In the context of the prevention model, the states are given by the preventive relationship $\pi_i \in \Phi$. The consequences are effort-dependent expected utilities

$$X(V) := \Big\{\pi_i(V)u(W_1(V),h_1) + (1-\pi_i(V))u(W_2(V),h_2) : 1 \leq i \leq n\Big\}.$$

Hence, one obtains for each fixed $V \in [0,1]$ the simple function

$$f(\pi_i|V) = \pi_i(V)u(W_1(V),h_1) + (1-\pi_i(V))u(W_2(V),h_2).$$

Making use of the functional representation (3.1), one obtains the objective function:

$$\mathcal{U}(V|\alpha,\delta) := \int_{\Phi} f(\pi_i|V)d\nu = (1-\delta)\ E_q[f(\cdot|V)] + \delta\Big\{\alpha Z_{max}(V) + (1-\alpha)Z_{min}(V)\Big\} \quad (3.2)$$

where

$$\begin{aligned} Z_{max}(V) &:= \max_{\pi_i\in\Phi}\ \Big\{\pi_i(V)u(W_1(V),h_1) + (1-\pi_i(V))u(W_2(V),h_2)\Big\} \\ Z_{min}(V) &:= \min_{\pi_i\in\Phi}\ \Big\{\pi_i(V)\ u(W_1(V),h_1) + (1-\pi_i(V))\ u(W_2(V),h_2)\Big\} \end{aligned} \quad (3.3)$$

and $E_q[f(\cdot|V)]$ denotes the expectation

$$\sum_{i=1}^{n} q_i f(\pi_i\ |V).$$

The patient's optimization problem is given by

$$\max_{V\in[0,1]} \mathcal{U}(V|\alpha,\delta).$$

The following corollary gives an alternative representation of the objective function $\mathcal{U}$.

Corollary 3.1. The objective function can be expressed in the form

$$\mathcal{U}(V|\alpha,\delta) = \pi_{\mathbb{CEU}}(V|\alpha,\delta)\ u(W_1(V),h_1) + (1-\pi_{\mathbb{CEU}}(V|\alpha,\delta))\ u(W_2(V),h_2) \quad (3.4)$$

where

$$\pi_{\mathbb{CEU}}(V|\alpha,\delta) = (1-\delta)\pi_q(V) + \delta(\alpha\pi_{max}(V) + (1-\alpha)\pi_{min}(V)) \quad (3.5)$$

and

$$\begin{aligned} \pi_q(V) &:= \sum_{i=1}^{n} q_i\pi_i(V), \\ \pi_{max}(V) &:= \arg\max_{\pi_i\in\Phi}\ \Big\{\pi_i(V)u(W_1(V),h_1) + (1-\pi_i(V))u(W_2(V),h_2)\Big\}, \\ \pi_{min}(V) &:= \arg\min_{\pi_i\in\Phi}\ \Big\{\pi_i(V)\ u(W_1(V),h_1) + (1-\pi_i(V))\ u(W_2(V),h_2)\Big\}. \end{aligned}$$

Proof. The proof is contained in the appendix.

Corollary 3.1 says that patients' objectives under Knightian uncertainty can be expressed in expected utility form with respect to the distorted probability $\pi_{\mathbb{CEU}}$. The distortion itself is a convex combination of an "expected probability" and a combination of worst and best-case probabilities. The following proposition gives an important technical property of the objective function and the existence of a solution for the patient's optimization problem.

Proposition 3.1. The patient's objective is continuous and the underlying optimization problem has a solution.

Proof. The proof is contained in the appendix.

The objective is continuous but not necessarily continuously differentiable. This can be seen by means of Example 3.5, which is contained in the appendix. The reason why the objective is not differentiable in the previous example lies in the fact that the minimizing preventive relationship changes from one preventive relationship to another preventive relationship at a point $\hat{V} \in (0, 1)$. Such a change of the minimizer (or maximizer) can occur when there are two preventive relationships $\hat{\pi}_1, \hat{\pi}_2 \in \Phi$ such that $\hat{\pi}_1$ crosses $\hat{\pi}_2$ from above, or from below, at some point $\hat{V} \in (0, 1)$. In this case, one calls $\hat{V}$ a crossing point. Subsequently, I give a formal definition of the term crossing point.

Definition 3.2. Consider two real-valued functions $f_i : D \to \mathbb{R}$ for $i = 1, 2$ where $D \subseteq \mathbb{R}^n$. A point $\hat{x} \in D$ is called a crossing point if there is $\delta > 0$ such that $f_1(\hat{x}) = f_2(\hat{x})$, $f_1(x) < f_2(x)$ for $x \in (\hat{x} - \delta, \hat{x})$ and $f_1(x) > f_2(x)$ for $x \in (, \hat{x}, \hat{x} + \delta)$.

So far, I have identified crossing points of functions in Φ as possible sources for points where the objective is not differentiable. A second source for crossing points can emerge in the context of the utility functions $u(W_1(V), h_1)$ and $u(W_2(V), h_2)$. Again, the objective is not necessarily differentiable at such crossing points. This is demonstrated by means of Example 3.6, which is also contained in the appendix.

Examples 3.5 and 3.6 demonstrate that crossing points are a possible source for points where the objective function is not differentiable. Henceforth, C_{prob} denotes the set of crossing points of functions in Φ and $C_{utility}$ denotes the set of crossing points of the two utilities $u(W_1(V), h_1)$ and $u(W_2(V), h_2)$ on $[0, 1]$. These sets can be defined formally as follows:

$$\begin{aligned} C_{prob} :=& \Big\{ \hat{V} \in [0,1] : \exists\, i, j \in 1, ..., n \; s.t. \; \hat{V} \text{ is a crossing point of } \pi_i \text{ and } \pi_j \Big\} \\ C_{utility} :=& \Big\{ \hat{V} \in [0,1] : \hat{V} \text{ is a crossing point of } u(W_1(V), h_1) \text{ and } u(W_2(V), h_2) \Big\} \end{aligned}$$

The following proposition gives conditions under which differential calculus can be used to analyze the patient's optimization problem.

Proposition 3.2. The patient's objective function is twice continuously differentiable when there are no crossing points, formally $C_{prob} = C_{utility} = \emptyset$. The objective is at least piecewise differentiable when both C_{prob} and $C_{utility}$ are finite.

Proof. The proof is contained in the appendix.

The condition $C_{prob} = \emptyset$ ensures that the preventive relationships in Φ can be ordered in a strict sense. This excludes the possibility that a preventive relationship $\pi_i \in \Phi$ yields a higher disease probability for a certain effort value $V_1 \in [0, 1]$ than another relationship π_j, and a lower disease probability than π_j for a different effort value $V_2 \in [0, 1] \backslash \{V_1\}$. The condition $C_{utility} = \emptyset$ ensures that for every $V \in [0, 1]$ the utility in one state of the world is always larger than in the other state of the world. Assume for instance the case

$$u(W_1(V), h_1) < u(W_2(V), h_2).$$

In this scenario, patients prefer to be healthy than to contract the illness irrespective of the effort level chosen.

Let V^* denote the set of solutions of the patient's optimization problem. If V^* is a singleton, one can differentiate between three different cases:

(1) *Corner solution 1:* No prevention is optimal.

(2) *Corner solution 2:* Maximum prevention is optimal.

(3) *Interior solution:* Partial prevention is optimal.

The following proposition gives conditions under which the objective is strictly concave.

Proposition 3.3 (Strict Concavity)**.** The objective function $\mathcal{U}$ is strictly concave if the following conditions are satisfied:

(SC_1) $C_{prob} = C_{utility} = \emptyset$

(SC_2) $u(W_1(V), h_1) < u(W_2(V), h_2)$ for all $V \in [0, 1]$

(SC_3) $\pi_{\mathbb{CEU}}$ is strictly convex

(SC_4) Both wealth functions W_i for $i = 1, 2$ are concave $W_i''(V) \leq 0$.

(SC_5) The inequality

$$W_1'(V)u'(W_1(V), h_1) - W_2'(V)u'(W_2(V), h_2) \geq 0 \tag{3.6}$$

holds for all $V \in [0, 1]$.

Proof. The proof is contained in the appendix.

Remark 3.2. Requirement (SC_3) is satisfied if every $\pi_i \in \Phi$ is convex and there is at least one $\pi_j \in \Phi$ that is strictly convex.

Remark 3.3. The statement of Proposition 3.3 remains true if the Assumptions (SC_3) and (SC_4) are replaced by the following set of requirements: $(\tilde{SC}_3)$ $\pi_{\mathbb{CEU}}$ is convex, $(\tilde{SC}_4)$ both wealth functions W_i are concave $W_i \leq 0$, and at least one wealth function is strictly concave.

Requirement (SC_1) excludes crossing points and ensures therefore that the objective is twice continuously differentiable. Assumption (SC_2) says that, irrespective of the effort level selected, patients always have a lower utility when they contract a disease than in a situation where they remain healthy. (SC_3) and (SC_4) are technical requirements. Condition (SC_5) can be expressed as

$$\frac{\frac{\partial u_1}{\partial V}}{\frac{\partial u_2}{\partial V}} \leq 1$$

where $u_1(V) = u(W_1(V), h_1)$ is patients' utility in the bad health state and $u_2(V) = u(W_1(V), h_1)$ denotes patients' utility in the good state. This representation has the following interpretation: on the left hand side is the marginal rate of substitution for more prevention in the bad versus the good health state. Hence, (SC_5) implies that patients would at least weakly prefer to exchange an additional unit of prevention in the bad health state for an additional unit of prevention in the good health state. The condition is automatically fulfilled if one of the marginal utilities $\frac{\partial u_i}{\partial V}$ is positive and the other negative.

The assumptions of Proposition 3.3 guarantee that there is either a unique interior maximizer or no interior maximizer. When there is no interior maximizer, the global optimum can be found at the boundary of the interval $[0, 1]$. The following proposition gives additional conditions under which no prevention or maximum prevention are feasible corner solutions.

Proposition 3.4 (Corner Solutions)**.** Let the patient's objective function $\mathcal{U}$ be strictly concave. Consider the following requirements:

(CS_1) $\lim_{V\to 0^+} \frac{\partial}{\partial V}\mathcal{U}(V|\alpha,\delta) < 0$

(CS_2) $\lim_{V\to 1^-} \frac{\partial}{\partial V}\mathcal{U}(V|\alpha,\delta) < 0.$

Requirement (CS_1) guarantees that $V^* = 0$ is a local maximizer. Condition (CS_2) makes sure that $V^* = 1$ is a local maximizer. If the first order condition has no solution, either $V^* = 0$ or $V^* = 1$ is the global maximizer. Moreover, there is either no corner solution or exactly one corner solution but there are never two corner solutions.

Proof. The proof is contained in the appendix.

3.3 Comparative Statics

In this section, I conduct a comparative static analysis for the prevention model presented in the previous section. This section is divided into two main parts. In the first part, I address the question of how preventive activities relate to the pessimism parameter α. It turns out that there is no clear-cut answer to this question, since the overall effect of pessimism on preventive effort depends on two concurrent effects, which will be explained in detail by means of simple numerical examples. As a next step, I treat the general case by looking at the overall effect of an increase in optimism on prevention. Subsequently, I give some general conditions under which the effect of pessimism can be clearly signed. In the second part, I relate the confidence parameter δ to preventive activities. It turns out that the comparative static analysis for the confidence parameter can be conducted in analogy to the case of the pessimism parameter α. Again, there are two concurrent effects at work, which entail a variety of different cases to be considered.
For technical reasons, I make the following set of assumptions throughout the rest of the paper.

Assumption 7. Patients' objectives satisfies the following requirements:

(CPS_1) The are no crossing points $C_{utility} = C_{prob} = \emptyset$.

(CPS_2) Patients' utility in the good state is always higher than in the bad state irrespective of the effort level $V[0,1]$ chosen. In formal terms,

$$u(W_2(V), h_2) - u(W_1(V), h_1) > 0.$$

(CPS_3) Patients' optimization problem has a unique solution.

Remark 3.4. Both the first and the second condition correspond to the Requirements (SC_1) and (SC_2) of Proposition 3.3. The last condition is imposed for technical reasons to simplify the analysis of the problem. Note that (CPS_3) becomes devoid of purpose if one assumes that the objective is strictly concave. Following Proposition 3.3, this can be achieved by implementing Requirements (SC_3) to (SC_5). Condition (CPS_3) is less restrictive than the Requirements (SC_3) to (SC_5) taken together, since a unique maximizer cannot be ruled out when the strict concavity conditions are violated.

As a direct consequence of Assumption 7, we can use differential calculus; there is either a unique interior maximizer or a unique corner solution. Henceforth, let $V^*(\alpha, \delta)$ be defined as the solution of the patient's optimization problem given the parameter constellation (α, δ):

$$V^*(\alpha, \delta) := \arg \max_{V \in [0,1]} \Big\{ \pi_{\mathbb{CEU}}(V|\alpha, \delta)\; u(W_1(V), h_1) \; + (1 - \pi_{\mathbb{CEU}}(V|\alpha, \delta))\; u(W_2(V), h_2) \Big\}.$$

Notation 3.1. Throughout this section, I condense the notation slightly, writing V^* instead of $V^*(\alpha, \delta)$ and $\mathcal{U}$ instead of $\mathcal{U} = \mathcal{U}(V^*)$ if not otherwise specified.

Pessimism and Preventive Activities

In this subsection, I analyze the effect of optimism on preventive activities. Assume for now that V^* is a unique interior solution. In this case, one can use the implicit function theorem to analyze the problem. It is

$$\frac{dV^*}{d\alpha} = -\frac{\frac{\partial^2 \mathcal{U}}{\partial\alpha\partial V}}{\frac{\partial^2 \mathcal{U}}{\partial V^2}}. \tag{3.7}$$

As the objective is assumed to be strictly concave, one can infer that the denominator of the α-derivative of V^* is negative and that the overall sign of (3.7) is determined by the sign of

$$\frac{\partial^2 \mathcal{U}}{\partial\alpha\partial V} = \Delta_u \frac{d^2}{d\alpha dV}\pi_{\mathbb{CEU}} + \left(\frac{d}{dV}\Delta_u\right)\frac{d}{d\alpha}\pi_{\mathbb{CEU}}$$

where

$$\Delta_u := u(W_1(V^*), h_1) - u(W_2(V^*), h_2).$$

Hence, the sign of (3.7) depends on the two determinants, Δ_1 and Δ_2, which are defined by

$$\Delta_1 := \Delta_u \frac{d^2}{d\alpha\, dV}\pi_{\mathbb{CEU}} \quad \text{and} \quad \Delta_2 := \left(\frac{d}{dV}\Delta_u\right)\frac{d}{d\alpha}\pi_{\mathbb{CEU}}.$$

Δ_1 is termed *perceived efficacy effect*, and Δ_2 is denoted as *expected marginal utility effect.* As a next step, I discuss both influencing factors Δ_1 and Δ_2 in detail, starting with Δ_1.

Perceived Efficacy Effect

Note that Δ_1 can be positive, negative or zero. This is demonstrated by means of the following example, which considers three scenarios where Δ_1 is considered in isolation.

This is done by selecting utilities $u(W_1(V), h_1)$ and $u(W_2(V), h_2)$ with the same marginals for every possible effort level $V \in [0, 1]$.

Example 3.1. Throughout this example, I specify the utilities by

$$u(W_1(V), h_1) = 8 - 2V^2 \quad \text{and} \quad u(W_2(V), h_2) = 12 - 2V^2.$$

Moreover, the confidence parameter is given by $\delta = \frac{1}{2}$, and the prior $q = (q_1, q_2)$ is defined by $q_1 = q_2 = \frac{1}{2}$. I compare three scenarios. In each scenario, a specific set of belief functions Φ_i is defined, and patients exhibit either extreme optimism $\alpha = 1$ or extreme pessimism $\alpha = 0$.

Scenario I

The first scenario shows that there are instances where a higher degree of pessimism yields a higher degree of preventive activities. Let $\Phi_1 = \{\pi_1, \pi_2\}$ where $\pi_1(V) = 1 - \frac{7}{15}V$ and $\pi_2(V) = \frac{1}{2} - \frac{2}{16}V$. Figure 3.4 displays the respective objectives in one diagram. Evidently, patients exert a higher level of effort in cases of extreme pessimism. Hence,

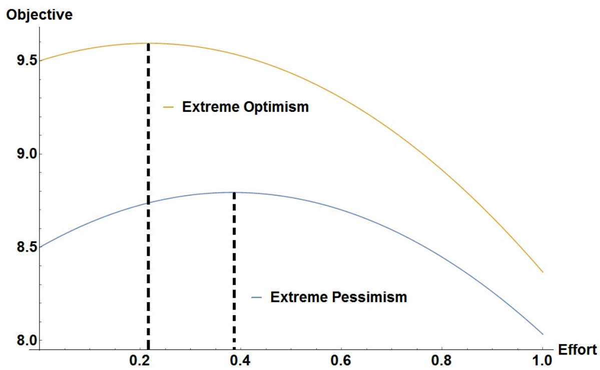


FIGURE 3.4: Pessimism Increases Preventive Activities

there are specifications where pessimism increases preventive activities.

Scenario II

The following scenario provides the opposite statement. Let $\Phi_2 = \{\pi_3, \pi_4\}$ where $\pi_3(V) = 1 - \frac{2}{15}V$ and $\pi_4(V) = \frac{1}{2} - \frac{7}{15}V$.[16] Figure 3.5 represents the respective objectives under extreme pessimism and extreme optimism.

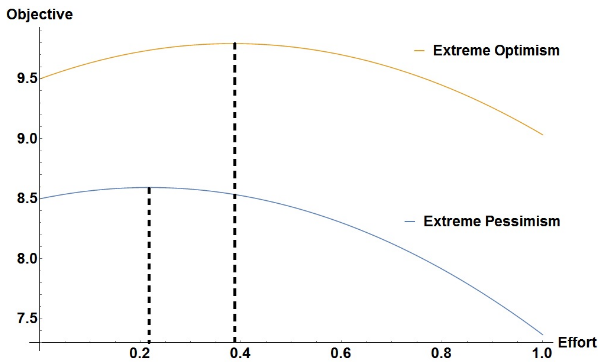


FIGURE 3.5: Pessimism Decreases Preventive Activities

Clearly, patients exert a lower level of effort in cases of extreme pessimism for this model specification. Hence, there are instances where pessimism decreases preventive activities.

Scenario III

The last scenario provides a specification where pessimism has no influence on patients' preventive activities. Let $\Phi_3 = \{\pi_5, \pi_6\}$ where $\pi_5(V) = 1 - \frac{7}{15}V$ and $\pi_6(V) = \frac{1}{2} - \frac{7}{15}V$. Figure 3.6 displays the objectives for the case of extreme pessimism and extreme optimism. Obviously, extreme pessimism and extreme optimism yield the same preventive activities.

In order to provide an explanation for the results of Example 3.1, a more profound analysis of the problem is required. Note that the Assumptions 1, 2, and 3 of the model framework are satisfied. Besides, there are no crossing points $C_{utility} = C_{prob} = \emptyset$. The only difference lies in the set of preventive relationships Φ_i for $= 1, 2, 3$. Since Δ_u is strictly negative

[16]Note that π_3 has the same intercept as π_1, and π_2 has the same intercept as π_4. Besides, the slope of π_3 corresponds to the slope of π_2, and the slope of π_4 corresponds to the slope of π_1.

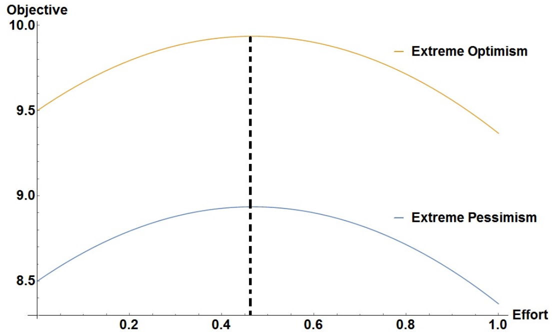
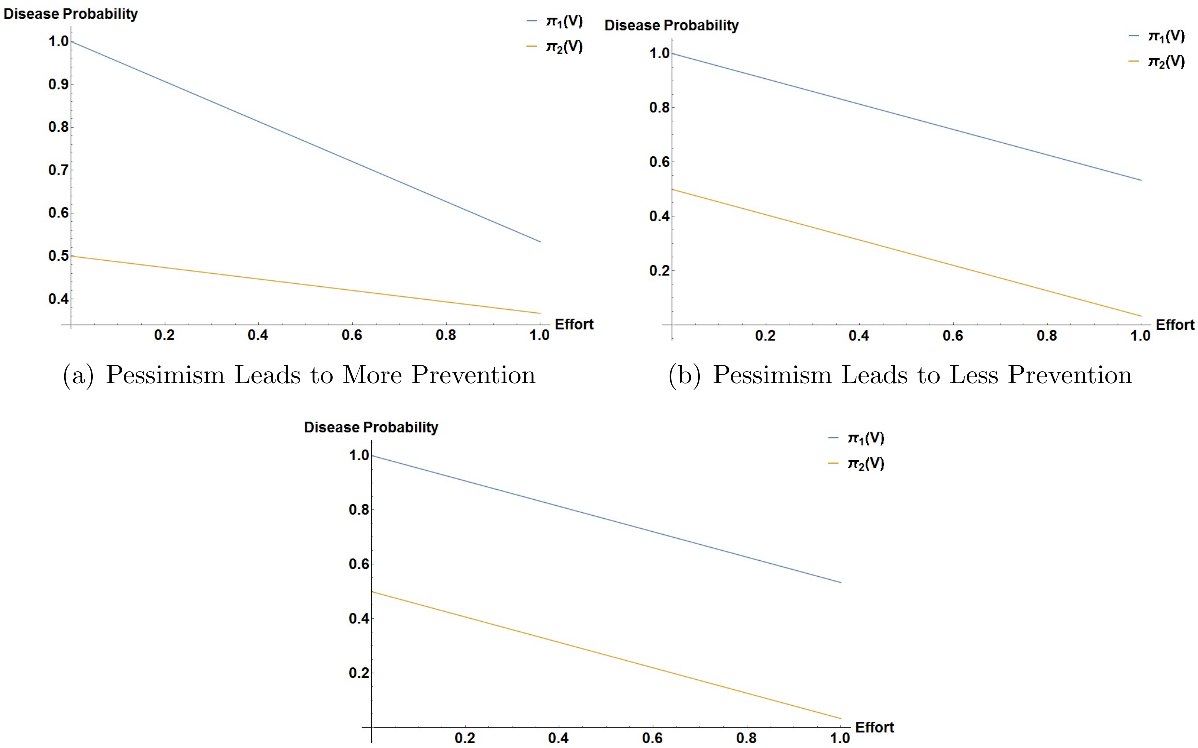


FIGURE 3.6: Pessimism Does Not Affect Preventive Activities

by Assumption (CPS_2), we can infer that π_{min} is given by π_1 in the first, by π_3 in the second, and by π_5 in the third part of the example. Similarly, π_{max} is given by π_2 in the first part of the example, by π_4 in the second part of the example, and by π_6 in the last part.

An obvious distinguishing feature between the three model specifications emerges when comparing the slopes of π_{min} and π_{max} on a case-by-case basis. Figure 3.7 gives a graphical representation for the functions in Φ_1, Φ_2 and Φ_3.



(a) Pessimism Leads to More Prevention

(b) Pessimism Leads to Less Prevention

(c) Pessimism has no Effect on Prevention

FIGURE 3.7: Perceived Efficacy Effect

In the first part of the example, π_{min} decreases more sharply than π_{max} for every effort level $V \in [0,1]$. In the second part, this relationship is reversed, whereas in the last part both π_{min} and π_{max} have identical slopes. How can we interpret the fact that π_{min} decreases more sharply than π_{max}? For every marginal increase in effort invested in the preventive measure, the marginal reduction of patients' perceived disease probability is larger in the pessimistic case than in the optimistic one. This observation can be formalized by means of the following definition.

Definition 3.3. Let π and $\hat{\pi}$ denote two differentiable preventive relationships. The relationship π is termed more effective than the relationship $\hat{\pi}$ if and only if

$$\left|\pi'(V)\right| > \left|\hat{\pi}'(V)\right| \quad \text{for all} \quad V \in [0,1].$$

Remember, Δ_2 is zero in Example 3.1 since both utilities $u(W_1(V), h_1)$ and $u(W_2(V), h_2)$ have identical slopes. Hence, the overall effect of optimism on preventive activities depends on the sign of

$$\frac{d^2}{d\alpha dV}\pi_{\mathbb{CEU}}(V_0|\alpha,\delta) \tag{3.8}$$

only. The derivative (3.8) describes how patients' perceived effectiveness of the preventive measure changes as they become more optimistic. In the first scenario, patients' perceived effectiveness of the preventive measure decreases with increasing optimism. As a result, pessimists deem the preventive measure more effective than optimists irrespective of the effort level chosen. In the second scenario, this relationship is reversed. Hence, optimists consider the preventive measure more effective than pessimists. Finally, in the third scenario, patients' perceived effectiveness is independent of the pessimism parameter α. As a consequence, optimists and pessimists consider the preventive measure as equally effective for all possible effort constellations.

The following proposition illustrates more clearly how patients' perceived effectiveness and preventive activities are interrelated.

Proposition 3.5 (Effectiveness and Prevention)**.** Let π and $\hat{\pi}$ be two differentiable preventive relationships. Moreover, define

$$V^{\pi} := \arg\max_{V\in[0,1]} \pi(V)u(W_1(V),h_1) + (1-\pi(V))u(W_2(V),h_2) \tag{3.9}$$

and

$$V^{\hat{\pi}} := \arg\max_{V\in[0,1]} \hat{\pi}(V)u(W_1(V),h_1) + (1-\hat{\pi}(V))u(W_2(V),h_2). \tag{3.10}$$

It is $V^{\hat{\pi}} \geq V^{\pi}$ if the following conditions are satisfied:

(1) $\hat{\pi}$ is more effective than π.

(2) $\hat{\pi}(V) \geq \pi(V)$ for all $V \in [0,1]$.

Proof. The proof is contained in the appendix.

Remark 3.5. Proposition 3.5 examines how preventive activities react when the underlying distorted probability π is replaced by a more effective distorted probability $\hat{\pi}$. It turns out that prevention increases weakly if $\hat{\pi}$ entails a higher perceived disease probability for every $V \in [0,1]$.

Expected Marginal Utility Effect

In this section, the expected marginal utility effect Δ_2 is analyzed. First of all, note that Δ_2 can be rewritten in the following way:

$$\Delta_2 = \frac{d}{d\alpha} E_{\pi_{\mathbb{CEU}}}\left[\frac{d}{dV}u\right] \tag{3.11}$$

where $E_{\pi_{\mathbb{CEU}}}$ denotes an expectation operator with respect to the distorted probability $\pi_{\mathbb{CEU}}$. Thus, Δ_2 describes how patients' expected marginal utility changes α as they become more optimistic. Like in the case of the perceived efficacy effect, Δ_2 can be positive, negative or zero. This is proved by means of the following numerical example. Note that the case where Δ_2 equals zero has been implicitly treated in Example 3.1. Therefore, it is sufficient to cover only cases where Δ_2 is either strictly positive or strictly negative.

Example 3.2. Throughout this example, I define the set of beliefs by $\Phi_4 = \{\pi_7, \pi_8\}$ where $\pi_7(V) = 1 - \frac{6}{15}V$ and $\pi_8(V) = \frac{1}{2} - \frac{6}{15}V$. The confidence parameter is given by $\delta = 1$ and the prior $q = (q_1, q_2)$ by $q_1 = q_2 = \frac{1}{2}$. I compare two scenarios. In each scenario, a pair of utilities $u(W_1(V), h_1)$ and $u(W_2(V), h_2)$ is defined. Moreover, I contrast extreme pessimism $\alpha = 0$ with extreme optimism $\alpha = 1$ in each scenario.

Scenario I

The first scenario provides a model specification where Δ_2 is strictly positive. The utilities are specified by $u(W_1(V), h_1) = 10 - 8V^3$ and $u(W_2(V), h_2) = 20 - 2V^3$. Figure 3.8 displays the patient's objective for the case of extreme pessimism $\alpha = 0$ and extreme optimism $\alpha = 1$.

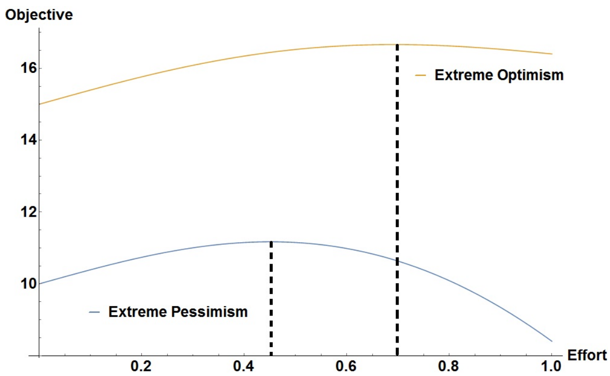


FIGURE 3.8: Less Prevention Under Pessimism

Figure 3.8 shows that, given the model specification above, patients' preventive activities are lower under extreme pessimism than under extreme optimism.

Scenario II

The following model specification provides the converse result. As a consequence, Δ_2 is strictly negative. Let the utilities be defined by

$$u(W_1(V), h_1) = 10 - 2V^3 \quad \text{and} \quad u(W_2(V), h_2) = 20 - 8V^3.$$

Figure 3.9 represents patients' objective functions under extreme pessimism and extreme optimism.

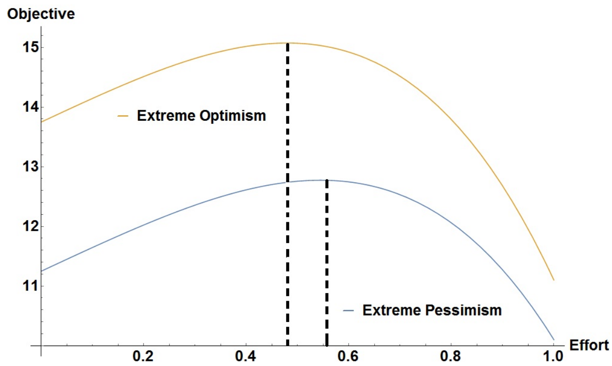


FIGURE 3.9: More Prevention Under Pessimism

Obviously, patients exert more effort under extreme pessimism than under extreme optimism.

The findings of Example 3.2 can be explained in the following way: Since the set of belief functions is the same in both scenarios, we can conclude that Δ_1 equals zero in both model specifications. Consequently, pessimists and optimists have the same perceived effectiveness of the preventive regime for every possible effort level $V \in [0, 1]$. As a result, the overall effect of α on prevention depends on the sign of Δ_2 only. Increasing optimism leads to decreasing preventive activities when the expected marginal utility from prevention is lower under optimism than under pessimism. But when is that the case? Obviously, the sign of Δ_2 is determined by the sign of

$$
\begin{aligned}
-\Delta_u' :=& \frac{d}{dV}\left(u(W_2(V), h_2) - u(W_1(V), h_1)\right) \\
=& W_2'(V)u'(W_2(V), h_2) - W_1'(V)u'(W_1(V), h_1)
\end{aligned}
$$

Note that $-\Delta_u' \leq 0$ if and only if Condition (SC_5) is satisfied. Remember, (SC_5) can be expressed in a marginal rate of substitution form

$$\frac{\frac{\partial u_1}{\partial V}}{\frac{\partial u_2}{\partial V}} \leq 1.$$

Hence, optimism decreases preventive activities as long as patients prefer to exchange a marginal unit of prevention in the bad state with a marginal unit of prevention in the good health state. The converse statement is true when patients prefer an additional marginal unit of prevention in the bad state. In other words, Δ_2 is negative when a marginal increase in prevention is better in the good health state. To be more precise, patients' utility in the good state increases more strongly than in the bad state when prevention is beneficial, and decreases less strongly than in the bad state when prevention is detrimental. This is exactly the case in the second part of Example 3.2.

Overall Effect of Pessimism on Prevention

The effect of pessimism on preventive activities is driven by two concurrent effects. The perceived efficacy effect describes how patients' perception of the efficacy of the preventive regime changes as α increases. Three different cases can occur, depending on whether a pessimistic patient deems a preventive measure more, less, or equally effective than a more optimistic patient. The expected marginal utility effect comes into being because an increase in α reduces the impact of patients' expected marginal utility on preventive activities. Again, three cases are possible depending on the constellation of patients' marginal utilities in the good and the bad health state. The overall effect of pessimism on prevention can be clearly identified when both Δ_1 and Δ_2 have the same sign, or when at least one of the two effects is zero. The effect cannot be clearly signed when Δ_1 and Δ_2 have opposite signs. This is the case when $\Delta_1 > 0$ and $\Delta_2 < 0$, or when $\Delta_1 < 0$ and $\Delta_2 > 0$. In both cases, the overall effect is determined by the magnitude of each individual effect. As a consequence, three scenarios can materialize: $|\Delta_1| > |\Delta_2|$, $|\Delta_1| < |\Delta_2|$ and $|\Delta_1| = |\Delta_2|$

In the first scenario, the perceived efficacy effect dominates the expected marginal utility effect. Whether optimism increases or decreases prevention depends on the sign of Δ_1. In the second scenario, the expected marginal utility effect dominates the perceived efficacy effect. As a result, prevention increases when Δ_2 is positive, prevention decreases when Δ_2 is negative, and prevention remains stable when Δ_2 is zero. Finally, in the third scenario, both effects have the same magnitude. Consequently, optimism does not affect preventive activities.

In the following, I give a set of conditions under which an increase in α can be clearly signed. First of all, remember that under strict concavity, which is ensured by the Conditions (SC_1) to (SC_5) of Proposition 3.3, we can infer that the expected marginal utility effect is non-positive.[17] As a consequence, we can conclude that the overall effect of optimism on prevention is negative when the perceived efficacy effect Δ_1 is negative or zero.

Corollary 3.2. The perceived efficacy effect is negative or zero $\Delta_1 \leq 0$ when the worst-case relationship π_{min} is more effective than the best-case relationship π_{max}.

Proof. The proof is contained in the appendix.

Remark 3.6. The slope ordering condition between π_{min} and π_{max} says that patients whose beliefs reflect a higher probability of disease perceive the preventive measure as more effective than those who base their decision on a lower perceived disease probability.

[17] This is implied by Condition (SC_5).

Remark 3.7. Δ_1 can be zero even if $\delta > 0$ due to the possibility of corner solutions.

The following proposition summarizes how optimism relates to preventive activities when the strict concavity conditions of Proposition 3.3 are imposed and the best- and worst-case preventive relationships can be ranked by their effectiveness.

Proposition 3.6 (Optimism)**.** Let $V^*(\alpha, \delta)$ be an interior solution. Under the Assumptions (SC_1) to (SC_5), the following comparative static results hold with respect to α:

(a) In cases of full confidence $\delta = 0$, a marginal increase in α has no effect on prevention.

(b) If patients give a positive weight $\delta > 0$ to extreme outcomes, the sign of the overall effect depends on the effectiveness ranking between π_{min} and π_{max}. When π_{min} is more effective than π_{max}, patients decrease preventive activities. In cases where π_{max} is more effective than π_{min}, the overall effect depends on the magnitude of Δ_1 and Δ_2. If the perceived efficacy effect Δ_1 is stronger than the expected marginal utility effect Δ_2, we can conclude that optimism increases prevention. The converse is true when the expected marginal utility effect dominates the perceived efficacy effect. When both effects have the same magnitude, it can be demonstrated that optimism does not affect preventive activities.

Remark 3.8. Preventive activities decrease in case of an interior solution if π_{min} is more effective than π_{max}. In case of the corner solution $V^* = 0$, there is no effect on optimal prevention. If $V^* = 1$, preventive activities either remain the same or decrease.

Confidence and Preventive Activities

In this section, I examine the relationship between confidence and preventive activities. It turns out that the analysis in this paragraph can be performed in analogy to the analysis of the optimism parameter α. If V^* is an interior solution, we obtain

$$\frac{dV^*}{d\delta} = -\frac{\frac{\partial^2 \mathcal{U}}{\partial \delta \partial V}}{\frac{\partial^2 \mathcal{U}}{\partial V^2}} \tag{3.12}$$

where $\frac{\partial^2 \mathcal{U}}{\partial V^2}$ is strictly negative. Hence, whether confidence increases or decreases preventive activities depends on the sign of

$$\frac{\partial^2 \mathcal{U}}{\partial \delta\, \partial V} = \Delta_u \frac{\partial^2}{\partial \delta\, \partial V} \pi_{\mathbb{CEU}} + \left(\frac{d}{dV} \Delta_u \right) \frac{\partial}{\partial \delta} \pi_{\mathbb{CEU}}.$$

One can see that the overall effect depends on two concurrent factors, Δ_3 and Δ_4, which are defined by

$$\Delta_3 := \Delta_u \frac{\partial^2}{\partial \delta\, \partial V} \pi_{\mathbb{CEU}} \quad \text{and} \quad \Delta_4 := \left(\frac{d}{dV} \Delta_u \right) \frac{\partial}{\partial \delta} \pi_{\mathbb{CEU}}.$$

In analogy to the comparative static section on the pessimism parameter α, I denote Δ_3 as *δ-perceived efficacy effect* and Δ_4 as *δ-expected marginal utility effect.* In the following, both effects are analyzed in detail.

δ-Perceived Efficacy Effect

The δ-perceived efficacy effect describes how patients adjust their perception of the preventive regime's effectiveness as they become less confident in the reference probability π_q. An increase in δ leads patients to give a higher weight to extreme outcomes and a lower weight to the reference belief π_q. This means patients become less confident that the reference function describes the true underlying preventive relationship. Whether Δ_1 is positive, negative, or zero depends on the effectiveness ranking between the reference function π_q and the extreme-outcome combination

$$\pi_\alpha := \alpha \pi_{min} + (1 - \alpha) \pi_{max}.$$

This is illustrated numerically by means of Example 3.7 in the appendix, where Δ_3 is contemplated in isolation. The following corollary more closely examines how the effectiveness ranking between π_q and π_α and the sign of Δ_3 are interrelated.

Corollary 3.3. Let $V^*(\alpha, \delta)$ be an interior solution. Δ_3 is positive if π_α is more effective than π_q. Δ_3 is negative if π_q is more effective than π_α, and zero if $\pi_\alpha = \pi_q$.

Proof. The proof is contained in the appendix.

Remark 3.9. Corollary 3.3 extends in the following way to the case of corner solutions: In a case where patients exert zero prevention, we can conclude that Δ_3 is non-negative if π_α is more effective than π_q. Δ_3 is zero if π_q is more effective than π_α. In cases where patients exert maximum effort, we can infer that Δ_3 is non-positive if π_q is more effective than π_α. Δ_3 equals zero if π_q is less effective than π_α.

An increase in δ can be interpreted as lowering confidence in the reference probability π_q. As a consequence, the underlying prevention function $\pi_{\mathbb{CEU}}(\alpha, \delta)$ shifts towards another prevention function $\hat{\pi}_{\mathbb{CEU}}(\alpha, \delta')$ that gives larger weight to the extreme-outcome part π_α. This implies that the overall effect on prevention depends on the effectiveness ordering between π_q and π_α. If π_α is more effective than π_q, we can conclude that preventive activities increase or remain the same as δ increases. This is because the patient gives a higher weight to the more effective part of his belief functional. The converse is true when the effectiveness ordering between π_q and π_α is reversed.

δ-Expected Marginal Utility Effect

The δ-expected marginal utility effect Δ_4 can be rewritten in the form

$$\Delta_4 = \frac{d}{d\delta}\pi_{\mathbb{CEU}} \cdot \Delta'_u = \frac{d}{d\delta}E_{\pi_{CEU}}\left[\frac{d}{dV}u\right] \tag{3.13}$$

where $E_{\pi_{CEU}}$ denotes the expectation with respect to the distorted probability $\pi_{\mathbb{CEU}}$, and $\frac{d}{dV}u$ denotes patients' marginal utility in the different health states. Δ_4 describes how patients' expected marginal utility is affected as their beliefs give a larger weight to the extreme-outcome combination π_α. In the following, I analyze this effect in detail.

Note that Δ_4 can be positive, negative, or zero. The overall sign of Δ_4 depends on the individual signs of $\frac{d}{d\delta}\pi_{\mathbb{CEU}}$ and Δ'_u. Note that Δ_4 is only zero when at least one of the following conditions holds:

(a) $\pi_q = \pi_\alpha$

(b) Both utilities $u(W_1(V), h_1)$ and $u(W_2(V), h_2)$ have the same marginal utility.

Scenario (a) corresponds to a situation where patients hold a subjective belief. In this context, the objective is independent of α and δ. Consequently, δ has no effect on preventive activities. In scenario (b), patients' marginal utilities are the same in both health states. Hence, patients' expected marginal utility is a constant and therefore independent of $\pi_{\mathbb{CEU}}$.

The derivative

$$\frac{d}{d\delta}\pi_{\mathbb{CEU}} = -\pi_q + \alpha\pi_{max} + (1-\alpha)\pi_{min}$$

is positive when the reference probability π_q induces a lower perceived disease probability than the extreme-outcome combination π_α. This is the case when patients are sufficiently pessimistic. To be more precise, let $\hat{\alpha}(V)$ denote the pessimism parameter for which both the reference probability and the extreme outcome combination yield the same perceived disease probability for a fixed level of effort $V \in [0,1]$. Then,

$$\pi_q(V) = \hat{\alpha}(V)\pi_1(V) + (1-\hat{\alpha}(V))\pi_n(V). \tag{3.14}$$

Solving equation (3.14) for $\hat{\alpha}(V)$, we obtain

$$\hat{\alpha}(V) = \frac{\pi_q(V) - \pi_n(V)}{\pi_1(V) - \pi_n(V)}.$$

Hence, the disease probability induced by π_q is smaller than the disease probability induced by π_α for $\alpha < \hat{\alpha}(V)$. The converse is true when π_q features a higher perceived disease probability than π_α. This is when patients are sufficiently optimistic with $\alpha > \hat{\alpha}(V)$.

The sign of Δ'_u is discussed in the comparative static section on the parameter α. Remember, Δ'_u is positive when patients prefer to exchange a marginal unit of prevention in bad health with a marginal unit of prevention in good health. Conversely, Δ'_u is negative when patients would at least weakly prefer to exchange prevention in the good state with prevention in the bad state. This proves the following corollary.

Corollary 3.4. Δ_4 is negative if patients are either sufficiently pessimistic $\alpha < \hat{\alpha}(V^*)$ and prefer to exchange prevention in the bad health state with prevention in the good health state, or if they are sufficiently optimistic $\alpha > \hat{\alpha}(V^*)$ and prefer to exchange prevention in the good health state with prevention in the bad health state. Δ_4 is positive if patients are either sufficiently pessimistic $\alpha < \hat{\alpha}(V^*)$ and prefer to exchange prevention in the good state with prevention in the bad state, or if they are sufficiently optimistic $\alpha > \hat{\alpha}(V^*)$ and prefer to exchange prevention in the bad state with prevention in good state.

Overall Effect of Confidence on Prevention

Clearly, whether confidence increases or decreases preventive activities depends on two concurrent effects, the δ-perceived efficacy effect and the δ-expected marginal utility effect. Similar to the analysis with respect to the pessimism parameter α, one can distinguish different cases. When both effects are strictly positive, or at least one effect is positive and the other effect is zero, we can conclude that the overall effect on prevention is positive. When both effects are zero, the overall effect is zero. When both effects are negative, or at least one effect is positive and the other negative, we can infer that the overall effect is negative. When one of the effects is strictly positive and the other strictly negative, the overall sign depends on the magnitude of each individual effect. When both effects have the same magnitude, the overall effect is zero. When one effect outweighs the other effect, the overall effect has the same sign as the effect with the larger magnitude.

In the following, I examine how preventive activities react to an increase in δ when the objective function is strictly concave.

Corollary 3.5. Under the Requirements (SC_1) to (SC_5), Δ_4 is negative if patients are sufficiently optimistic $\alpha > \hat{\alpha}(V^*)$, positive if patients are sufficiently pessimistic $\alpha < \hat{\alpha}(V^*)$, and zero for the intermediate pessimism parameter $\alpha = \hat{\alpha}(V^*)$.

Proof. The proof is contained in the appendix.

As a next step, we can analyze the overall effect of confidence on prevention under the assumption that π_α and π_q can be ranked according to their effectiveness.

Proposition 3.7 (Confidence)**.** Let V^* be an interior solution. Moreover, π_α and π_q can be ranked according to their effectiveness. Under the Requirements (SC_1) to (SC_5), preventive activities react in the following way to marginal increases in δ.

(a) Prevention increases if both π_α is more effective than π_q and patients are sufficiently pessimistic $\alpha < \hat{\alpha}$. For $\alpha = \hat{\alpha}$ preventive activities remain unchanged. In cases of strong enough optimism $\alpha > \hat{\alpha}$, the overall effect depends on the magnitude of Δ_3 and Δ_4. If the δ-perceived efficacy effect is stronger than the expected marginal utility effect, we can conclude that an increase in δ entails intensified prevention. The converse is true when the δ-expected marginal utility effect dominates the δ-perceived efficacy effect. When both effects have the same magnitude, preventive activities remain unchanged.

(b) Prevention decreases if both π_α is less effective than π_q and patients are sufficiently optimistic $\alpha > \hat{\alpha}$. For $\alpha = \hat{\alpha}$ preventive activities remain unchanged. In cases of sufficient pessimism $\alpha < \hat{\alpha}$, the overall effect depends on the magnitude of Δ_3 and Δ_4. If the δ-perceived efficacy effect is stronger than the δ-expected marginal utility effect, we can conclude that an increase in δ yields lower preventive activities. The converse is true when the δ-expected marginal utility effect dominates the δ-perceived efficacy effect. When both effects have the same magnitude, preventive activities remain unchanged.

Proof. The proof is a direct consequence of Corollaries 3.3 and 3.5. □

In short, Proposition 3.7 says that the effect of confidence on preventive activities depends on two factors. The first factor is the effectiveness ranking between the reference probability π_q and the extreme-outcome combination π_α. The second factor is patients' attitude towards ambiguity relative to $\hat{\alpha}$.

3.4 Preventive Effort When Patients Receive New Information

Modeling Information

This section studies how information on the underlying preventive regime affects patients' preventive activities. Throughout the rest of the paper, I assume that the Conditions (SC_1) to (SC_5) of Proposition 3.3 are satisfied. Hence, there is always a unique interior maximizer or a unique corner solution. Information is modeled by means of signal s conveyed by the physician. The signal is the realization of a random variable S with values in $\{1, ..., n\}$. The index of the true underlying preventive relationship is denoted by $\theta \in \{1, ..., n\}$. Moreover, the patient knows the conditional distribution $\mathbb{P}^{S|\theta}$ of the signal given θ. Henceforth, I denote with $p_{ij} = \mathbb{P}(S = i|\theta = j)$ the conditional probability that the physician conveys relationship i to be the true relationship given $\theta = j$. In addition, p_{ij} denotes the probability that the physician conveys the wrong relationship i for $i \neq j$, and p_{ii} is the probability that the physician conveys a correct signal.

As a next step, patients' beliefs are updated. In the special case $\delta = 0$, the neo-additive belief reduces to a purely Bayesian belief $q = (q_1, ..., q_n)$, which is updated via Bayes' Rule. Under the assumption that the patient observes the signal $s = i$, the posterior probability q_j^{Bayes} of q_j is given by

$$q_j^{Bayes} = \frac{q_j p_{ij}}{\sum\limits_{k=1}^{n} q_k \; p_{ik}} \tag{3.15}$$

Given the updated prior

$$q^{Bayes} = (q_1^{Bayes}, ..., q_n^{Bayes})$$

patients maximize

$$\pi_{q^{Bayes}}(V) \; u(W_1(V), h_1) + (1 - \pi_{q^{Bayes}}(V)) \; u(W_2(V), h_2). \tag{3.16}$$

In the general case $\delta \in [0, 1]$, patients revise a neo-additive belief. Contrary to the Bayesian scenario, there are multiple ways to update non-additive probabilities. Eichberger et al. [2010] suggest three different ways of updating neo-additive capacities: an optimistic updating rule, a pessimistic updating rule, also called a Dempster-Shafer updating rule, and the so-called generalized Bayesian updating rule. All of these rules are motivated by prominent updating rules for capacities discussed in the literature. By Proposition 1, page 93 in Eichberger et al. [2010], one knows that, under each of these rules, the update of a neo-additive capacity $\nu(\alpha, \delta)$ is still neo-additive with new optimism and confidence parameters $\nu(\alpha', \delta')$, and a reference probability q^{Bayes} that corresponds to the Bayesian update of the prior reference probability q.

In cases where patients resort to the optimistic updating rule, the parameter α is updated to $\alpha^O = 1$. Under the pessimistic updating rule, patients revise α to $\alpha^P = 0$, and under the generalized Bayesian updating rule α is not affected by the patient's updating process, hence $\alpha^{GB} = \alpha$. The update of the confidence parameters depends on the signal realization s. Assume that the patient observes the signal $s = i$. Then, the confidence parameter δ is updated to

$$\delta^O = \frac{\delta\alpha}{(1-\delta)q_i + \alpha\delta}$$

in case of the optimistic updating rule. The confidence parameter under the pessimistic updating rule is given by

$$\delta^P = \frac{\delta(1-\alpha)}{(1-\delta)q_i + (1-\alpha)\delta}.$$

Under the generalized Bayesian updating rule, we obtain the revised confidence parameter

$$\delta^{GB} = \frac{\delta}{(1-\delta)q_i + \delta}.$$

The following corollary characterizes patients' objectives for each of the previously discussed updating rules.

Corollary 3.6. After observing the signal realization s, patients maximize the objective

$$\pi^U_{\mathbb{CEU}}(V|\cdot)\ u(W_1(V), h_1) + (1 - \pi^U_{\mathbb{CEU}}(V|\cdot))\ u(W_2(V), h_2)$$

where $\pi^U_{\mathbb{CEU}}(V|\cdot)$ is a rule-dependent distorted probability. It is

$$\pi^U_{\mathbb{CEU}} = (1-\delta^O)\pi^{Bayes}_q + \delta^O \pi_{max}$$

in case of the optimistic updating rule. For the pessimistic updating rule, we obtain

$$\pi^U_{\mathbb{CEU}} = (1-\delta^P)\pi^{Bayes}_q + \delta^P \pi_{min}.$$

In cases where the generalized Bayesian updating rule applies, we have

$$\pi_{\mathbb{CEU}} = (1-\delta^{GB})\pi^{Bayes}_q + \delta^{GB} \pi_{\alpha}.$$

Proof. The proof is contained in the appendix.

Information and Preventive Activities

The following proposition compares patients holding a Bayesian belief $\delta = 0$ with patients holding a Non-Bayesian belief $0 < \delta \leq 1$ after observing the signal realization s. Henceforth, I denote patients with $\delta = 0$ as Bayesian patients and patients with $\delta > 0$ as Knightian or Non-Bayesian patients.

Proposition 3.8. Assume that there are two types of patients: a Bayesian patient with a belief q and a Non-Bayesian patient holding a neo-additive belief with the same reference belief q. Moreover, let $V^* \in (0, 1)$ be an interior solution for the Bayesian patient.

(a) The Non-Bayesian patient exerts less effort than the Bayesian patient under the optimistic updating rule if π_{max} is less effective than π_q^{Bayes}.

(b) The Non-Bayesian patient exerts more effort than the Bayesian patient under the pessimistic updating rule if π_{min} is more effective than π_q^{Bayes}.

(c) The Non-Bayesian patient exerts less effort than the Bayesian patient under the generalized Bayesian updating rule if π_q^{Bayes} is more effective than π_α and $\pi_q^{Bayes}(V) > \pi_\alpha(V)$ for all $V \in [0, 1]$.

(d) The Non-Bayesian agent exerts more effort than the Bayesian patient under the generalized Bayesian updating rule if π_q^{Bayes} is less effective than π_α and $\pi_q^{Bayes}(V) < \pi_\alpha(V)$ for all $V \in [0, 1]$.

Proof. The proof is contained in the appendix.

The results of Proposition 3.8 show that the underlying updating rule plays a crucial role in explaining heterogeneous choices in preventive activities. If we consider the Bayesian agent as a rational, representative patient, the Non-Bayesian patient's behavior might be regarded as a deviation from this representative patient. Depending on the underlying updating rule, the Non-Bayesian agent's choice might lead to more or fewer preventive activities compared to the prevention level selected by the Bayesian patient.

The proof of Proposition 3.8 is based on Proposition 3.5, which says that replacing the distorted probability π_{CEU} with a more effective distorted probability $\hat{\pi}_{CEU}$ leads to intensified prevention if $\hat{\pi}_{CEU}$ exhibits a higher perceived disease probability than π_{CEU} for every possible effort level $V \in [0,1]$. Due to the fact that the inequality

$$\pi_{max}(V) \leq \pi_q^{Bayes}(V) \leq \pi_{min}(V)$$

holds for every $V \in [0,1]$, we can infer that the latter condition is satisfied for the optimistic and pessimistic updating rule.[18] In cases where the generalized updating rule applies, π_q^{Bayes} and π_{CEU}^U can be ranked according to their disease probabilities when there is a clear ranking between the extreme-outcome combination π_α and the Bayesian update π_q^{Bayes}. Example 3.8 in the appendix demonstrates that there are instances where such a clear ranking does not exist even when there are no crossing points $C_{prob} = \emptyset$. This raises the question: under which conditions is a clear ordering between π_q and π_α possible? The following corollary provides the answer.

Corollary 3.7. There is a maximum pessimism parameter α_{min} such that $\pi_q < \pi_\alpha$ for all $\alpha < \alpha_{min}$. Furthermore, there is a minimal pessimism parameter $\alpha_{max} \geq \alpha_{min}$ such that $\pi_q > \pi_\alpha$ for all $\alpha > \alpha_{max}$. There is no clear ordering between π_q and π_α when $\alpha \in (\alpha_{min}, \alpha_{max})$.

Proof. The proof is contained in the appendix.

In other words, π_q always features a lower perceived disease probability than π_α if patients are sufficiently pessimistic, and π_q always exhibits a higher perceived disease probability than π_α if patients are sufficiently optimistic. For intermediate values of the pessimism parameter, there is no clear ordering between π_α and π_q.

[18] In particular, we have $\pi_q^{Bayes}(V) > \pi_{CEU}^U(V)$ for all $V \in [0,1]$ in cases where the optimistic updating rule applies and $\pi_q^{Bayes}(V) < \pi_{CEU}^U(V)$ for all $V \in [0,1]$ in cases where the pessimistic updating rule applies.

Another interesting question related to the comparison between Bayesian and Knightian patients is how the preventive gap between Bayesian and Non-Bayesian patients reacts to the arrival of new information.

Definition 3.4. Let $V^{old}(\alpha, \delta)$ denote the solution of patients' optimization problem before observing the signal and let $V^{new}(\alpha, \delta)$ be the solution after observing the signal. Information increases the gap between Bayesian and Non-Bayesian patients if

$$|V^{old}(\alpha, 0) - V^{old}(\alpha, \delta)| < |V^{new}(\alpha, 0) - V^{new}(\alpha, \delta)|.$$

Information decreases the gap between Bayesian and Non-Bayesian patients if

$$|V^{old}(\alpha, 0) - V^{old}(\alpha, \delta)| > |V^{new}(\alpha, 0) - V^{new}(\alpha, \delta)|.$$

The gap between Bayesian and Non-Bayesian patients is not affected by the arrival of new information if

$$|V^{old}(\alpha, 0) - V^{old}(\alpha, \delta)| = |V^{new}(\alpha, 0) - V^{new}(\alpha, \delta)|.$$

Does the preventive gap between Bayesian and Non-Bayesian patients always decrease as new information becomes available? The following numerical example shows that this is not necessarily the case. In particular, one can find feasible model specifications such that both an increase and a reduction of the preventive gap is possible.

Example 3.3. In the following, three preventive relationships are considered with

$$\pi_1(V) = 1 - \frac{2}{3}V, \quad \pi_2(V) = \frac{1}{2} - \frac{1}{4}V, \quad \text{and} \quad \pi_3(V) = \frac{1}{4} - \frac{1}{8}V.$$

Obviously, $\pi_{min} = \pi_1$ and $\pi_{max} = \pi_3$. The utilities are defined by

$$u(W_1(V), h_1) = 10 - 2V^2 \quad \text{and} \quad u(W_2(V), h_2) = 20 - 2V^2.$$

The conditional probabilities $p_{ij} = \mathbb{P}(S = i|\theta = j)$ are specified by the stochastic matrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{3}{5} & \frac{1}{5} \end{pmatrix}.$$

The initial prior $q = (q_1, q_2, q_3)$ is given by $q_1 = \frac{1}{5}$, $q_2 = \frac{2}{5}$ and $q_3 = \frac{2}{5}$. Assuming that the patient receives the signal $s = 3$, the Bayesian update q^{Bayes} of q is given by

$$\begin{aligned} q_1^{Bayes} &= \frac{\frac{1}{25}}{\frac{1}{25} + \frac{6}{25} + \frac{2}{25}} = \frac{1}{9} \\ q_2^{Bayes} &= \frac{\frac{6}{25}}{\frac{1}{25} + \frac{6}{25} + \frac{2}{25}} = \frac{2}{3} \\ q_3^{Bayes} &= \frac{\frac{2}{25}}{\frac{1}{25} + \frac{6}{25} + \frac{2}{25}} = \frac{2}{9}. \end{aligned}$$

As a next step, we can determine patients' objectives in the Bayesian case $\delta = 0$. Using simple algebra, one can show that the objective function is given by

$$\mathcal{U}_1^{old}(V) = 15 + \frac{17}{6}V - 2V^2$$

before observing the signal. This is a quadratic function with the global maximizer $V_1^{old} = \frac{17}{24} \approx 0.71$. Similarly, one can show that the objective is given by

$$\mathcal{U}_1^{new}(V) = 15 + \frac{145}{54}V - 2V^2$$

after observing the signal with the maximizer $V_1^{new} = \frac{145}{216} \approx 0.67$. For the Non-Bayesian case, it is assumed that $\alpha = 1$ and $\delta = \frac{1}{2}$. Besides, the patient makes use of the generalized Bayesian updating rule. Remember, under the generalized Bayesian updating rule the parameter α remains unchanged and δ is updated to

$$\delta^{GB} = \frac{\delta}{(1-\delta)q_3 + \delta} = \frac{5}{7}.$$

The objective function before observing the signal is given by

$$\mathcal{U}_2^{old}(V) = \frac{65}{4} + \frac{49}{24}V - 2V^2$$

with the global maximizer $V_2^{old} \approx 0.51$. After observing the signal, the objective is given by

$$\mathcal{U}_2^{new}(V) = \frac{235}{14} + \frac{1255}{756}V - 2V^2$$

with the global maximizer $V_2^{new} \approx 0.41$. It is $|V_1^{old} - V_2^{old}| \approx 0.2$ and $|V_1^{new} - V_2^{new}| \approx 0.26$. This demonstrates that

$$|V_1^{old} - V_2^{old}| < |V_1^{new} - V_2^{new}|.$$

Hence, there are instances where information increases the difference in preventive activities between Bayesian and Non-Bayesian patients.

The following model specification demonstrates the converse results. Let $\alpha = 0.75$. The patient observes again $s = 3$. The remaining parameters are the same as in the first part of the example. Since the Bayesian objective is independent of α, we have the same objective in the Bayesian case. The Non-Bayesian objective is given by

$$\mathcal{U}_3^{old}(V) = \frac{245}{16} + \frac{87}{32} - 2V^2$$

before observing the signal. The global maximizer is $V_3^{old} \approx 0.68$. After observing the signal, one obtains the objective

$$\mathcal{U}_3^{new}(V) = \frac{865}{56} + \frac{1135}{432}V - 2V^2$$

with the global maximizer $V_3^{new} \approx 0.66$. It is $|V_1^{old} - V_3^{old}| \approx 0.03$ and $|V_1^{new} - V_3^{new}| \approx 0.01$. Consequently,

$$|V_1^{old} - V_3^{old}| > |V_1^{new} - V_3^{new}|.$$

Hence, there are instances where the gap between Bayesian and Non-Bayesian patients decreases with the arrival of new information.

Example 3.3 is interesting because it demonstrates that Bayesian and Non-Bayesian agents might react entirely differently to new information. Moreover, it becomes clear that information can reinforce or attenuate extreme behavior, depending on the underlying parameter constellations. Undoubtedly, extreme behavior is not a desirable consequence of information campaigns or health counseling.

Excessive Preventive Behavior and Preventive Inertia

The previous section provides an outline on how Bayesian and Non-Bayesian agents react to the arrival of new information by comparing their respective preventive activities. Non-Bayesian agents deviate from Bayesian agents and exert lower or higher levels of prevention. An issue of major importance is how strongly patients deviate from the true underlying relationship π_θ before and after observing the signal realization s. If patients knew π_θ, they would select effort by solving

$$\max_{V \in [0,1]} \pi_\theta(V) u(W_1(V), h_1) + (1 - \pi_\theta(V)) u(W_2(V) h_2). \tag{3.17}$$

In the following, V_θ^* denotes the solution of (3.17). By assumption, V_θ^* is unique.

Definition 3.5. Let $V^*(\alpha, \delta)$ be the solution of patients' optimization problem under Knightian uncertainty. Patients exhibit excessive preventive behavior if they select a higher level of effort under Knightian uncertainty than under a situation where they know true relationship π_θ. Formally, $V^*(\alpha, \delta) > V_\theta^*$. Patients exhibit preventive inertia if they select a lower level of effort under Knightian uncertainty than under a situation where they know π_θ. Formally, $V^*(\alpha, \delta) < V^*_\theta$.

An important question is how excessive preventive behavior or preventive inertia of Bayesian patients relates to excessive preventive behavior or preventive inertia of Knightian patients. Before treating the general case, the following special case is examined.

Patients Can Perfectly Infer the Correct Relationship

Patients can perfectly infer the correct preventive relationship from a signal realization when there is a signal i such that

$$p_{il} = P(S = i|\theta = l) = 1$$

for some $i, l \in \{1, ..., n\}$. In the special case $i = l$, the physician provides with probability one the correct signal realization as $s = i$ is observed. Otherwise, if $i \neq l$, physicians always communicate the wrong preventive relationship. Still, patients can infer the correct relationship from the wrong signal since they know how to relate the signal to the correct relationship π_θ. In cases where $s = i$ is observed, the Bayesian update of q is given by

$$q_j^{Bayes} = \begin{cases} 1 & \text{for} \quad j = \theta \\ 0 & \text{for} \quad j \neq \theta. \end{cases}$$

Consequently, Bayesian patients solve the "correct" optimization problem after the signal is observed. Thus, excessive preventive behavior and preventive inertia vanish by processing the signal realization s. As a next step, we examine cases where patients beliefs are represented by a neo-additive capacity with $\delta > 0$. By proposition 1 in Eichberger et al. [2010], we can infer that, under the optimistic updating rule, the updated neo-additive capacity is of the form

$$\nu^O(A) = (1 - \delta^O)\pi_E(A) + \delta^O$$

where $\pi_E(A)$ denotes the Bayesian update of the reference probability π, E is the conditioning event, and A is the event measured by the underlying capacity. In our example, $\pi_E(A)$ corresponds to q_i^{Bayes} as the signal $s = i$ is observed. Therefore, we can deduce that

$$\nu^O(s = \theta) = (1 - \delta^O) + \delta^O = 1$$

Note that the same statement holds for the optimistic and pessimistic updating rule under Knightian uncertainty. Consider first the case of the optimistic updating rule. Hence, patients assign a probability of one to the correct relationship after observing the signal. In cases where patients rely on the pessimistic updating rule, a similar reasoning applies. The capacity is of the form

$$\nu^P(A) = (1 - \delta^P)\pi_E(A)$$

Again, since $\pi_E(A)$ corresponds to q_i^{Bayes}, we can infer that

$$\nu^P(s = \theta) = (1 - \delta^P) = 1$$

since $\delta^P =$

such statement is not true when $\delta > 0$. This can be seen by looking at the updated distorted probability $\pi^U_{\mathbb{CEU}}$ for each of the updating rules discussed at the beginning of this section. Since the revised confidence parameters δ^O, δ^P and δ^{GB} are in general not equal to zero, we can conclude that $\pi^U_{\mathbb{CEU}}$ differs from

$$\pi_{q^{Bayes}} = \pi_\theta.$$

Hence, Non-Bayesian patients still deviate from the true underlying relationship even in a highly idealized world where physicians always communicate the correct preventive relationship. As a consequence, preventive inertia and excessive preventive behavior persist under Knightian uncertainty and cannot be eliminated by information campaigns and or physician counseling. Under the assumptions of Proposition 3.8, we can clearly identify whether Knightian patients exhibit preventive inertia or excessive preventive behavior after observing the signal.

Corollary 3.8. Let $i \in \{1, .., n\}$ be a signal such that patients can perfectly infer the correct relationship

$$p_{ij} = P(S = i|\theta = j) = 1.$$

Then, excessive preventive behavior and preventive inertia vanish if $\delta = 0$ and persist for $\delta > 0$. Knightian patients exhibit excessive preventive behavior under the pessimistic updating rule if π_{min} is more effective that π_q, as well as the generalized Bayesian updating rule if π_q^{Bayes} is more effective than π_α and $\alpha < \alpha_{min}$. Knightian patients exhibit preventive inertia under the optimistic updating rule if π_{max} is less effective than π_q, as well as the generalized Bayesian updating rule, if π_q is less effective than π_α and $\alpha > \alpha_{max}$.

Proof. The proof is a direct consequence of Proposition 3.8 and Corollary 3.7. □

What happens if the idealized assumption that patients can perfectly infer the correct signal realization is abandoned?

Physicians Provide a Wrong and a Correct Signal with Positive Probability

In this general case, preventive inertia and excessive preventive behavior persist even for Bayesian patients. This is because $\pi_{CEU}^U \neq \pi_\theta$. The following corollary investigates how excessive preventive behavior of Bayesian patients relates to excessive preventive behavior of Knightian patients after observing the signal.

Corollary 3.9. Let $V^*(\alpha, 0)$ be an interior solution for the Bayesian patient after observing the signal realization. Besides, let $V^*(\alpha, 0)$ feature excessive preventive behavior. Formally, $V^*(\alpha, 0) > V_\theta^*$. Knightian patients exhibit stronger ex-post excessive preventive behavior than Bayesian patients under

- the pessimistic updating rule if π_{min} is more effective than π_{max} and

- under the generalized Bayesian updating rule if π_q^{Bayes} is more effective than π_α and patients are sufficiently pessimistic $\alpha < \alpha_{min}$.

Knightian patients exhibit a lower level of ex-post excessive preventive behavior than Bayesian patients

- under the optimistic updating rule if π_{max} is less effective than π_{min}, and
- under the generalized Bayesian updating rule if π_q^{Bayes} is less effective than π_α and patients are sufficiently optimistic $\alpha > \alpha_{max}$.

Similarly, we can examine how preventive inertia of Knightian patients relates to preventive inertia of Bayesian patients after the signal is observed.

Corollary 3.10. Let $V^*(\alpha, 0)$ be an interior solution for the Bayesian patient after observing the signal realization. Besides, let $V^*(\alpha, 0)$ feature preventive inertia. Formally, $V^*(\alpha, 0) < V_\theta^*$. Knightian patients exhibit stronger ex-post preventive inertia than Bayesian patients

- under the optimistic updating rule if π_{max} is less effective than π_{min}, and
- under the generalized Bayesian updating rule if π_q^{Bayes} is less effective than π_α and patients are sufficiently optimistic $\alpha > \alpha_{max}$.

Knightian patients exhibit a lower level of ex-post preventive inertia than Bayesian patients

- under the pessimistic updating rule if π_{min} is more effective than π_{max}, and
- under the generalized Bayesian updating rule if π_q^{Bayes} is more effective than π_α and patients are sufficiently pessimistic $\alpha < \alpha_{min}$.

Proof. The proof of Corollary 3.9 and 3.10 is a direct consequence of Proposition 3.8. $\square$

Evidently, information can reinforce extreme behavior when the wrong signal is communicated. Assume for instance a situation where a Bayesian patient would already exhibit excessive preventive behavior before observing the signal. Moreover, the patient trusts highly the physician to communicate the correct signal. This implies conditional probabilities p_{ii} close to 1 for all $i = 1, ..., n$. In cases where the physician wrongly communicates a relationship, inducing excessive preventive behavior, patients adjust their beliefs by assigning a larger posterior probability to this relationship. As a consequence, it is very likely that excessive preventive behavior is reinforced.[19] On the other hand, if the physician communicates the correct signal, we can conclude that π_q comes closer to the true relationship π_θ in the sense that the posterior q^{Bayes} assigns a larger posterior probability to the correct relationship π_θ. But does this mean that excessive preventive behavior and preventive inertia automatically diminish? The following example illustrates that this is not the case, even when δ equals zero.

Example 3.4. Reconsider Example 3.3 by replacing the initial prior q with

$$\hat{q} = \left(\frac{3}{10}, \frac{3}{10}, \frac{2}{5}\right).$$

The patient observes the signal $s = 2$ and the true underlying relationship is given by

$$\pi_\theta(V) = \pi_2(V) = \frac{1}{2} - \frac{1}{4}V.$$

If patients knew the true relationship π_{theta}, they would maximize

$$\mathcal{U}^{real}(V) = \pi_2(V)u(W_1(V), h_1) + (1 - \pi_2(V))u(W_2(V), h_2)$$

Using simple algebra, we obtain

$$\mathcal{U}^{real}(V) = 15 + \frac{5}{2}V - 2V^2$$

[19]The statement is wrong when the updating process strongly reduces the probability of other preventive relationships which would, by themselves, induce even stronger excessive preventive behavior than the relationship communicated by the physician.

with the maximizer

$$V_\theta^* = \frac{5}{8} = 0.625.$$

Patients exhibit preventive inertia if $V^*(\alpha, \delta) < 0.625$ and excessive preventive behavior if $V^*(\alpha, \delta) > 0.625$. The Bayesian update of $\hat{q}$ is given by

$$\hat{q}_1^{Bayes} = \frac{\frac{3}{25}}{\frac{3}{25} + \frac{3}{25} + \frac{2}{25}} = \frac{3}{8}$$
$$\hat{q}_2^{Bayes} = \frac{\frac{3}{25}}{\frac{3}{25} + \frac{3}{25} + \frac{2}{25}} = \frac{3}{8}$$
$$\hat{q}_3^{Bayes} = \frac{\frac{2}{25}}{\frac{3}{25} + \frac{3}{25} + \frac{2}{25}} = \frac{1}{4}$$

Henceforth, it is assumed that δ equals zero. Consequently, patients' objectives are independent of α. Before observing the signal, the objective is given by

$$\mathcal{U}^{old}(V|\alpha, 0) := \frac{29}{2} + \frac{13}{4}V - 2V^2$$

with the maximizer $\hat{V}_1 = \frac{13}{16} \approx 0.81$. After observing the signal, we obtain the objective

$$\mathcal{U}^{new}(V|\alpha, 0) := \frac{55}{4} + \frac{15}{4}V - 2V^2$$

with the maximizer $\hat{V}_2 = \frac{15}{16} \approx 0.94$. Obviously, patients exhibit excessive preventive behavior before and after observing the signal realization. Besides, we have $|V_\theta^* - \hat{V}_1| = \frac{3}{16} \approx 0.19$ and $|V_\theta^* - \hat{V}_2| = \frac{5}{16} \approx 0.31$. This implies

$$|V_\theta^* - \hat{V}_1| < |V_\theta^* - \hat{V}_2| = \frac{5}{16} \approx 0.31.$$

Hence, excessive preventive behavior is reinforced, even in the "reduced" Bayesian case and under the "favorable" assumption that the physician provides the correct signal realization.

The reason for this observation lies in the prior $\hat{q}$ and the signal structure. Observe that

$\hat{q}_1 = \hat{q}_2$ and

$$P(S = 2|\theta = 1) = P(S = 2|\theta = 2).$$

This implies that the updates of $\hat{q}_1$ and $\hat{q}_2$ coincide. Hence, if the correct signal realization is provided, patients increase their posterior probability for π_2. At the same time, the posterior probability for π_1 increases with the same magnitude. The posterior probability for the third relationship π_3 decreases. Since π_i is more effective than π_j for $i < j$, and since $\pi_1(V) > \pi_2(V) > \pi_3(V)$ for all $V \in [0, 1]$, Proposition 3.5 shows that

$$V^*_{\theta=1} > V^*_{\theta=2} > V^*_{\theta=3}.$$

Due to the fact that the posterior assigns a larger probability to $\theta = 1$ and $\theta = 2$ and a smaller probability to $\theta = 3$, we can infer that preventive activities increase after observing the signal. Since patients already exhibit excessive preventive behavior before observing the signal, we can draw the conclusion that information intensifies excessive preventive behavior.

Example 3.4 demonstrates that information does not necessarily bring patients' preventive activities closer to their optimal levels, even in a favorable environment where physicians communicate the correct relationship and probabilities are not distorted. Even worse, there is the possibility that information induces patients to intensify excessive preventive behavior or preventive inertia. The following paragraph shows that there are instances where information does not affect preventive activities.

Communication of Uninformative Signals

In the following, a special case is considered where patients receive "uninformative" signals. Henceforth, a signal $s = i$ is called "uninformative" if its conditionals are given by

$$p_{ij} = P(S = i|\theta = j) = \frac{1}{n} \quad \text{for } j = 1, ..., n.$$

Given this specification of the signal structure, we obtain

$$q_j^{Bayes} = \frac{q_i \frac{1}{n}}{\sum_{j=1}^{n} q_j \frac{1}{n}} = q_i$$

as the Bayesian update of q. Hence, information does not affect the prior distribution. In a sense, such a signal is completely uninformative for a Bayesian patient since it precludes the possibility to draw any further inferences on the true underlying θ. As a consequence, the objective function of Bayesian patients $\delta = 0$ remains unaffected and excessive preventive behavior or preventive inertia persist in the same magnitude as without observing the signal. But how do Non-Bayesian patients react to uninformative signals? Again, the answer depends on the updating rule. Since δ^O, δ^P and δ^{GB} are in general not equal to zero, we can infer that $\pi_{CEU} \neq \pi_{\mathbb{CEU}}^{U}$. Hence, contrary to the Bayesian case, Knightian patients adjust their beliefs even in the light of uninformative signals.

3.5 Conclusion

This paper studies how patients adjust their primary preventive activities in the light of new information when the relationship between preventive effort and disease probabilities is characterized by Knightian uncertainty. Patients are assumed to be Choquet-expected utility maximizers with beliefs that are represented by so-called non-extreme outcome capacities.

In a first step, I derive conditions for the existence and and uniqueness of interior and corner solutions of the underlying optimization problem. Subsequently, I conduct a comparative static analysis with respect to the pessimism parameter α and the confidence parameter δ. It turns out that the effect of optimism on preventive activities depends on two concurrent effects which are denoted as "perceived efficacy effect" and "expected marginal utility effect". The perceived efficacy effect covers the fact that optimistic patients might judge the preventive regime's capability to reduce the underlying probability of disease differently from pessimistic patients. This is captured by the slope of the effort-dependent distorted probability function $\pi_{\mathbb{CEU}}(V)$. The sign of the effect depends on the effectiveness ranking between the worst-case relationship π_{min} and the best-case relationship π_{max}. For instance, if π_{max} is less effective than π_{min}, we can infer that optimistic patients deem the preventive regime less capable of reducing the disease probability than pessimistic ones. In this case, the perceived efficacy effect is negative and optimistic patients reduce preventive activities. The expected marginal utility effect captures how patients' expected marginal utility changes as they become more optimistic. It is negative as long as patients prefer to exchange a marginal unit of prevention in the bad health state with a marginal unit of prevention in the good health state. The overall effect of optimism on prevention is determined by the sum of both individual effects.

Variation in the confidence parameter δ can be analyzed in a similar fashion. Again, there are two concurrent effects, termed "δ-perceived efficacy effect" and "δ-expected marginal utility effect", with similar interpretations. The δ-perceived efficacy effect captures the fact that confidence variations might entail a shift in the assessment of the preventive regime's capability to reduce the probability of disease. Similarly, the expected marginal utility effect captures how patients' expected marginal utility from prevention changes as they become less confident. Again, the overall effect is given by the sum of both individual effects.

As a next step, I introduce information in the form of a random signal provided by the physician. Using the pessimistic, optimistic and generalized Bayesian updating rule for non-extreme outcome capacities, I examine how preventive activities of Knightian patients relate to preventive activities of Bayesian patients whose beliefs are represented by a standard subjective probability. It turns out that, after observing the signal, preventive activities of Knightian patients are consistently lower than those of Bayesian patients under the optimistic updating rule if the best-case relationship π_{max} has a lower perceived effectiveness than the updated Bayesian relationship π_q^{Bayes}. The same holds for the generalized Bayesian updating rule if the extreme-outcome relationship π_α is less effective than π_q^{Bayes} and patients are sufficiently optimistic. Preventive activities of Knightian patients exceed those of Bayesian patients for the pessimistic updating rule if the worst-case relationship π_{min} is more effective than the updated Bayesian relationship π_q^{Bayes}. The same result holds for the generalized Bayesian updating rule if π_q^{Bayes} is less effective than π_α and patients are sufficiently pessimistic. As a next step, I introduce the term "preventive gap" as the difference between Bayesian and Knightian patients with respect to their preventive activities. It turns out that information can increase or decrease the preventive gap between Bayesian and Non-Bayesian patients.

Finally, I compare patients to an important benchmark case, which is a situation where the true underlying relationship π_θ is known. "excessive preventive behavior" describes a situation where patients exert a higher level of effort than in the benchmark case. Similarly, the term "preventive inertia" refers to a situation where patients exert a lower level of effort than in the benchmark case. It turns out that information can reinforce or attenuate excessive preventive behavior and preventive inertia. This observation has important policy implications, since it demonstrates that extensive information campaigning potentially reinforces extreme preventive behavior among patients. This is a problematic finding, since it questions the justification of information campaigns, at least for certain subgroups of patients.

3.6 Mathematical Proofs

Throughout this section, I write u' and refer to the partial derivative $\frac{\partial u}{\partial V}$. Similarly, I write u'' and refer to the second-order partial derivative $\frac{\partial^2 u}{\partial V^2}$ if not specified differently.

Proof of Corollary 3.1. The proof is straightforward. One obtains

$$
\begin{aligned}
E_q[f(\cdot|V)] &= \sum_{i=1}^{n} q_i \Big\{ \pi_i(V) u(W_1(V), h_1) + (1 - \pi_i(V)) u(W_2(V), h_2) \Big\} \\
&= u(W_1(V), h_1) \cdot \Big(\sum_{i=1}^{n} q_i \pi_i(V) \Big) \\
&+ u(W_2(V), h_2) \cdot \Big(1 - \sum_{i=1}^{n} q_i \pi_i(V) \Big) \\
&= \pi_q(V) u(W_1(V), h_1) + (1 - \pi_q(V)) u(W_2(V), h_2).
\end{aligned}
$$

Defining

$$
\begin{aligned}
\pi_{max}(V) &:= \arg\max_{\pi_i \in \Phi} \Big\{ \pi_i(V) u(W_1(V), h_1) + (1 - \pi_i(V)) u(W_2(V), h_2) \Big\} \\
\pi_{min}(V) &:= \arg\min_{\pi_i \in \Phi} \Big\{ \pi_i(V)\ u(W_1(V), h_1) + (1 - \pi_i(V))\ u(W_2(V), h_2) \Big\},
\end{aligned}
$$

one can express Z_{min} and Z_{max} as

$$
\begin{aligned}
Z_{max}(V) &= \pi_{max}(V) u(W_1(V), h_1) + (1 - \pi_{max}(V)) u(W_2(V), h_2) \\
Z_{min}(V) &= \pi_{min}(V) u(W_1(V), h_1) + (1 - \pi_{min}(V)) u(W_2(V), h_2)
\end{aligned}
$$

With this notation, the objective can be rewritten as

$$
\mathcal{U}(V|\alpha, \delta) = \pi_{\mathbb{CEU}}(V|\alpha, \delta) u(W_1(V), h_1) + (1 - \pi_{\mathbb{CEU}}(V|\alpha, \delta)) u(W_2(V), h_2) \tag{3.18}
$$

where

$$
\pi_{\mathbb{CEU}}(V|\alpha, \delta) = (1 - \delta) \pi_q(V) + \delta(\alpha \pi_{max}(V) + (1 - \alpha) \pi_{min}(V)) \tag{3.19}
$$

□

Proof of Proposition 3.1. The objective is continuous if the functions $E_q[f(\cdot|V)]$, Z_{min} and Z_{max} are continuous. By Corollary 3.1, it follows that

$$
E_q[f(\cdot|V)] = \pi_q(V) u(W_1(V), h_1) + (1 - \pi_q(V)) u(W_2(V), h_2).
$$

The function π_q is continuous since it is a sum of continuous functions. Since u is continuous, we can conclude that $E_q[f(\cdot|V)]$ is continuous, since sums and products of continuous functions are continuous. Hence, what remains to be shown is the continuity of the functions Z_{min} and Z_{max}. This can be shown because the minimum and the maximum of continuous functions is again continuous. Let $(V_n)_{n\in\mathbb{N}}$ be a sequence in $[0,1]$ with $V_n \to V^*$ for $n \to \infty$. Due to the fact that $[0,1]$ is closed, it follows that $V^* \in [0,1]$. Since π_i is continuous for $i = 1, ..., n$, we can infer that $\pi_i(V_n) \to \pi_i(V^*)$ for $n \to \infty$. This implies that $\pi(V_n)$ converges to $\pi_i(V^*)$ for all $i = 1, .., n$. If all sequences converge, we can conclude that the sequence of minima and the sequence of maxima also converge. Moreover, the limit of $\min\{\pi_1(V_n), ..., \pi_n(V_n)\}$, or of $\max\{\pi_1(V_n), ..., \pi_n(V_n)\}$ must be contained in the set of limiting values $\{\pi_1(V^*), ..., \pi_n(V^*)\}$. Otherwise, we could find a natural number $n_0 \in \mathbb{N}$ such that $\min\{\pi_1(V_n), ..., \pi_n(V_n)\} \notin \{\pi_1(V_n), ..., \pi_n(V_n)\}$ for $n \geq n_0$. This is a contradiction proving that Z_{min} and Z_{max} are continuous.

Since $[0,1]$ is compact, and since the objective $\mathcal{U}(V|\alpha,\delta)$ is continuous on $[0,1]$. We can use Weierstrass' theorem to show that a solution of $\max\limits_{V\in[0,1]} \mathcal{U}(V|\alpha,\delta)$ exists. □

Proof of Proposition 3.2. In this proof, I call a function C^2 if it is twice continuously differentiable on its domain. The objective $\mathcal{U}$ is C^2 if the functions $E_q[f(\cdot|V)]$, Z_{min} and Z_{max} are C^2. By Corollary 3.1, we obtain

$$E_q[f(\cdot|V)] = \pi_q(V)u(W_1(V), h_1) + (1 - \pi_q(V))u(W_2(V), h_2).$$

The function π_q is C^2, since it is a sum of C^2-functions. Since u is C^2, we can conclude that $E_q[f(\cdot|V)]$ is C^2, because sums and products of C^2-functions are C^2. Hence, what remains to be shown is that Z_{min} and Z_{max} are C^2 when there are now crossing points and at least piecewise continuously differentiable if the sets C_{prob} and $C_{utility}$ are finite. Consider first a point $\hat{V} \in [0,1]$ that is not a crossing point of functions in Φ or the utilities $u(W_1(V), h_1)$ and $u(W_2(V), h_2)$. This means $\hat{V} \notin C_{prob} \cup C_{utility}$. Then, there is a $\delta > 0$ such that

$$\pi_{i_1}(V) > \pi_{i_2}(V)... > \pi_{i_n}(V)$$

for all $V \in (\hat{V} - \delta, \hat{V} + \delta)$ where $i_j \in \{1, ..., n\}$ for $j \in \{1, ..., n\}$. This means we can find a neighborhood of points around $\hat{V}$ such that the ordering between the functions $\pi_i \in \Phi$ remains stable in this neighborhood. Now, we can differentiate between the following cases:

(a) $u(W_1(\hat{V}), h_1) > u(W_2(\hat{V}), h_2)$

(b) $u(W_1(\hat{V}), h_1) = u(W_2(\hat{V}), h_2)$

(c) $u(W_1(\hat{V}), h_1) < u(W_2(\hat{V}), h_2)$

Consider first case (a). Since u is continuous, we can conclude that there is $\varepsilon_1 > 0$ such that $u(W_1(V), h_1) > u(W_2(V), h_2)$ for all $V \in (\hat{V} - \varepsilon_1, \hat{V} + \varepsilon_1)$. This implies

$$Z_{min}(V) = \pi_{i_n}(V)u(W_1(V), h_1) + (1 - \pi_{i_n}(V))u(W_2(V), h_2)$$

for all $V \in (\hat{V} - \min\{\varepsilon_1, \delta\}, \hat{V} + \min\{\varepsilon_1, \delta\})$. This shows that Z_{min} is, as sum and product of C^2-functions, C^2 on $(\hat{V} - \min\{\varepsilon_1, \delta\}, \hat{V} + \min\{\varepsilon_1, \delta\})$, and therefore also C^2 in $\hat{V}$. A similar argument holds for case (c) with the difference that there is an $\varepsilon_2 > 0$ such that $u(W_1(V), h_1) < u(W_2(V)h_2)$ for all $V \in (\hat{V} - \varepsilon_2, \hat{V} + \varepsilon_2)$. Hence,

$$Z_{min}(V) = \pi_{i_1}(V)u(W_1(V), h_1) + (1 - \pi_{i_1}(V))u(W_2(V), h_2)$$

for all $V \in (\hat{V} - \min\{\varepsilon_2, \delta\}, \hat{V} + \min\{\varepsilon_2, \delta\})$. Again, we can conclude that Z_{min} is C^2 on $(\hat{V} - \min\{\varepsilon_2, \delta\}, \hat{V} + \min\{\varepsilon_2, \delta\})$. In case (b), three subcases can occur.

(d) there is $\varepsilon_3 > 0$ such that $u(W_l(V), h_l) < u(W_k(V), h_k)$ for all $V \in (\hat{V} - \varepsilon_3, \hat{V} + \varepsilon_3)\backslash\{\hat{V}\}$ and $l, k \in \{1, 2\}$ with $l \neq k$

(e) there is $\varepsilon_4 > 0$ such that $u(W_l(V), h_l) = u(W_k(V), h_k)$ for all $V \in (\hat{V} - \varepsilon_4, \hat{V} + \varepsilon_4)$ and $l, k \in \{1, 2\}$ with $l \neq k$

(f) there is $\varepsilon_5 > 0$ such that $u(W_l(V), h_l) < u(W_k(V), h_k)$ for all $V \in (\hat{V} - \varepsilon_5, \hat{V})$ and $u(W_l(V), h_l) > u(W_k(V), h_k)$ for all $V \in (\hat{V}, \hat{V} + \varepsilon_5)$

Subsequently, I consider w.l.o.g. the case $l = 1$ and $k = 2$. In case (d), we can conclude that

$$Z_{min}(V) = \pi_{i_1}(V)u(W_1(V), h_1) + (1 - \pi_{i_1}(V))u(W_2(V), h_2)$$

for all $V \in (\hat{V} - \min\{\varepsilon_3, \delta\}, \hat{V} + \min\{\varepsilon_3, \delta\})$. Hence, Z_{min} is C^2 in $\hat{V}$. In case (e), we have

$$Z_{min}(V) = u(W_1(V), h_1) = u(W_2(V), h_2)$$

for all $V \in (\hat{V} - \min\{\varepsilon_4, \delta\}, \hat{V} + \min\{\varepsilon_4, \delta\})$. As in the previous cases, we can conclude that Z_{min} is C^2 in $\hat{V}$. In the remaining case (f), we can infer that $\hat{V}$ is a crossing point of the utility functions $u(W_1(V), h_1)$ and $u(W_2(V), h_2)$. Formally, $\hat{V} \in C_{utilities}$. Since $C_{utilities} = \emptyset$ by assumption, we can exclude case (f). Hence, Z_{min} is C^2 in $\hat{V}$. The same proof applies for the function Z_{max} with the difference that one needs to replace the minimizing relationship π_{min} with the respective maximizing relationship π_{max} for each of the cases (a) to (f). Since $C_{prob} \cup C_{utility}$ is a finite set, we obtain that the objective is C^2 up to finitely many points. If $C_{prob} \cup C_{utility} = \emptyset$, the objective is C^2. □

Proof of Proposition 3.3. Due to Requirement (SC_1), Proposition 3.2 shows that the objective is twice continuously differentiable. The second order condition of $\mathcal{U}$ is given by:

$$
\begin{aligned}
\frac{\partial^2 \mathcal{U}}{\partial V^2} =& \pi''_{\mathbb{CEU}}(V|\alpha,\delta)\cdot(u(W_1(V),h_1)-u(W_2(V),h_2)) \\
&+2\pi'_{\mathbb{CEU}}(V|\alpha,\delta)(W_1'(V)u'(W_1(V),h_1)-W_2'(V)u'(W_2(V),h_2)) \\
&+\pi_{\mathbb{CEU}}(V|\alpha,\delta)W_1''(V)u'(W_1(V),h_1)+(1-\pi_{\mathbb{CEU}}(V|\alpha,\delta))W_2''(V)u'(W_2(V),h_2) \\
&+\pi_{\mathbb{CEU}}(V|\alpha,\delta)W_1'(V)^2u''(W_1(V),h_1)+(1-\pi_{\mathbb{CEU}}(V|\alpha,\delta))W_2'(V)^2u''(W_2(V),h_2)
\end{aligned}
\tag{3.20}
$$

By Requirement (SC_2), we obtain $u(W_2(V),h_2) > u(W_1(V),h_1)$. Together with Requirement (SC_3), it follows that

$$\pi''_{\mathbb{CEU}}(V|\alpha,\delta)(u(W_1(V),h_1)-u(W_2(V),h_2)) < 0 \tag{3.21}$$

By using Requirement (SC_5) and the fact that $\pi'_{\mathbb{CEU}} < 0$, we obtain

$$2\pi'_{\mathbb{CEU}}(V|\alpha,\delta)\cdot(W_1'(V)u'(W_1(V),h_1)-W_2'(V)u'(W_2(V),h_2)) \le 0. \tag{3.22}$$

Due to Requirement (SC_4), we can infer

$$\pi_{\mathbb{CEU}}(V|\alpha,\delta)W_1''(V)u'(W_1(V),h_1)+(1-\pi_{\mathbb{CEU}}(V|\alpha,\delta))W_2''(V)u'(W_2(V),h_2) \le 0. \tag{3.23}$$

Furthermore, we can conclude from $u'' \le 0$ and $\pi_{\mathbb{CEU}}(V|\alpha,\delta) \in [0,1]$ that

$$\pi_{\mathbb{CEU}}(V|\alpha,\delta)W_1'(V)^2u''(W_1(V),h_1)+(1-\pi_{\mathbb{CEU}}(V|\alpha,\delta))W_2'(V)^2u''(W_2(V),h_2) \le 0. \tag{3.24}$$

Due to the inequalities (3.21), (3.22), (3.23), and (3.24), the claim is proved. □

Proof of Proposition 3.4. The Lagrangian is of the form

$$L(V,\lambda_1,\lambda_2) := \pi_{\mathbb{CEU}}(V|\alpha,\delta)u(W_1(V),h_1)+(1-\pi_{\mathbb{CEU}}(V|\alpha,\delta))u(W_2(V),h_2)+\lambda_1V+\lambda_2(1-V)$$

The first order condition
$$\frac{\partial L}{\partial V}(V_0) = 0$$

yields
$$\frac{\partial \mathcal{U}}{\partial V}(V_0|\alpha, \delta) + \lambda_1 - \lambda_2 = 0.$$

Furthermore, we obtain the non-negativity conditions $\lambda_i \geq 0$ for $i = 1, 2$, the inequality constraints $V \geq 0$ and $V \leq 1$, as well as, the following complementary slackness conditions:

$$\lambda_1(-V_0) = 0$$
$$\lambda_2(V_0 - 1) = 0$$

In order to solve the optimization problem, four different cases must be considered.

1. $\lambda_1 > 0$ and $\lambda_2 > 0$: In this case, we have $V_0 = 0$ and $V_0 = 1$, which is a contradiction.

2. $\lambda_1 = 0$ and $\lambda_2 = 0$. This is the case of an interior solution; the first order condition is equivalent to $\frac{\partial \mathcal{U}}{\partial V} = 0$.

3. $\lambda_1 = 0$ and $\lambda_2 > 0$: In this case, we can conclude $V_0 = 1$. Plugging these conditions into the first order condition, we obtain in the limit

$$\lim_{V_0 \to 1^-} \frac{\partial \mathcal{U}}{\partial V}(V_0|\alpha, \delta) - \lambda_2 = 0.$$

   Since λ_2 is required to be strictly positive, we need to rely on the condition

$$\lim_{V_0 \to 1^-} \frac{\partial \mathcal{U}}{\partial V}(V_0|\alpha, \delta) > 0$$

   to guarantee that the corner solution $V_0 = 1$ is a feasible solution.

4. $\lambda_1 > 0$ and $\lambda_2 = 0$: From the complementary slackness conditions, we can conclude that $V_0 = 0$. Resorting to an argument similar to the one in the previous case, we

obtain the following limit for the first order condition:

$$\lim_{V_0 \to 0^+} \frac{\partial \mathcal{U}}{\partial V}(V_0|\alpha, \delta) + \lambda_1 = 0$$

Since λ_1 is assumed to be strictly positive, it follows that the condition

$$\lim_{V_0 \to 0^+} \frac{\partial \mathcal{U}}{\partial V}(V_0|\alpha, \delta) < 0$$

needs to be satisfied in order to guarantee that $V_0 = 0$ is a feasible corner solution.

It remains to be shown that there can never be more than one corner solution. Assume that the conditions $\lim_{V_0 \to 1^-} \frac{\partial \mathcal{U}}{\partial V}(V_0|\alpha, \delta) > 0$ and $\lim_{V_0 \to 0^+} \frac{\partial \mathcal{U}}{\partial V}(V_0|\alpha, \delta) < 0$ hold at the same time. Since $\mathcal{U}$ is continuously differentiable, we can deduce that the derivative $\mathcal{U}_V$ is continuous on $[0, 1]$. By the intermediate value theorem, we can conclude that there exists $\hat{V} \in (0, 1)$ such that $U_V(\hat{V}) = 0$. Furthermore, the derivative changes its sign from negative to positive at $\hat{V}$. Therefore $\hat{V}$ must be a minimizer. This is a contradiction to the assumption that $\mathcal{U}$ is strictly concave. □

Proof of Proposition 3.5. Consider two preventive relationships π and $\hat{\pi}$, where $\hat{\pi}$ is more effective than π. The derivative of

$$\pi(V)u(W_1(V), h_1) + (1 - \pi(V))u(W_2(V), h_2)$$

with respect to V is given by

$$\begin{aligned} \frac{\partial \mathcal{U}}{\partial V} &= \pi'(V)[u(W_1(V), h_1) - u(W_2(V), h_2)] \\ &\quad + \pi(V)W_1'(V)u'(W_1(V), h_1) + (1 - \pi(V))W_2'(V)u'(W_2(V), h_2). \end{aligned}$$

I define

$$g(V) := \pi'(V)[u(W_1(V), h_1) - u(W_2(V), h_2)]$$

and

$$
\begin{aligned}
b(V) :=&\pi(V)W_1'(V)u'(W_1(V),h_1) + (1-\pi(V))W_2'(V)u'(W_2(V),h_2)\\
=&W_2'(V)u'(W_2(V),h_2) + \pi(V)\left(W_1'(V)u'(W_1(V),h_1) - W_2'(V)u'(W_2(V),h_2)\right)\\
=&W_2'(V)u'(W_2(V),h_2) + \pi(V)\Delta_u'
\end{aligned}
$$

Since $\Delta_u = u(W_2(V),h_2) - u(W_1(V),h_1) \geq 0$ and $\pi'(V) < 0$, we can conclude $g(V) \geq 0$. Similarly, it follows that $b(V) \leq 0$. Due to the fact that $\Delta_u' \geq 0$ by Condition (SC_5) and $\hat{\pi}(V) \geq \pi(V)$ for all $V \in [0,1]$, we can conclude that

$$
b(V) \leq \hat{\pi}(V)W_1'(V)u'(W_1(V),h_1) + (1-\hat{\pi}(V))W_2'(V)u'(W_2(V),h_2).
$$

Moreover, since $\pi'(V) \geq \hat{\pi}'(V)$ for all $V \in [0,1]$, it follows that

$$
g(V) \leq \hat{\pi}'(V)[u(W_1(V),h_1) - u(W_2(V),h_2)].
$$

Consequently,

$$
g(V) + b(V) \leq \frac{\partial \mathcal{U}^{\hat{\pi}}}{\partial V}
$$

where $\frac{\partial \mathcal{U}^{\hat{\pi}}}{\partial V}$ denotes the derivative of

$$
\hat{\pi}(V)u(W_1(V),h_1) + (1-\hat{\pi}(V))u(W_2(V),h_2)
$$

with respect to V. As a consequence, we obtain

$$
\frac{\partial \mathcal{U}}{\partial V} \leq \frac{\partial \mathcal{U}^{\hat{\pi}}}{\partial V}.
$$

Since for an interior maximum, the derivative $\frac{\partial \mathcal{U}}{\partial V}$ changes its sign from positive to negative, we can conclude that preventive effort given π is smaller than prevention under relationship $\hat{\pi}$.

If $V^* = 0$ is the global maximizer of $\mathcal{U}$ on $[0,1]$, we can deduce that $\mathcal{U}$ is strictly decreasing in V. In this case there are two possibilities. The first one is that the partial derivative of $\mathcal{U}^{\hat{\pi}}$ is still strictly decreasing in $V = 0$. Then, the solution of $\max\limits_{V\in[0,1]} \mathcal{U}^{\hat{\pi}}$ is still given by $V^* = 0$. Hence, a higher effectiveness has no effect on preventive activities. The second possibility is that $\mathcal{U}^{\hat{\pi}}$ is not strictly decreasing in $V = 0$ anymore. In this case, the first order condition has a zero $V^* \in (0,1)$. Since $\mathcal{U}^{\hat{\pi}}$ is strictly concave, V^* is the unique global maximizer. If $V^* = 1$ is the global maximizer of $\mathcal{U}$ on $[0,1]$, we obtain that $\mathcal{U}$ is strictly increasing on $[0,1]$. Besides, we have $0 \leq \frac{\partial \mathcal{U}}{\partial V} \leq \frac{\partial \mathcal{U}^{\hat{\pi}}}{\partial V}$. Hence, $\mathcal{U}^{\hat{\pi}}$ is still strictly increasing in V. Thus, $V^* = 1$ remains the solution. □

Proof of Corollary 3.2. It is

$$\Delta_1 = \Delta_u \frac{d^2}{d\alpha\, dV} \pi_{\mathbb{CEU}} = \Delta_u \delta \left[\pi'_{max}(V) - \pi'_{min}(V)\right]$$

By assumption, we have $\Delta_u < 0$ and $\pi'_{max}(V) - \pi'_{min}(V) > 0$. Therefore, we can conclude that $\Delta_1 < 0$ for $\delta > 0$ and $\Delta_1 = 0$ for $\delta = 0$. □

Proof of Proposition 3.6. The proof is straightforward. The result of part (a) is trivial since the objective is independent of α in cases of full confidence. If the agent gives a positive weight to extreme outcomes, the parameter α influences prevention. This is the case for part (b). As α increases, the patient becomes more optimistic. In cases where π_{min} has a higher perceived effectiveness than π_{max}, we can conclude that the belief function $\pi_{\mathbb{CEU}}(\cdot|\alpha,\delta)$ shifts towards another belief function $\pi_{\mathbb{CEU}}(\cdot|\alpha',\delta)$ which has a lower perceived effectiveness. By Corollary 3.2, we obtain $\Delta_1 \leq 0$. Moreover, Δ_2 is negative due to Requirement (SC_5). Together, this implies that prevention either remains unchanged or decreases. Where there is an interior solution, we have $\Delta_1 < 0$. Where there is a corner solution, we have $\Delta_1 = 0$ for $V^* = 0$ and $\Delta_1 \leq 0$ for $V^* = 1$. On the other hand, if π_{max} is more effective than π_{min}, we can deduce that $\Delta_1 \geq 0$. Hence, the overall effect is positive if $\Delta_1 > \Delta_2$, negative if $\Delta_1 < \Delta_2$, and zero if $\Delta_1 = \Delta_2$. □

Proof Corollary 3.3. It is

$$\Delta_3 = \Delta_u \cdot \frac{\partial^2 \pi_{\mathbb{CEU}}(V|\alpha,\delta)}{\partial\delta\partial V} = \Delta_u \cdot \big(-\pi_q'(V) + \alpha\pi_1'(V) + (1-\alpha)\pi_n'(V)\big)$$

Consequently, it is necessary to differentiate between the following cases:

(1) $\Delta_3 > 0$ iff $\alpha\pi_1'(V) + (1-\alpha)\pi_n'(V) < \pi_q'(V)$

(2) $\Delta_3 < 0$ iff $\alpha\pi_1'(V) + (1-\alpha)\pi_n'(V) > \pi_q'(V)$

(3) $\Delta_3 = 0$ iff $\alpha\pi_1'(V) + (1-\alpha)\pi_n'(V) = \pi_q'(V)$

Now, the statement of the corollary follows directly from Definition 3.3. □

Proof of Corollary 3.5. Due to Condition (SC_5), we can infer that $\Delta_u' \geq 0$. Therefore, the sign of the δ-expected marginal utility effect is determined by the sign of $\frac{d}{d\delta}\pi_{\mathbb{CEU}}$. It is

$$\text{sign}\left(\frac{d}{\delta}\pi_{\mathbb{CEU}}\right) = \begin{cases} +1 & \text{for} \quad \alpha < \hat{\alpha}(V^*) \\ 0 & \text{for} \quad \alpha = \hat{\alpha}(V^*) \\ -1 & \text{for} \quad \alpha = \hat{\alpha}(V^*). \end{cases}$$

□

Proof of Corollary 3.6. Since neo-additive capacities remain neo-additive with revised parameters α' and δ', we can replace the ex-ante capacity $\nu(\alpha,\delta)$ with the updated capacity $\nu(\alpha',\delta')$. By using Corollary 3.1, we can see that patients maximize

$$\mathcal{U}(V|\alpha',\delta') = \pi_{\mathbb{CEU}}(V|\alpha',\delta')u(W_1(V),h_1) + (1-\pi_{\mathbb{CEU}}(V|\alpha',\delta'))u(W_2(V),h_2)$$

where

$$\pi_{\mathbb{CEU}}(V|\alpha',\delta') = (1-\delta')\pi_q(V) + \delta'(\alpha'\pi_{max}(V) + (1-\alpha)\pi_{min}(V))$$

with rule-dependent α' and δ'. Replacing these parameters by α^O and δ^O in case of the optimistic updating rule, by α^P and δ^P in case of the pessimistic updating rule, and by α^{GB} and δ^{GB} in case of the generalized Bayesian updating rule concludes the proof. □

Proof of Proposition 3.8. In this proof, each updating rule is examined separately.

Optimistic Updating Rule

Under the optimistic updating rule, the Non-Bayesian patient maximizes the objective with respect to the preventive relationship

$$\pi^O_{\mathbb{CEU}}(V) = (1-\delta^O)\pi_q^{Bayes}(V) + \delta^O \pi_{max}(V).$$

First of all, note that

$$\pi_q^{Bayes} - (1-\delta^O)\pi_q^{Bayes}(V) - \delta^O\pi_{max}(V) = \delta^O(\pi_q^{Bayes}(V) - \pi_{max}(V)) \geq 0.$$

Since

$$(\pi_q^{Bayes})'(V) \leq (1-\delta^O)\pi_q^{Bayes}(V)' + \delta^O\pi_{max}(V)',$$

if π_{max} is less effective than π_q^{Bayes}, we can conclude that $\pi^O_{\mathbb{CEU}}$ is less effective than π_q^{Bayes}. By using Proposition 3.5, we obtain the claims for the optimistic updating rule.

Pessimistic Updating Rule

Similarly, for the pessimistic updating rule, the Non-Bayesian agent relies on the relationship

$$\pi^P_{\mathbb{CEU}}(V) = (1-\delta^O)\pi_q^{Bayes}(V) + \delta^O\pi_{min}(V).$$

It is

$$\pi_q^{Bayes} - (1-\delta^O)\pi_q^{Bayes}(V) - \delta^O\pi_{min}(V) = \delta^O(\pi_q^{Bayes}(V) - \pi_{min}(V)) \leq 0.$$

Since π_{min} is more effective than π_q, we can infer that

$$(\pi_q^{Bayes})'(V) \geq (1-\delta^P)\pi_q^{Bayes}(V)' + \delta^P \pi_{min}(V)'.$$

Hence, π_q^{Bayes} is less effective than $\pi_{\mathbb{CEU}}^P$. Again, by using Proposition 3.5, we can infer the claim for the pessimistic updating rule.

Generalized Bayesian Updating Rule

In case of the generalized Bayesian updating rule, the Non-Bayesian patient maximizes an expected utility with respect to the distorted probability

$$\pi_{\mathbb{CEU}}^{GB}(V) = (1-\delta^{GB})\pi_q^{Bayes}(V) + \delta^{GB}(\alpha\pi_{min}(V) + (1-\alpha\pi_{max}(V)).$$

Again, we can consider the difference

$$\begin{aligned}
&\pi_q^{Bayes}(V) - (1-\delta^{GB})\pi_q^{Bayes}(V) - \delta^{GB}(\alpha\pi_{max}(V) + (1-\alpha)\pi_{min}(V)) \\
&= (1-\delta^{GB})(\pi_q^{Bayes}(V) - \alpha\pi_{max}(V) - (1-\alpha)\pi_{min}(V))
\end{aligned}$$

As a consequence, we can distinguish three possible cases:

(1) $\pi_q^{Bayes}(V) > \pi_\alpha(V)$

(2) $\pi_q^{Bayes}(V) = \pi_\alpha(V)$

(3) $\pi_q^{Bayes}(V) < \pi_\alpha(V)$

In case (1), we have $\pi_q^{Bayes} > \pi_{\mathbb{CEU}}^{GB}$, in case (2), we can infer $\pi_q^{Bayes} = \pi_{\mathbb{CEU}}^{GB}$, and in case (3), we can deduce that $\pi_q^{Bayes} < \pi_{\mathbb{CEU}}^{GB}$. The remaining part of the proof follows directly from Proposition 3.5. □

Proof of Corollary 3.7. Remember that the condition $\pi_q^{Bayes}(V) < \pi_\alpha(V)$ is fulfilled when patients are sufficiently pessimistic with $\alpha < \hat{\alpha}(V)$. Similarly, the condition $\pi_q^{Bayes}(V) >$

$\pi_\alpha(V)$ is fulfilled when patients are sufficiently optimistic with $\alpha > \hat{\alpha}(V)$. Hence, a clear ordering between π_q and π_α is possible whenever $\alpha < \hat{\alpha}(V)$ for all $V \in [0,1]$ or $\alpha > \hat{\alpha}(V)$ for all $V \in [0,1]$. Defining

$$\alpha_{min} := \min_{V\in[0,1]} \hat{\alpha}(V) \quad \text{and} \quad \alpha_{max} := \max_{V\in[0,1]} \hat{\alpha}(V),$$

we obtain $\pi_q(V) < \pi_\alpha(V)$ for all $V \in [0,1]$ if and only if $\alpha < \alpha_{min}$. Similarly, $\pi_q(V) > \pi_\alpha(V)$ for all $V \in [0,1]$ if and only if $\alpha > \alpha_{max}$. □

3.7 Examples

Example 3.5. Let $\Phi = \{\pi_1, \pi_2\, \pi_3\}$ with $\pi_1(V) = 1 - \frac{1}{2}V$, $\pi_2(V) = 1 - V^2$, and $\pi_3(V) = 1 - V$. Obviously, each π_i is twice continuously differentiable and strictly decreasing on $[0,1]$. Figure 3.10 illustrates these functions graphically in one diagram.

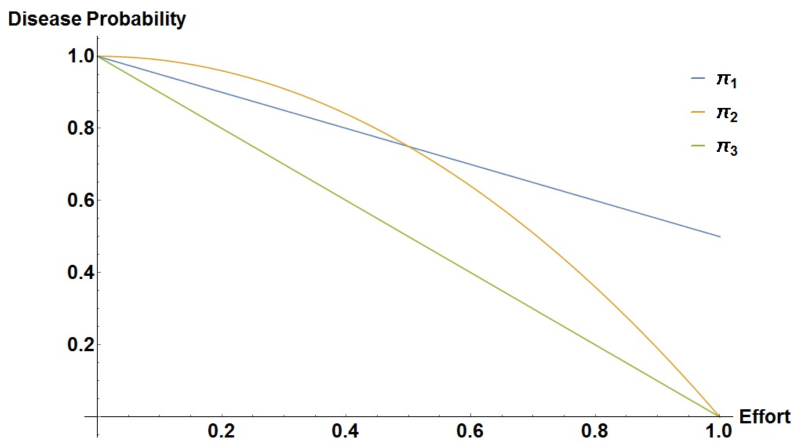


FIGURE 3.10: Preventive Relationships with a Single Crossing Point

At the point $\hat{V} = \frac{1}{2}$, π_1 and π_2 intersect. Moreover, we can see that π_1 is smaller than π_2 for all $0 \leq V < \hat{V}$ and larger than π_2 for all $\hat{V} < V \leq 1$. Assume furthermore, that the utility in the bad state is always lower than the utility in the good state

$$u(W_1(V), h_1) < u(W_2(V), h_2)$$

for all $V \in [0,1]$. Then, we obtain for π_{min}:

$$\pi_{min}(V) = \begin{cases} 1 - V^2 & \text{for} \quad 0 \leq V < \hat{V} \\ \frac{3}{4} & \text{for} \quad V = \hat{V} \\ 1 - \frac{1}{2}V & \text{for} \quad V > \hat{V} \end{cases}$$

For π_{max}, we have $\pi_{max} = \pi_3$. This implies that π_{max} is continuously differentiable. Obviously, π_{min} is not continuously differentiable due to a kink at $V = \frac{1}{2}$. Formally,

$$\lim_{V \to \hat{V}^-} \frac{d}{dV} \pi_{min}(V) = -1 \neq \lim_{V \to \hat{V}^+} \frac{d}{dV} \pi_{min}(V) = -\frac{1}{2}.$$

This implies that the objective is not necessarily continuously differentiable at $\hat{V} = \frac{1}{2}$. This becomes clear, as we consider the special case $\alpha = 0$, $\delta = 1$, $u(w,h) = w$, $W_2(V) = 5 - V$, $W_2(V) = 10 - V$ and $q_1 = q_2 = \frac{1}{2}$. Then, the objective is given by

$$\mathcal{U}(V|\alpha = 0, \delta = 1) = \begin{cases} 10 - \frac{7}{2}V & \text{for} \quad 0 \leq V \leq \frac{1}{2} \\ \frac{33}{4} & \text{for} \quad V = \frac{1}{2} \\ 10 - V - 5V^2 & \text{for} \quad \frac{1}{2} < V \leq 1. \end{cases}$$

For the limit of the derivatives from the left and from the right, we obtain

$$\lim_{V \to \hat{V}^-} \frac{d}{dV} \mathcal{U}(V|\alpha = 0, \delta = 1) = -\frac{7}{2} \neq \lim_{V \to \hat{V}^+} \frac{d}{dV} \mathcal{U}(V|\alpha = 0, \delta = 1) = -6$$

This proves that $\mathcal{U}$ is not differentiable at $\hat{V} = \frac{1}{2}$.

Example 3.6. Let $\pi_1(V) = 1 - V$, $\pi_2(V) = 1 - \frac{1}{4}V$, $W_1(V) = 10 - \frac{1}{2}V$, $W_2(V) = 10 - V^5$, $u(w,h) = w$, $\delta = 1$, $q_1 = q_2 = \frac{1}{2}$, and $\alpha = 0$. Obviously, the Assumptions 1,2 and 3 are fulfilled for this model specification. Moreover, there is no crossing point of π_1 and π_2 on the domain $D = [0,1]$. The utilities $u(W_1(V), h_1) = W_1(V)$ and $u(W_2(V), h_2) = W_2(V)$

have a crossing point at $\hat{V} = \sqrt[4]{\frac{1}{2}}$. The objective function is of the form

$$\mathcal{U}(V|\alpha = 0, \delta = 1) = \pi_{min}(V)u(W_1(V), h_1) + (1 - \pi_{\min}(V))u(W_2(V), h_2)$$

where

$$\pi_{min}(V) = \begin{cases} \pi_1(V) & \text{for} \quad 0 \leq V \leq \hat{V} \\ \pi_2(V) & \text{for} \quad \hat{V} \leq V \leq 1 \end{cases}$$

Figure 3.11 gives a graphical representation of the objective function. Obviously, the

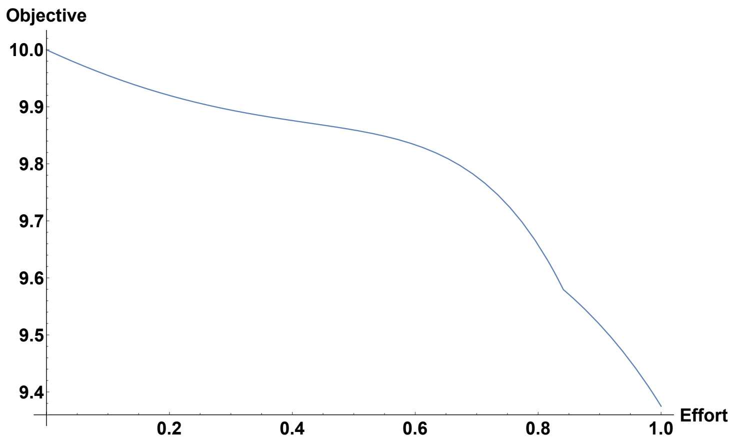


FIGURE 3.11: Objective Function with a Kink

objective has a kink at $\hat{V}$. Hence, $\mathcal{U}$ is not differentiable at the point $\hat{V}$.

Example 3.7. Throughout this example, assume that

$$u(W_1(V), h_1) = 10 - 2V^3 \quad \text{and} \quad u(W_2(V), h_2) = 20 - 2V^3$$

as well as $q_1 = q_2 = \frac{1}{2}$. I contemplate three scenarios. For each scenario, I compare the following parameter constellations:

(1) Full Confidence: $\delta = 0$

(2) Extreme Optimism: $\delta = 1$ and $\alpha = 1$

(3) Extreme Pessimism: $\delta = 1$ and $\alpha = 0$.

Furthermore, I specify for each scenario a different set of belief functions Φ_i for $i = 5, 6, 7$.

Scenario I

Let the set of preventive relationships be given by $\Phi_5 = \{\pi_9, \pi_{10}\}$ where $\pi_9(V) = 1 - \frac{2}{15}V$ and $\pi_{10}(V) = \frac{1}{2} - \frac{7}{15}V$. Figure 3.12 displays the objective for each scenario.

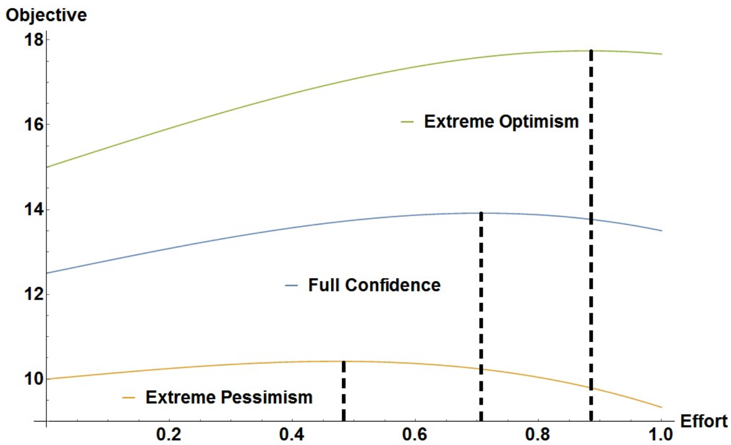


FIGURE 3.12: Confidence Increases Preventive Activities for Pessimists

Observe that a higher degree of confidence[20] yields lower preventive activities when patients are optimistic, and more preventive activities when patients are pessimistic. The following model specification demonstrates the converse result.

Scenario II

Let $\Phi_6 = \{\pi_{11}, \pi_{12}\}$ where $\pi_{11}(V) = 1 - \frac{7}{15}V$ and $\pi_{12}(V) = \frac{1}{2} - \frac{3}{15}V$. Figure 3.13 represents the objective for the case of full confidence, extreme optimism and extreme pessimism.

Given this new constellation of beliefs, we obtain that confidence increases preventive activities when patients are optimistic and decreases prevention when patients are pessimistic. The following model specification shows that Δ_3 can be zero.

[20]This means δ decreases.

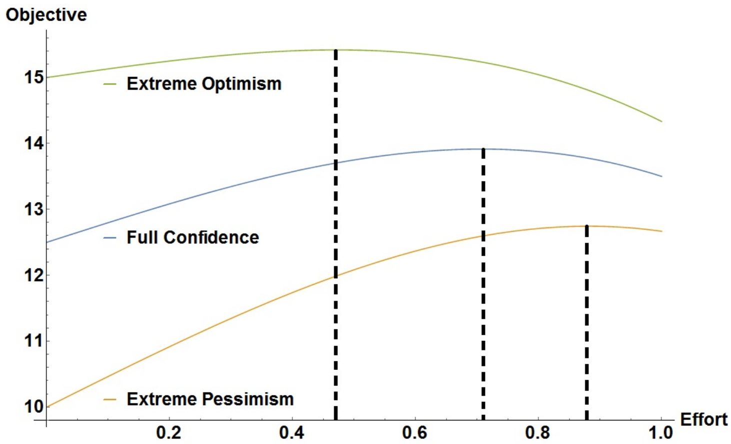


FIGURE 3.13: Confidence Increases Preventive Activities for Optimists

Scenario III

Let $\Phi_7 = \{\pi_{13}, \pi_{14}\}$ where $\pi_{13}(V) = 1 - \frac{2}{15}V$ and $\pi_{14}(V) = 1 - \frac{2}{15}V$. The following diagram displays the objective function in cases of full confidence, extreme optimism and extreme pessimism.

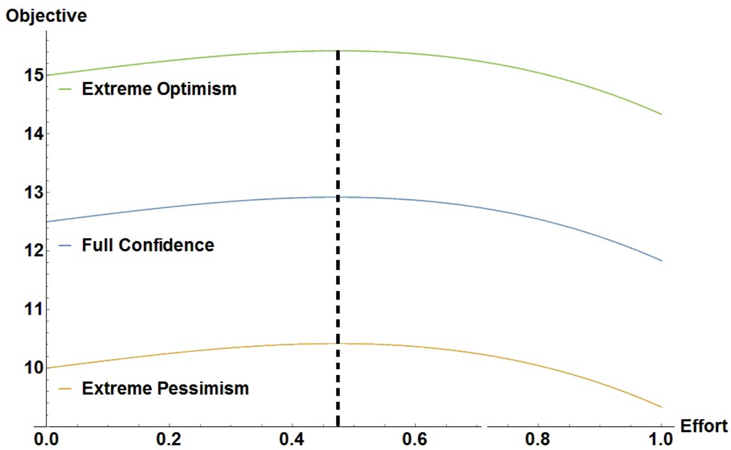


FIGURE 3.14: Confidence Has No Influence on Preventive Activities

Obviously, there is no influence of confidence on preventive activities given this constellation of parameters.

Example 3.8. Let $\pi_1(V) = 1 - \frac{3}{4}V$, $\pi_2(V) = \frac{1}{2} - \frac{1}{4}V$, and $\pi_3(V) = \frac{1}{4} - \frac{1}{5}V$. The prior q is given by $q = (0, 1, 0)$ and $\alpha = \frac{1}{2}$. Hence, we have $\pi_q = \pi_2$ and $\pi_\alpha(V) = \pi_{\frac{1}{2}}(V) = \frac{5}{8} - \frac{19}{40}V$. Figure 3.15 displays the extreme outcome function $\pi_{\frac{1}{2}}$ and the reference function π_q in one diagram. Obviously, $\pi_{\frac{1}{2}}$ and π_q have a crossing point at $V_0 = \frac{5}{9}$. As a consequence,

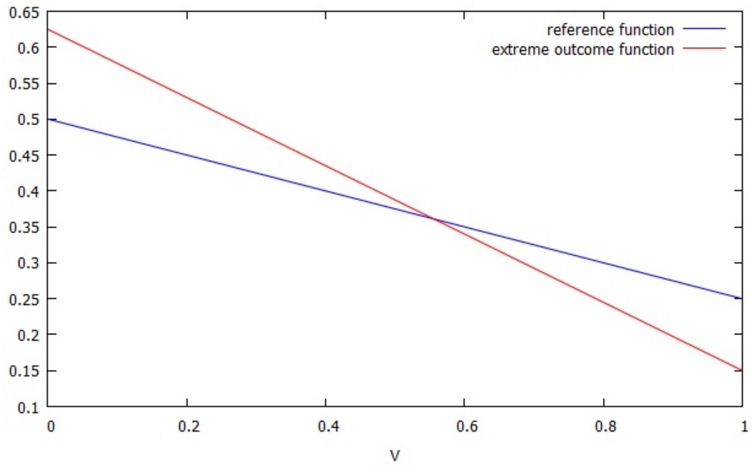


FIGURE 3.15: Example Showing That There Is No Clear Ranking Between π_α and π_q

$\pi_{\frac{1}{2}}(V) > \pi_q(V)$ for $V < V_0$ and $\pi_{\frac{1}{2}}(V) < \pi_q(V)$ for $V > V_0$.

Chapter 4

Value of Information

4.1 Introduction

For decades economists have been studying the relationship between decision-making under uncertainty and the so-called value of information. A famous and well-known result in this context is Blackwell's theorem (Blackwell 1953), stating that an experiment is more valuable than another if and only if the same experiment is more informative than the latter. In order to obtain this equivalence (e.g. Crémer 1982), a standard assumption has been that decision-makers are subjective expected utility (SEU) maximizers, cf. Savage [1954].

Over the last decades, SEU preferences have been subject to severe criticism, e.g. Ellsberg [1961] showed with the prominent Ellsberg paradox, that a decision-maker may display preferences which do not allow for subjective probabilities, thus showing an incompatibility with Savage's SEU theory. A well-established model of decision-making under uncertainty incorporating behavioral patterns that are in line with preferences displayed in the Ellsberg paradox, is given by the so-called maxmin expected utility (MEU) model. Maxmin preferences were axiomatized by Gilboa and Schmeidler [1989].

In recent years, multiple prior models have been used extensively in a wide range of economic fields like finance and behavioral economics (Gilboa et al. 2010, Riedel 2009). Given the importance of multiple priors and MEU in particular, it is worthwhile to know whether Blackwell's theorem extends to this class of preferences. This has been the objective of Çelen [2012], who offers a simple proof for the validity of Blackwell's theorem under MEU preferences.

In this comment, we demonstrate that Çelen's proof relies on a value of information for MEU preferences that is not defined via backward induction and thus is incompatible with the intertemporal extension of MEU introduced by Epstein and Schneider [2003]. In particular, optimal strategies in Çelen's framework prescribe decisions conditional on signal realizations that a MEU decision-maker will not find optimal to adhere to once those signal realizations have been observed. In this sense, Çelen's framework features dynamic inconsistency.

4.2 Framework and Definition of the Value of Information in Çelen's Model

In the following, we adopt Çelen's framework and notation. Let $\Omega := \{\omega_1, \ldots, \omega_n\}$ be the finite set of *states of the world* and $X := \{a_1, \ldots, a_\chi\}$ the finite set of *actions* available to a decision-maker. Moreover, we denote by $\Delta(\Omega)$ and $\Delta(X)$ the set of all probability distributions defined on Ω and X, respectively. Let further $u : \Omega \times X \to \mathbb{R}$ be a *utility function* and $\boldsymbol{u}$ with $u_{ij} = u(\omega_i, a_j)$ the corresponding utility matrix. An SEU decision-maker is characterized by $(\boldsymbol{\pi}, \boldsymbol{u})$, where $\boldsymbol{\pi} \in \Delta(\Omega)$ is a prior over the states.

An *experiment* is a tuple $(S, \boldsymbol{p})$ with the signal space $S = \{s_1, \ldots, s_\sigma\}$ and the Markov matrix $\boldsymbol{p}$ with $p_{ij} = \Pr(s_j|\omega_i)$ for $s_j \in S$. Çelen introduces a *strategy* as a vector valued mapping $f : S \to \Delta(X)$, thus characterizing all (mixed) actions the decision-maker plans to take after observing certain signal realizations s. The $\sigma \times \chi$-matrix $\boldsymbol{f}$ is defined such that $(f_{i1}, \cdots, f_{i\chi}) := f(s_i)$.

In this framework, Çelen determines the *value of the experiment* $(S, \boldsymbol{p})$ *for a given strategy* $\boldsymbol{f}$ as

$$\mathcal{U}^{\boldsymbol{f}}_{(\pi,u)}(S, \boldsymbol{p}) = \sum_j \Pr(s_j) \sum_i \Pr(\omega_i | s_j) \sum_k f_{jk} u(\omega_i, a_k) \tag{4.1}$$

$$= \sum_j \sum_i p_{ij} \pi_i \sum_k f_{jk} u_{ik} \qquad \text{(by Bayes' rule)} \; . \tag{4.2}$$

With a strategy $\boldsymbol{f}^*$ maximizing (4.2), Çelen defines $\mathcal{U}^*_{(\pi,u)}(S, \boldsymbol{p}) = \mathcal{U}^{\boldsymbol{f}^*}_{(\pi,u)}(S, \boldsymbol{p})$ as the value of the experiment for an SEU decision-maker.

Building on this, Çelen extends the definition of the value of an experiment to the class of MEU preferences. For that purpose, he characterizes an MEU decision-maker by (A, u), where $A \subset \Delta(\Omega)$ is a convex and compact *set* of priors. As a counterpart of $\mathcal{U}^*_{(\pi,u)}(S, \boldsymbol{p})$, he defines

$$\mathcal{W}^*_{(A,u)} = \max_{\boldsymbol{f}} \min_{\boldsymbol{\pi} \in A} \; \mathcal{U}^{\boldsymbol{f}}_{(\pi,u)}(S, \boldsymbol{p}) \tag{4.3}$$

as the value of an experiment $(S, \boldsymbol{p})$ for an MEU decision-maker. It is expression (4.3) that Çelen relies on in his proof of the generalized Blackwell theorem.

4.3 A Recursively Defined MEU Value of Information

It is insightful to note that Çelen's framework basically describes an intertemporal setting with two periods, a useful distinction that could be concealed by the fact that the decision-maker only acts once. In the second period, *after* observing a signal realization, the decision-maker takes a (mixed) action. In the first period, *before* observing the signal realization, the value of the experiment $(S, \boldsymbol{p})$ is determined. Çelen accounts for the intertemporal structure insofar as he considers strategies, that is complete contingent plans for appropriate play after observing signal realizations.

His formulation, however, is in contrast to the usual intertemporal formulation of MEU preferences that was provided by Epstein and Schneider [2003]. One of the main characteristics of the recursive definition of intertemporal MEU in Epstein and Schneider [2003] is the compatibility with backward induction. We follow their approach and present here an alternative definition of the value of information for MEU preferences. According to backward induction, the first step to define a value of information is to determine an optimal action *after* observing signal realization s_j, $j = 1, \ldots, \sigma$, which is given by

$$g_j^* \in \underset{g \in \Delta(X)}{\operatorname{argmax}} \min_{\boldsymbol{\mu} \in \mathcal{M}(s_j)} \mathbb{E}_{\boldsymbol{\mu}}[u] \ . \tag{4.4}$$

Here, $\mathcal{M}(s_j)$ is the set of *posteriors* after observing signal realization s_j, formally

$$\mathcal{M}(s_j) := \{p(\cdot|s_j) \, : \, p \in A\} \ , \tag{4.5}$$

where $p(\cdot|s_j)$ denotes the conditional probability of the prior $p \in \Delta(\Omega)$ given the signal s_j. We obtain $p(\cdot|s_j)$ via Bayes' rule and update each prior p in this way.[1]

According to the principle of backward induction, we determine the value of information in the first period on the assumption of optimal actions in the second period. Thus, we suggest the following definition of the value of an experiment $(S, \boldsymbol{p})$ for MEU preferences:

$$\tilde{\mathcal{W}}^*_{(A,u)} = \min_{\boldsymbol{\pi} \in A} \sum_{i,j} \pi_i p_{ij} \sum_k g^*_{jk} u_{ik} \ . \tag{4.6}$$

Here, g_j^* denotes an optimal decision after observing signal realization s_j, given in (4.4). This alternative way of defining the value of an experiment is in line with the intertemporal

[1] Epstein and Schneider [2007] show that further restrictions on the set $\mathcal{M}$ can be made. For the sake of simplicity, you may think of full Bayesian updating.

model of recursive utility under multiple priors as pointed out in Epstein and Schneider [2003, 2007].

The key characteristic of (4.6) is that optimal actions are determined with the MEU rule for each signal realization s_j individually. In particular, the worst posterior in (4.4) in general depends on the signal realization s_j. This is in contrast to (4.3). By following the derivation of the SEU counterpart, essentially the step from (4.1) to (4.2), Çelen silently assumes that the worst prior from the ex-ante perspective coincides with the preimage of *all* worst posteriors, irrespective of the signal realization. For the SEU decision-maker this argumentation is innocent as there is a unique prior, and thus a unique posterior as well. For the MEU decision-maker, however, this argument is in conflict with backward induction.

In the appendix, we demonstrate that the conflict of Çelen's framework with intertemporal recursive utility can be made even more concrete. We provide an example in which the optimal strategy derived in Çelen's framework prescribes actions that are different from what a MEU decision-maker will actually do after observing those signals realizations.[2] This supports our claim that the value of information for MEU preferences should be defined by (4.6). By construction, our definition of the value of information is compatible with dynamic consistency.

[2] One could think that the reason we observe this form of dynamic inconsistency is the missing assumption of *rectangularity* of the prior set, a key assumption in Epstein and Schneider [2003] to ensure dynamic consistency within an intertemporal setting of recursive utility. But this is not the case. Even though Çelen's setting is not fully transferable to the setting of Epstein and Schneider, in particular the analysis in Epstein and Schneider [2007] suggests that rectangularity is no issue in this setting, simply because the learning process is defined via conditional one-step-ahead conditionals, as required by Epstein and Schneider [2003]. The reason for the peculiar properties of Çelen's framework lies in the fact, that utilities are defined in a non-recursive way. His framework is thus incompatible with Epstein and Schneider right from the start.

4.4 Results and Discussion

We have shown that Çelen's proof of the Blackwell theorem only applies to a value of information that is defined in a non-recursive utility framework. We have offered a definition for the value of information derived via backward induction, thus compatible with the dynamic consistent intertemporal axiomatization of Epstein and Schneider [2003]. Consequently, we suggest that the proof of the Blackwell theorem should deal with expression (4.6) as the definition of the value of information for MEU preferences. This proof is still pending.

4.5 Example Demonstrating That Çelen's Value of Information Is No Dynamically Consistent

After observing a certain signal realization, an MEU decision-maker will in general *not* adhere to actions she determined to be optimal before the signal realization has been observed. In other words, an optimal strategy f^* determined in Çelen's framework in general prescribes, for all signal realization contingencies, actions that are different from what an MEU decision-maker will actually do after observing those signal realizations.

We demonstrate this with a simple example. We restrict the number of states of the world $\Omega = \{\omega_1, \omega_2\}$, actions, $X = \{a_1, a_2\}$ and signal realizations $S = \{s_1, s_2\}$ to two. By that, we can write a prior $\boldsymbol{\pi}$ as $(\pi_1, 1-\pi_1)$. Moreover, the Markov matrix $\boldsymbol{p}$ is fully specified by $p_{11} = p_{22} = \lambda$ and $p_{12} = p_{21} = 1 - \lambda$ with $1/2 < \lambda < 1$. We assume $1/2 < \lambda < 3/4$. Due to the restriction on two signal realizations, we can write $f_{12} = 1 - f_{11}$ and $f_{21} = 1 - f_{22}$. To further simplify our example, we specify payoffs by $u_{11} = 1$, $u_{12} = -1$, $u_{22} = 2$ and $u_{21} = 0$. This is a simple example of a setting in which the decision-maker wants to learn the true ω because action a_1 is optimal if $\omega = \omega_1$ and action a_2 is optimal if $\omega = \omega_2$.

With these specifications, (4.2) reduces to

$$\mathcal{U}^{\boldsymbol{f}}_{(\pi,u)}(S,\boldsymbol{p}) = (2f_{11} - 2f_{22} - 1)\,\pi_1 \;+\; 2 - 2\,(1-f_{22})\,\lambda - 2f_{11}(1-\lambda)\ . \tag{4.7}$$

For the set of priors, we specify $A = \{(\pi_1, 1-\pi_1)\ :\ 1/4 \leq \pi_1 \leq 3/4\}$.

In order to determine Çelen's optimal strategy $\boldsymbol{f}^*$ in equation (4.3), we first calculate, for a given strategy $\boldsymbol{f}$, the prior that minimizes (4.7). This is given by

$$\boldsymbol{\pi}^{\text{worst}} = \begin{cases} \left(\frac{1}{4},\frac{3}{4}\right) & \text{if } f_{11} - f_{22} > \frac{1}{2} \\ \text{any } \boldsymbol{\pi} \in A & \text{if } f_{11} - f_{22} = \frac{1}{2} \\ \left(\frac{3}{4},\frac{1}{4}\right) & \text{if } f_{11} - f_{22} < \frac{1}{2}\ . \end{cases} \tag{4.8}$$

Building on this, we can derive the optimal strategy $\boldsymbol{f}^*$. We calculate

$$\boldsymbol{f}^*(s_1) = (1,0) \quad , \quad \boldsymbol{f}^*(s_2) = \left(\tfrac{1}{2},\tfrac{1}{2}\right)\ . \tag{4.9}$$

In words, the optimal strategy in the Çelen framework consists of taking action a_1 if $s = s_1$ and mixing over actions a_1 and a_2 with equal weights if $s = s_2$.

We now demonstrate that an MEU decision-maker that determines her optimal strategy via (4.3) would actually revise her optimal plan as soon as the signal materializes. The decision rule after observing a signal realization s_j is given by (4.4), where g is a randomization over actions a_1 and a_2, and $\mathcal{M} \subset \Delta(\Omega)$ is the set of posteriors that depends on the set of priors A, the likelihood p and the signal realization s_j observed. In our example, the expected value of the decision g under the posterior $\boldsymbol{\mu}$ is

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\mu}}[u] &= \mu_1\,(g_1 \cdot 1 + (1-g_1)\cdot(-1)) + (1-\mu_1)\,(g_1 \cdot 0 + (1-g_1)\cdot 2) \\ &= (4g_1 - 3)\,\mu_1 + 2(1-g_1)\ .\end{aligned}$$

Using the notation $\underline{\mu}_1 := \min_{\boldsymbol{\mu}\in\mathcal{M}} \mu_1$ and $\bar{\mu}_1 := \max_{\boldsymbol{\mu}\in\mathcal{M}} \mu_1$, the worst posterior is

$$\boldsymbol{\mu}^{\text{worst}} = \begin{cases} \left(\underline{\mu}_1, 1-\underline{\mu}_1\right) & \text{if } g_1 > \frac{3}{4} \\ \text{any } \boldsymbol{\mu} \in \mathcal{M} & \text{if } g_1 = \frac{3}{4} \\ (\bar{\mu}_1, 1-\bar{\mu}_1) & \text{if } g_1 < \frac{3}{4}\ . \end{cases} \tag{4.10}$$

We rewrite $\mathbb{E}_{\boldsymbol{\mu}}[u] = g_1\,(4\mu_1 - 2) + 2 - 3\mu_1$. From that we can infer that g_1 is chosen minimal if the relevant posterior fulfills $\mu_1 < 1/2$ and g_1 is chosen maximal if the relevant posterior fulfills $\mu_1 > 1/2$. Our assumption $\lambda < 3/4$ implies for all signal realizations $\underline{\mu}_1 < 1/2$ and $\bar{\mu}_1 > 1/2$. Thus for $g_1 \geq 3/4$, it is optimal to lower g_1 as much as possible. Considering the case $g_1 \leq 3/4$, it is optimal to increase g_1 as much as possible. Taken together, this shows $g^* = (3/4, 1/4)$, independent of the signal realization. In words, the optimal action of the MEU decision-maker (both after receiving $s = s_1$ and $s = s_2$) is to mix over actions a_1 and a_2 with the ratio 3 to 1.

As we have shown above, this is different from the behavior prescribed in Çelen's framework, given in expression (4.9). This shows the dynamic inconsistency and thus illustrates the incompatibility of Çelen's framework with the recursive setting of Epstein and Schneider [2003, 2007].

Chapter 5

Implementing Ambiguity - the Monopoly Market

5.1 Models with Perfect and Imperfect Information

Every economic model is an attempt to provide a simplified description of the world with the inherent objective to obtain a better understanding of real-world phenomena as well as complex economic interrelations and interactions. Since every model is just a simplification of the real world, it neglects a certain number of influencing factors or determinants and is therefore strictly speaking "always wrong", compare Box and Draper [1987], page 424. As such, constructing a model goes hand in hand with a process of formal abstraction in which the modeler is required to identify and select a certain number of influencing factors, or variables, that he or she deems relevant for the underlying analysis. The primitives of each model specify these variables and establish logical relationships between them. Deardorff [2001] gives the following definition of an economic model: An economic model is "a collection of assumptions, often expressed as equations relating variables, from which inferences can be derived about economic behavior and performance."

Let me hereafter introduce a simple and well-known example for a stylized "economic model" that is going to serve as a template for the considerations of this chapter.

Example 5.1 (Monopoly with Linear Demand under Certainty)**.** Consider a monopoly market and a firm that produces a homogeneous good. Market demand is linear and given by

$$D(p) = \max\{0, a - bp, 0\}$$

where $a, b > 0$ are parameters and p denotes the price of the good. The parameter a can be interpreted as the maximum number of possible customers whereas the slope parameter b captures how fast demand decreases when prices increase. Moreover, the firm faces a marginal cost of $c > 0$ for producing one unit of the good and no fixed costs. The firm is assumed to set a price such that its profit

$$\Pi(p) = (p - c)D(p)$$

is maximized. The profit function is continuous, since sums and products of continuous functions are continuous. Moreover, the objective is piecewise continuously differentiable with a kink at $p_0 = \frac{a}{b}$. The following proposition describes the solution of the monopolist's

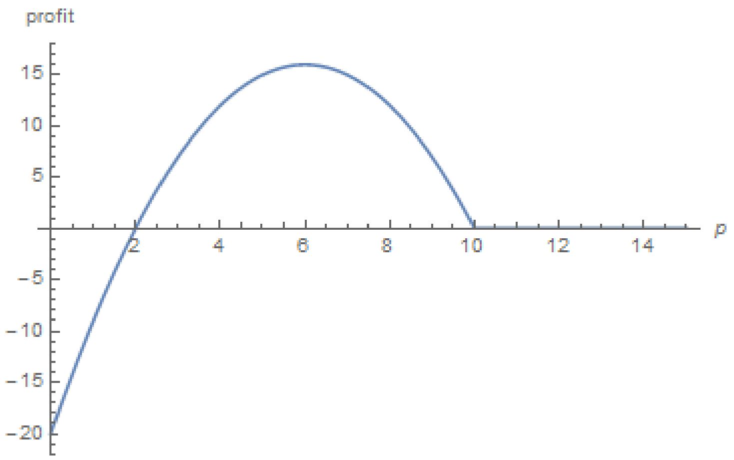


FIGURE 5.1: Objective in the Certainty Case

optimization problem in the certainty case.

Corollary 5.1. The solution of the monopolist's optimization problem under certainty is given by

$$p^{certainty} = \begin{cases} p^M = \frac{a+bc}{2b} & a > bc \\ [p_0, \infty) & a \leq bc \end{cases}$$

Proof. The proof is contained in the last section of this chapter.

The result of Proposition 5.1 can be interpreted in the following way: As long as the marginal cost parameter is small enough, the monopolist charges the monopoly price p^M. As marginal costs increase, the monopoly price p^M increases as well. As soon as the monopoly price exceeds the threshold value $p_0 = \frac{a}{b}$, demand is zero. As a consequence, the best the monopolist can do in such cases is to secure a profit of zero by setting a price $p \geq p_0$.

Implicitly, this model features assumptions on the firm's state of knowledge with respect to the underlying variables and relationships. To be more precise, the calculation of the monopoly price presumes that the firm manager knows the relationship between prices and market demand, as well as the underlying cost structure of the firm. The starting point for the introduction of ambiguity into preexisting economic models consists in identifying each agent's degree of information with respect to the underlying parameters. Doing so, one can roughly distinguish between two different cases which are treated in the following. The case of *perfect information*, where the decision-maker knows a certain parameter or functional relationship, and the case of *imperfect information* where this knowledge is at least partially absent. The simple monopoly model is an example for perfect information, since all relevant parameters are assumed to be known.
Whether it is reasonable to drop the assumption of perfect information for a specific variable or not strongly depends on the modeling context and the narrative of the underlying problem. Assume, for instance, that the monopolist can rely on a number of high quality market studies, and that all of these studies indicate that market demand in the

current state of the economy is given by $D(p)$. Moreover, assume that the monopolist has been operating successfully in the market for many years and that the results of the market research studies correspond to the monopolist's experience from past years. In such a case, it seems inappropriate to assume that, given the overwhelming evidence, the monopolist would perceive market demand as uncertain in the current state of the economy. On the other hand, if we consider situations where performing market studies is too costly or time-consuming, or where the monopolist issues a completely new product into the market, one can more easily support the hypothesis that the assumption of perfect information is doubtful.

Imperfect information prescribes the presence of uncertainty. According to Knight [1921], uncertainty can be classified into two categories: risk and ambiguity. In cases where uncertainties are captured by risk, decision-makers ignore a crucial parameter or relationship, but they know the set of possible outcomes for this parameter and the probability for each of these outcomes. The following section extends Example 5.1 to a scenario where the monopolist ignores the overall number of consumers a.

Example 5.2 (Monopoly with Linear Demand under Risk)**.** Suppose the monopolist knows that there are two scenarios: In the first scenario, the overall number of consumers is high and denoted by $a^H > 0$; in the second scenario, there is a smaller, but positive, number of consumers $0 < a_L < a_H$.[1] Besides, the monopolist knows the probability of each scenario. Let $s \in (0, 1)$ denote the probability that a_H is the true underlying parameter.[2] Then, market demand is given by $D_1(p) = \max\{0, a_H - bp, 0\}$ with probability s and by $D_2(p) = \max\{0, a_L - bp, 0\}$ with probability $1 - s$. Under the assumption that the monopolist maximizes his expected profit, we obtain the following objective:

$$E_s[\Pi](p) := (p - c)\left(sD_1(p) + (1 - s)D_2(p)\right)$$

[1]If we had $a_H = a_L$, we would be back in the certainty case.

[2]At this point, I exclude the extreme cases $s = 0$ and $s = 1$, since the solution of these cases corresponds to the certainty case treated in Example 5.1. This can be achieved by replacing the demand function $D(p)$ with $D_1(p)$ or $D_2(p)$ respectively.

Throughout the rest of the paper, I condense the notation slightly writing $E_s[\Pi]$ instead of $E_s[\Pi](p)$ if not otherwise specified. Knowing that sums and products of continuous functions are continuous, we can infer that the monopolist's profit function is also continuous. Furthermore, the objective is only piecewise continuously differentiable. This is because it has two kinks at $p_0 = \frac{a_L}{b}$ and $p_1 = \frac{a_H}{b}$, see Figure 5.2 for an illustration. The

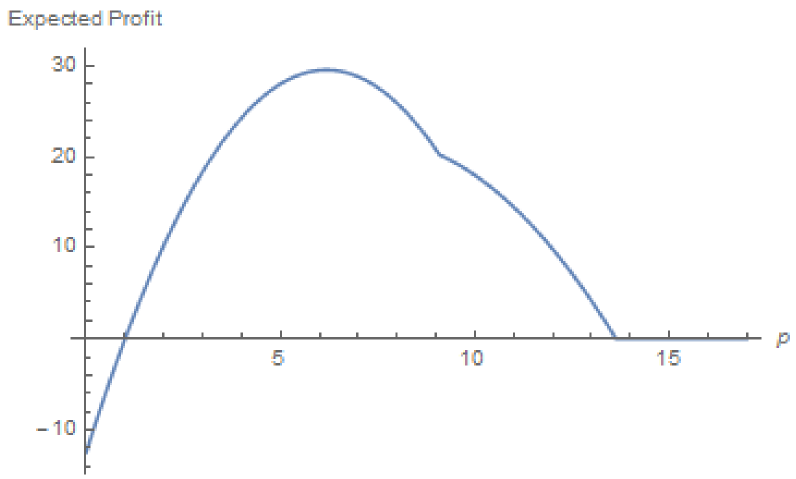


FIGURE 5.2: Objective Function Under Risk

following proposition characterizes the solution of the monopoly model under risk.

Proposition 5.1. The following prices are possible solutions of the monopolist's optimization problem under risk:

1. $p^* = \frac{sa_H+(1-s)a_L+bc}{2b}$
2. $p^{**} = \frac{a_H+bc}{2b}$
3. $p^{***} \in [p_1, \infty]$.

Which of these candidates is the global maximizer depends on the underlying parameter constellations and the threshold value

$$\hat{s} = \frac{(b\,c - a_L)^2}{(a_L - a_H)^2}.$$

The price p^* is the only solution of the optimization problem if one of the following sets of conditions is met:

1. C_1: $bc < 2a_L - a_H$

2. C_2: $a_H \geq \max\{bc, 2a_L - bc\}$, $2a_L - sa_H - (1-s)a_L \geq bc$, and $s \in [0, \hat{s}) \cup \{1\}$

3. C_3: $a_H \geq \max\{bc, 2a_L - bc\}$, $2a_L - sa_H - (1-s)a_L \geq bc$, and $\hat{s} > 1$

The price p^{**} is the only solution of the optimization problem if one of the following sets of requirements holds:

4. C_4: $a_H \geq \max\{bc, 2a_L - bc\}$, $2a_L - sa_H - (1-s)a_L \geq bc$, and $s \in (\hat{s}, 1)$

5. C_5: $a_H \geq \max\{bc, 2a_L - bc\}$ and $2a_L - sa_H - (1-s)a_L \leq bc$

Both p^* and p^{**} are the solution of the optimization problem if the following assumptions are met:

6. C_6: $a_H \geq \max\{bc, 2a_L - bc\}$, $2a_L - sa_H - (1-s)a_L \geq bc$, and $s \in \{\hat{s}, 1\}$

Every $p^{***} \in [p_1, \infty]$ is a solution of the optimization problem if the following condition is satisfied:

7. C_7: $a_H < bc$

Proof. The proof is contained in the last section of this chapter.

The result of Proposition 5.1 can be summarized in the subsequent manner: The optimal monopoly price crucially depends on marginal costs. There are three cases. In the first case, marginal costs are so small such that the demand remains positive in both potential scenarios a_H and a_L. The second case deals with intermediate values of c; it can either occur that the monopoly price yields positive demand for both scenarios, or that the demand in the low number of consumers scenario a_L is zero but remains positive in the high number of consumers scenario a_H. Whether p^* or p^{**} is the solution depends on the probability s and the threshold value $\hat{s}$. If s equals one, the model reduces to a model of certainty with $a = a_H$. In this case $p^* = p^{**}$ is the optimum. For small values of s, the monopolist expects the low consumer scenario a_L to materialize. Consequently, the monopolist's objective gives a small weight to scenario a_H and a large weight to scenario a_L. Under these assumptions, the price p^* is optimal. The threshold value $\hat{s}$ has no bite when both scenarios a_L and a_H are sufficiently similar. In cases where difference between a_H and a_L is large enough[3], p^{**} is optimal if the monopolist's belief s gives sufficiently weight to scenario a_H. Finally, if marginal costs are so large that the corner solution price p^{**} exceeds the value $p_1 = \frac{a_H}{b}$, we can infer that $D_1(p^{**}) = D_2(p^{**}) = 0$. Hence, the best the monopolist can do is to secure a profit of zero. This can be achieved by any price $p^{***} \in [p_1, \infty)$.

Note that the higher the probability s of the high demand scenario, the higher the resulting monopoly price p^*. In cases where one of the corner solutions p^{**} or p^{***} is optimal, the monopoly price is independent of s.

5.2 Why Ambiguity?

Whether the monopolist's imperfect information with respect to an underlying variable should be modeled by a decision-theoretic framework featuring risk, or by a framework featuring ambiguity, is a challenging question that cannot be answered completely satisfactorily. Under ambiguity, the decision-maker knows the set of possible outcomes but is confronted with a lack of (reliable) probabilistic information. In fact, ambiguity presumes that, due to the absence of crucial information, the monopolist is incapable of assigning a well-defined probability to the events $\{a_H\}$ and $\{a_L\}$. Therefore, the real underlying issue is to assess whether we can justify the existence of a well-defined probability distribution that the monopolist may use as a basis for his decision-making process.

The existence problem is at the same time mathematical and philosophical in nature and

[3] The threshold value $\hat{s}$ converges to zero if the difference $\Delta_a = |a_H - A_L|$ converges to infinity.

broadly discussed in the literature. From a mathematical point of view, a probability is defined via Kolmogorov's axioms, see Kolmogorov [1950] for more details. Kolmogorov's axioms result in a sound formal theory but leave room for interpretation when it comes to answering the question where probabilities come from and how to interpret them. This is exactly the point where the philosophical discussion on probabilities sets in. The problem of appropriately defining probabilities has been addressed by different schools of thought. A discussion of these can be found in many textbooks on decision-theory and philosophy. In the following, I summarize and discuss briefly the main interpretations of probabilities introduced in Peterson [2009] starting with Laplace's classical definition of probabilities, see Laplace and Truscott [1814].

According to Laplace's notion of probabilities, a probability is defined as a ratio, namely the number of favorable cases divided by the total number of cases where each case is presumed to be "equally possible". As an example of the classical definition of a probability, consider, for instance, an urn containing r_1 red balls and r_2 blue balls. Then, the probability of drawing a red ball is the number of favorable cases r_1 divided by the total number of cases $r_1 + r_2$. The classical definition of probabilities has its limitations. It is, for instance, not clear whether events can always be divided in such a way that they are equally possible. Moreover, the procedure of dividing the state space into equally possible events needs to be done by using symmetry arguments, such as the principle of insufficient reason, in order to avoid the problem of circular logical arguments,[4] compare Hájek [2012]. Secondly, the classical definition only applies to finite state spaces.

The second notion for probabilities is the so-called frequentist definition. Frequentism presumes that probabilities directly arise from empirical observations; therefore, a probability is defined as relative frequency. Formally, the probability of an event A is defined as the number of observations in which A occurred divided by the total number of observations, see for instance Peterson [2009], pp. 136-137. The problem of the frequency definition is that different series of observations give rise to different probabilities. An

[4]Otherwise, one would implicitly assume that "equally possible" has the same meaning as "equally probable".

extension of the frequentist view is proposed by Venn [1888], who attempts to resolve the problems of the relative frequency definition by assuming that the probability of an event equals the limit of the relative frequency if the underlying experiment was repeated infinitely many times. The problem of this alternative definition is that it is not clear whether it is always possible to repeat an experiment infinitely many times, or whether the relative frequency really converges, since we cannot observe the whole sequence of trials.

The third interpretation of probabilities is the propensity definition pioneered by Karl R. Popper.[5] Peterson [2009], pp.139-140, provides the subsequent definition of a probability according to the propensity approach:

> *"[...] probabilities can be identified with certain features of the external world, namely the disposition or tendency of an object to give rise to a certain effect."*

The propensity interpretation is criticized for a variety of reasons. The first one is that the term "propensity" is dubious, since it cannot be defined in a clear manner. The second objection frequently put forward against propensities is that they entail a temporal structure which precludes Bayes' theorem.[6] See for instance Humphreys [1985] for a critique on the propensity approach.

The fourth interpretation of probabilities is the so-called logical or epistemic notion of probabilities. According to Hájek [2012], the logical interpretation dates back to Keynes [1921], Johnson [1921], Jeffreys [1939], and Carnap [1962]. The logical approach assigns probabilities to hypotheses which are, by definition, unverified logical statements. More precisely, the logical interpretation presumes that probabilities can be logically deduced from evidence E supporting a certain hypothesis H, compare Peterson [2009], page 141. Mathematically speaking, the logical probability $p(H|E)$ assigns a probability to the event that the hypothesis H is true. Assume, for instance, that we intend to assign a probability to the hypothesis that the so-called "giant impact hypothesis" regarding the moon's

[5]Compare Popper [1957] and Popper [1959].
[6]Compare Peterson [2009], pp. 140-141.

origin is true. In this case, the underlying logical probability describes how strongly the giant impact hypothesis is supported, given current scientific evidence at hand. One focus of the criticism put forward against logical probabilities is the way evidence is connected to probabilities in Carnap's approach[7] and the fact that probabilities are solely evidence-dependent excluding the possibility of guesses.[8]

At this point, I want to mention that there are different perceptions in the literature on how to define objective probabilities. Anscombe and Aumann [1963] tie the definition of objective probabilities to interpersonal agreement. When two persons' subjective beliefs coincide, then they are called objective in the sense of Anscombe-Aumann. This approach is criticized in Gilboa [2009], pp. 138-139. Loosely speaking, the critique is that according to Anscombe-Aumann's definition objectiveness might just be the result of a coincidental match of beliefs. But why should we think that coincidence can be a foundation of an objective probability? In this thesis, I adopt a more restrictive view on what should be called an objective probability. Therefore, I adopt the definition proposed in Peterson [2009], page 133, that objective probabilities are derived from facts in the external world and not from personal judgment. With this definition in mind, I subsume the classical, frequentist, logical, and propensity approaches to probability under the term "objective probabilities".

In contrast to objective probabilities, subjective probabilities arise from the agents themselves, the "subjects", and can be interpreted as individual assessments of risky situations. Therefore, subjective probabilities are frequently denoted as "beliefs". The subjective definition does not require the agents to derive probabilities from facts or knowledge, the "objects", related to the problem. Major contributions for the subjective interpretation originate from Ramsey [1931], De Finetti [1937], and Savage [1954]. Briefly summarized, the subjective approach allows us the elicitation of beliefs by offering hypothetical bets. Note that according to Ramsey [1931], agents are not required to "know" their subjective probability. It is rather that, if agents conform to Savage's axioms, one can infer their

[7]There is the need to introduce some weighting between different pieces of evidence, compare Hájek [2012], page 11, for more details.

[8]See Peterson [2009], page 141.

subjective beliefs by observing their betting behavior. The subjective interpretation of probabilities is criticized for the fact that beliefs are attached to preferences. In a sense, it is not excluded that a preference for certain states of the world influences an agent's personal probabilistic judgments, compare Hájek [2012] for more details.

Following the discussion of the different schools of thought on probability, I return to the initial question of this section, namely which prerequisites need to be satisfied such that the implementation of ambiguity into preexisting economic models is justified. In order to answer this question, we should keep in mind that the concept of ambiguity obviously originates from the subjective school of thought on probabilities. Both the subjective school and the advocates of ambiguity models share the view that beliefs can be purely subjective in nature. But wherein lies the difference? Remember, the development of non-expected utility models originates from Ellsberg's discovery that agents' behavior might contradict the existence of a single subjective belief.

Taking these considerations together, one realizes that modelers face two major obstacles when implementing ambiguity preferences.

The first obstacle is to provide a sound justification of why objective probabilities seem inappropriate in the context of the model setup. Given the strict definition of objective probabilities in Peterson [2009], one can see that the derivation of such probabilities presumes the availability of facts or evidence from the external world. Conversely, this means that the justification of objective probabilities is not guaranteed whenever agents are insufficiently informed about these facts and therefore prone to rely on individual probabilistic judgments.

Frequently, the rationale in this step boils down to one, or several, of the following reasons: First of all, agents might lack crucial data; data might be absent or inaccessible and its procurement time-consuming or associated with high costs. Secondly, even if data was available, it could originate from an unreliable source of information and therefore be judged as insufficient or imprecise. Finally, there is the possibility that agents dispose of contradictory information, or information highlighting different aspects of the underlying uncertainty.

The second obstacle to ambiguity is to argue why agents violate SEU. Suppose that a set of conditions was identified such that the hypothesis of subjective probabilities can be supported. If the agents conform to the axioms of Savage [1954], one can "replace" objective probabilities with subjective ones and conduct an analysis based on the expected utility calculus. Since SEU is widely considered as the benchmark model of decision making under uncertainty, it is indispensable to point out why the agents might not conform to this model. The main argument frequently put forward to support this hypothesis is the Ellsberg paradox by Ellsberg [1961]. It demonstrates that, in situations where probabilities are not objectively given, decision-makers might display preferences that refute the hypothesis of subjective probabilities. The validity of the Ellsberg paradox has been confirmed by a variety of experimental studies such as Camerer and Weber [1992] or Halevy [2007].

On the whole, modelers are confronted with two major challenges when arguing in favor of ambiguity models. The first one is to support the hypothesis that probabilities are unlikely to be objective, and the second one is to refute the SEU model.

At the end of this section, I want to point out that whenever one decides to generalize a baseline model by making use of non-expected utility models, you find yourself implicitly in the tradition of the Non-Bayesian school of thought on probabilities and decision making. The decision whether or not to include ambiguity into economic models is not necessarily a question of "right" or "wrong" that can be easily determined along a series of clear-cut objective criteria. It is rather a commitment to the Non-Bayesian approach which needs defense on the grounds of plausible reasoning.

5.3 Implications of Ambiguity for Monopoly Pricing

The previous section of this chapter addresses the question whether it is appropriate to introduce ambiguity into a preexisting baseline model or not. In the following, I discuss the implications of the most prominent models of decision making under uncertainty for the baseline monopoly example. Note that the problem of monopoly pricing under ambiguity is addressed by a novel strain of literature in Industrial Organization. Asano and Shibata [2011] consider a continuous-time dynamic pricing model where a monopolist can make irreversible quality investments. The authors find that, contrary to risk, the presence of Knightian uncertainty decreases prices and optimal quality. Bergemann and Schlag [2011] investigate monopoly pricing under MEU and the minimax regret criterion. The authors find that both decision criteria lead to lower monopoly prices than under a certainty. Using a maxmin rule, a recent paper by Zheng et al. [2015] investigates optimal non-linear monopoly pricing under ambiguity. There is a continuum of different types of buyers. Each type is characterized by a certain valuation for the product. Knightian uncertainty is modeled by so-called ε-contaminations, see Eichberger and Kelsey [1999] for more details. The authors find that, under Knightian uncertainty, the monopolist assumes a larger portion of buyers to have the lowest possible valuation of the product. Besides, the monopolist adjusts her pricing strategy under ambiguity by offering a larger discount to all consumers.

The objective of this section is to investigate the pricing problem in a simple static framework for the most prominent models of decision making under uncertainty. In this context, I assume that the monopolist is risk-neutral and faces demand ambiguity with respect to the intercept a. Moreover, I abstract from quality investments. Contrary to Asano and Shibata [2011] and Bergemann and Schlag [2011], I can demonstrate that Knightian uncertainty can increase or decrease optimal prices. Besides, I can show that, due to the possibility of corner solutions, there are instances where ambiguity has no effect on optimal prices.

An important point to be considered is whether the monopolist's attitude towards ambiguity should be included into the model framework or not. Naturally, this question is closely related to the research agenda intended by the introduction of ambiguity. If, for instance, the aim of research solely consists in comparing a risky environment with an ambiguous situation characterized by extreme pessimism, then the MMEU model developed by Gilboa and Schmeidler [1989] might be the right choice, since it features a simple and parsimonious formal structure, which is sufficient to address the underlying question. In terms of the monopoly example the underlying question would be how the monopoly price is affected if the monopolist features extreme pessimism with respect to the maximum number of consumers.

Example 5.3 (Monopoly with Linear Demand and MMEU-Preferences)**.** MMEU-preferences presume that the monopolist's beliefs are represented by a nonempty, closed, and convex set of priors $\mathcal{P}$, see Gilboa and Schmeidler [1989], page 145. Since I consider only a model with two outcomes, we can associate every prior $(s, 1-s)$ with the number $s \in [0,1]$. By using this simplified notation, we can represent $\mathcal{P}$ by means of a compact interval $\mathcal{P} = [\underline{s}, \overline{s}]$ where $0 \leq \underline{s} < \overline{s} \leq 1$. Under MMEU, the monopolist's optimization problem is given by

$$\max_{p \geq 0} \min_{s \in [\underline{s}, \overline{s}]} E_s[\Pi] \tag{5.1}$$

where

$$E_s[\Pi] = (p-c)(sD_1(p) + (1-s)D_2(p))$$

denotes the expected profit given the belief $(s, 1-s)$. As in the risk case, the objective function has two kinks at $p_0 = \frac{a_L}{b}$ and $p_1 = \frac{a_H}{b}$. The following corollary states the monopolist's worst-case priors.

Corollary 5.2. The prior s_{worst} inducing worst-case expected profits is given by

$$s_{worst} := \begin{cases} \overline{s} & \text{for} \quad p < c \\ [\underline{s}, \overline{s}] & \text{for} \quad p = c \\ \underline{s} & \text{for} \quad c < p < \frac{a_H}{b} \\ [\underline{s}, \overline{s}] & \text{for} \quad p \geq \frac{a_H}{b} \end{cases}$$

Proof. The proof is contained in the last section of this chapter.

By using Corollary 5.2, we obtain the following representation for the reduced objective:

$$E_{s_{worst}}[\Pi] =: \Psi(p) = \begin{cases} E_{\overline{s}}[\Pi] & \text{for} \quad 0 \leq p < c \\ E_{\underline{s}}[\Pi] & \text{for} \quad p \geq c \end{cases}$$

Note that

$$E_{\underline{s}}[\Pi] = E_{\overline{s}}[\Pi] = E_{s}[\Pi] = 0$$

for $p = c$ and $p \geq \frac{a_H}{b}$. Therefore, we can replace $E_{s_{worst}}[\Pi]$ by $E_{\underline{s}}[\Pi]$ in cases where p equals the marginal cost parameter c or where p exceeds the threshold value $\frac{a_H}{b}$. The monopolist's problem can be expressed as

$$\max_{p \geq 0} \Psi(p). \tag{5.2}$$

The subsequent corollary characterizes the solution of problem (5.1).

Corollary 5.3. The following prices are possible solutions of the monopolist's optimization problem under MMEU-preferences:

1. $p^{*}_{pess} = \frac{\underline{s}a_H + (1-\underline{s})a_L + bc}{2b}$
2. $p^{**}_{pess} = \frac{a_H + bc}{2b}$
3. $p^{***}_{pess} \in [p_1, \infty]$

Which of these candidates is the global maximizer depends on the Conditions C_1 to C_7 and the threshold value $\hat{s}$ defined in Proposition 5.1. Note that it is necessary to replace the probability s with the worst-case prior $\underline{s}$ in Conditions C_2 to C_6. The price p^{*}_{pess} is the only solution of the optimization problem if one of the Conditions C_1, C_2, or C_3 is satisfied. The price p^{**}_{pess} is the only solution of the optimization problem if the Condition C_4 or the Condition C_5 holds. Both p^{*}_{pess} and p^{**}_{pess} are global maximizers if Requirement C_6 is met. Every $p^{***}_{pess} \in [p_1, \infty]$ is a solution of the optimization problem if C_7 holds.

Proof. The proof is contained in the last section of this chapter.

Note that the risky monopoly price p^*, which is defined in Example 5.2, is at least as large as the pessimistic monopoly price p^*_{pess} in cases where the belief s is not smaller than the worst-case belief $\underline{s}$. This condition is automatically fulfilled when the monopolist relies on a full prior set $\mathcal{P} = [0, 1]$. In cases where $\mathcal{P}$ is a strict subset of $[0, 1]$, we can conclude that p^* is smaller than p^*_{pess} if $s < \underline{s}$. Such a scenario is not very plausible, since it presumes that the monopolist exhibits stronger pessimistic under risk than in the worst-case scenario under ambiguity. The concept of a prior set, however, implicitly assumes that $\mathcal{P}$ contains all conceivable priors from an ex-ante perspective. For this reason, one could argue that, in such cases, the prior set $\mathcal{P}$ lacks a crucial belief.

What we learned so far from the MMEU-approach is that extreme pessimism with respect to the maximum number of consumers leads to a lower monopoly price than in a scenario where the monopolist holds a subjective belief s with $s > \underline{s}$. These results are consistent with the findings of Asano and Shibata [2011] and Bergemann and Schlag [2011]. A legitimate question, which cannot be answered by the MMEU-model, is, what happens if the monopolist displays optimism instead of pessimism. It remains unclear how ambiguity affects monopoly pricing for intermediate cases of optimism and pessimism. Is it that an increase in optimism always yields a higher monopoly price? A model accommodating different attitudes towards ambiguity is the α-MEU model by Ghirardato et al. [2004].

Example 5.4 (Monopoly under α-MEU)**.** A monopolist making use of the α-MEU heuristics maximizes the objective

$$\max_{p \geq 0} \left\{ \alpha \min_{s \in [\underline{s}, \overline{s}]} E_s[\Pi] + (1 - \alpha) \max_{s \in [\underline{s}, \overline{s}]} E_s[\Pi] \right\} \tag{5.3}$$

where

$$E_s[\Pi] = (p - c)(sD_1(p) + (1 - s)D_2(p))$$

denotes the expected profit given the belief $(s, 1-s)$. Following a reasoning similar to the one developed in Example 5.3, I identify the monopolist's worst- and best-case priors before deriving optimal prices. Since the worst-case priors are the same as in the MMEU case, see Corollary 5.2, I proceed by characterizing the best-case priors.

Corollary 5.4. The monopolist's best-case priors

$$s_{best} := \arg\max_{s \in [\underline{s}, \overline{s}]} E_s[\Pi]$$

are given by

$$s_{best} = \begin{cases} \underline{s} & \text{for} \quad p < c \\ [\underline{s}, \overline{s}] & \text{for} \quad p = c \\ \overline{s} & \text{for} \quad c < p < \frac{a_H}{b} \\ [\underline{s}, \overline{s}] & \text{for} \quad p \geq \frac{a_H}{b}. \end{cases}$$

Proof. The proof is contained in the last section of this chapter.

By using Corollary 5.2 and Corollary 5.4, the α-MEU-objective can be rewritten in the following way:

$$\Psi_\alpha(p) = \begin{cases} \alpha E_{\overline{s}}[\Pi] + (1-\alpha) E_{\underline{s}}[\Pi] & \text{for} \quad 0 \leq p < c \\ \alpha E_{\underline{s}}[\Pi] + (1-\alpha) E_{\overline{s}}[\Pi] & \text{for} \quad p \geq c \end{cases} \tag{5.4}$$

Corollary 5.5. Objective (5.4) can be rewritten in the subsequent manner:

$$\Psi_\alpha(p) = \begin{cases} E_{s_1(\alpha)}[\Pi] & \text{for} \quad 0 \leq p < c \\ E_{s_2(\alpha)}[\Pi] & \text{for} \quad p \geq c \end{cases}$$

where

$$s_1(\alpha) := \alpha \overline{s} + (1-\alpha)\underline{s} \quad \text{and} \quad s_2(\alpha) := \alpha \underline{s} + (1-\alpha)\overline{s}.$$

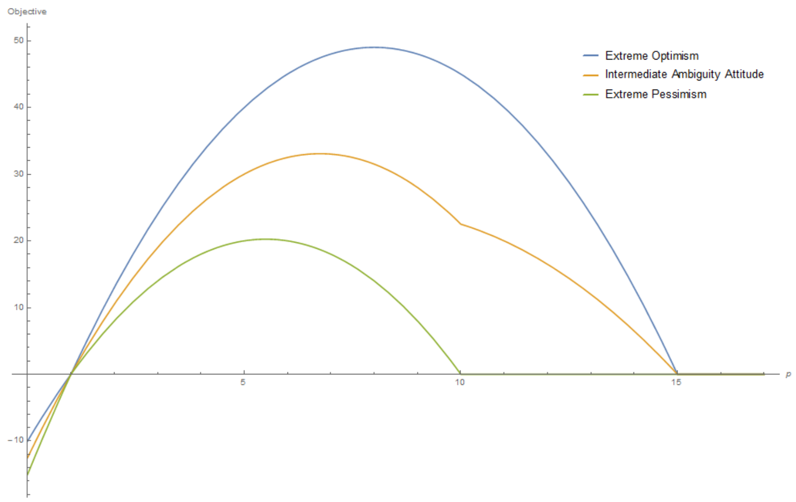


FIGURE 5.3: Objectives for $\alpha = 0$, $\alpha = \frac{1}{2}$, and $\alpha = 1$

Moreover, we can interpret $s_i(\alpha)$ as a distorted probability

$$0 \leq s_i(\alpha) \leq 1.$$

Proof. The proof is contained in the last section of this chapter.

The following corollary characterizes the solution of the monopolist's optimization problem under α-MEU.

Corollary 5.6. The following prices are possible solutions of the monopolist's optimization problem under α-MEU preferences:

1. $p_\alpha^* = \frac{s_2(\alpha)a_H + (1-s_2(\alpha))a_L + bc}{2b}$
2. $p_\alpha^{**} = \frac{a_H + bc}{2b}$
3. $p_\alpha^{***} \in [p_1, \infty]$

Which of these candidates is the global maximizer depends on the Conditions C_1 to C_7 and the threshold value $\hat{s}$ defined in the risk case. Note that it is vital to replace the probability s with $s_2(\alpha)$ in Conditions C_2 to C_6. The price p_α^* is the only solution of the optimization problem if one of the Conditions C_1, C_2, or C_3 is satisfied. The price p_α^{**}

is the only solution of the optimization problem if Requirement C_4 or Requirement C_5 holds. Both p^*_α and p^{**}_α are the solution of the optimization problem if Condition C_6 is met. Every $p^{***}_\alpha \in [p_1, \infty]$ is a solution of the optimization problem if Condition C_7 holds.

Proof. The proof is contained in the last section of this chapter.

Now, we can answer the initial question of this paragraph. How does the monopoly price change if the monopolist is more optimistic? Taking the derivative of p^*_α with respect to α yields

$$\frac{\partial p^*_\alpha}{\partial \alpha} = \frac{(\underline{s} - \overline{s})(a_H - a_L)}{2b} < 0.$$

Hence, a higher degree of optimism yields a higher monopoly price in cases where the interior solution p^*_α is optimal. The monopoly price is independent of α in cases where the corner solutions p^{**}_α and p^{***}_α are optimal. Moreover, the α-MEU approach contains the the MMEU-example as a special case for $\alpha = 1$.

The problem of the α-MEU model is that it still lacks a sound axiomatic foundation. As a consequence, agents' choices are based on a heuristic approach. This complicates the empirical validation of the model as well. If an axiomatic foundation is desirable or necessary, the modeler needs to discard the α-MEU model. Another argument against the α-MEU and MMEU model is that both frameworks cannot accommodate how strongly decision-makers are exposed to ambiguity. Is it really realistic to assume that the magnitude of ambiguity has no influence on the monopolist's decisions? Imagine, for instance, that the monopolist can rely on a data set on market demand, which has some predictive power. However, due to personal experience, the monopolist knows that she cannot fully trust this data set. The α-MEU approach would prescribe that the monopolist ignores the data set. Instead, she would look at a convex combination of the worst and best possible outcome and make her decision according to this decision rule. Prominent models of decision making under uncertainty accommodating the magnitude of ambiguity are the Choquet-expected utility model, the variational model of preferences, and the KMM model.

Among these models, the Choquet-expected utility model provides the most striking difference from a structural point of view, since it is the only decision-theoretic model dismissing the assumption of additive probabilities. The variational and the KMM model retain additivity[9] and rely therefore on a set of priors. Hence, the special appeal of the CEU-model results from the fact that it can model beliefs with a single structural component, the underlying capacity. One of the main advantages of the Choquet model lies in its tremendous flexibility in modeling beliefs. An interesting class of capacities are neo-additive capacities, since they allow for the separation of confidence, or the degree of ambiguity, and a decision-maker's attitude towards ambiguity. In addition, neo-additive capacities incorporate the α-MEU heuristics and the MMEU model as special cases.

Example 5.5 (Monopoly under Neo-Additive Capacities)**.** In this example, a more formal approach is required to derive the monopolist's objective. To begin with, assume that there is a state space Ω consisting of two elements ω_1 and ω_2 with the following interpretations: The event $\{\omega_1\}$ occurs when the monopolist faces the high consumer scenario a_H. Similarly, the event $\{\omega_2\}$ occurs in the low consumer scenario a_L. Moreover, the monopolist's beliefs are represented by a finite set of priors $\mathcal{P} \subset \Delta(\Omega)$ where $\Delta(\Omega)$ denotes the simplex

$$\Delta(\Omega) := \Big\{(s_1, s_2) : s_1 + s_2 = 1,\ s_1 \geq 0,\ s_2 \geq 0\Big\}.$$

Since the state space consists of two elements only, we can represent every prior in $(s_1, s_2) \in \Delta(\Omega)$ by its first component $s_1 \in [0, 1]$, which is the probability of the high demand scenario. Using this simplified notation, we can describe the prior set $\mathcal{P}$ by

$$\mathcal{P} = \Big\{s_1, ..., s_n : s_i \in [0, 1],\ s_i \neq s_j \text{ for } i \neq j\Big\}.$$

[9]To be more precise: the KMM model relies on a set of second-stage priors.

Without loss of generality, I assume that the probabilities in $\mathcal{P}$ are ordered in the following way:

$$0 \leq \underline{s} := s_1 < s_2 < ... < s_n =: \overline{s} \leq 1$$

If the monopolist knew the true underlying probability s, she could proceed by maximizing her expected profit with respect to s. In this case, the solution would correspond to the risk case treated in Example 5.1. In this example, s is unknown. In fact, the monopolist can derive for every fixed price $p \geq 0$ and every prior $s_i \in \mathcal{P}$ an expected profit of the form

$$E_{s_i}[\Pi](p) := (p-c)\Big(s_i D_1(p) + (1-s_i)D_2(p)\Big).$$

Henceforth, $X(p)$ denotes the collection of these expected profits. Formally,

$$X(p) := \Big\{E_{s_i}[\Pi](p) : \ s_i \in \mathcal{P}\Big\}.$$

In the following, I assume that there is a second-stage belief ν defined on the set of first-stage priors $\mathcal{P}$, which reflects the monopolist's uncertainty with respect to the beliefs in the prior set $\mathcal{P}$. The value $\nu(s)$ denotes the monopolist's belief, or subjective probability, that $(s, 1-s)$ is the true distribution of possible demand realizations in the future. The belief ν is assumed to be a neo-additive capacity. The following definition of a neo-additive capacity is adopted from Eichberger et al. [2009], page 359:

Definition 5.1. Let $q = (q_1, ..., q_n)$ be a probability measure on $(\mathcal{P}, \Sigma)$ where Σ denotes a σ-algebra of events on $\mathcal{P}$. Then, for real numbers α and δ we can define a neo-additive capacity ν by $\nu(\varnothing) = 0$, $\nu(\Phi) = 1$, $\nu(A) = \delta\alpha + (1-\delta)q(A)$ where $A \in \Sigma$ is a nonempty and strict subset of $\mathcal{P}$.

The monopolist is assumed to be a Choquet-expected utility maximizer, see Schmeidler [1989] for an axiomatization of Choquet expected utilities. The Choquet model presumes that a decision-maker maximizes a Choquet integral with respect to a capacity.

Chateauneuf et al. [2007] demonstrate for the class of neo-additive beliefs that the respective Choquet integral can be expressed in the following way:[10]

$$\int_{\mathcal{P}} f d\nu = (1-\delta)E_q[f] + \delta\Big(\alpha \max\{x : f^{-1}(x) \notin \mathcal{N}\} + (1-\alpha)\min\{x : f^{-1}(x) \notin \mathcal{N}\}\Big) \quad (5.5)$$

In this context, $f : \mathcal{P} \to \mathbb{R}$ denotes a measurable function with finite range. $\mathcal{N} = \{A \in \Sigma : \nu(A) = 0\}$ is the collection of null-events of ν, $E_q[f]$ is the expectation of f with respect to the probability distribution q, $\max\{x : f^{-1}(x) \notin \mathcal{N}\}$ denotes the best case of f given x is not the realization of a null-event and $\min\{x : f^{-1}(x) \notin \mathcal{N}\}$ denotes the worst-case of f given x is not the realization of a null-event.
In this example, f depends on prices and assigns to each prior $s_i \in \mathcal{P}$ an expected profit $E_{s_i}[\Pi](p)$. Formally, we consider for each fixed p a mapping

$$f(s_i|p) := E_{s_i}[\Pi](p). \quad (5.6)$$

This definition of f corresponds to the intuition that the monopolist maximizes a generalized average of expected profit functions. The difference to a standard risk approach is that this average is taken with respect to a distorted probability ν. Using the definition of $f(\cdot|p)$, we can derive the monopolist's objective. It is

$$\begin{aligned} \mathcal{U}(p) := \int_{\mathcal{P}} f(\cdot|p) d\nu = (1-\delta) \sum_{i=1}^{n} q_i E_{s_i}[\Pi](p) \\ + \delta \left[\alpha \max_{s_i \in \mathcal{P}} E_{s_i}[\Pi](p) + (1-\alpha) \min_{s_i \in \mathcal{P}} E_{s_i}[\Pi](p)\right] \end{aligned} \quad (5.7)$$

The expectation $\sum_{i=1}^{n} q_i E_{s_i}[\Pi](p)$ can be rewritten in the following way:

[10]See Lemma 3.1 in Chateauneuf et al. [2007], page 541.

$$\begin{aligned}
\sum_{i=1}^{n} q_i E_{s_i}[\Pi](p) &= \sum_{i=1}^{n}(p-c)q_i(s_iD_1(p)+(1-s_i)D_2(p)) \\
&= (p-c)\sum_{i=1}^{n}(q_is_iD_1(p)+q_i(1-s_i)D_2(p)) \\
&= (p-c)(\tilde{s}D_1(p)+(1-\tilde{s})D_2(p)) \\
&= E_{\tilde{s}}[\Pi](p)
\end{aligned}$$

where

$$\tilde{s} = \sum_{i=1}^{n} q_i s_i.$$

Hence, we obtain the following representation of the objective:

$$\mathcal{U}(p) = (1-\delta)E_{\tilde{s}}[\Pi] + \delta\left(\alpha \min_{s\in\mathcal{P}} E_s[\Pi] + (1-\alpha)\max_{s\in\mathcal{P}} E_s[\Pi]\right) \tag{5.8}$$

The following corollary gives the solution of the monopolist's problem under neo-additive capacities.

Corollary 5.7. The following prices are possible solutions of the monopolist's optimization problem under neo-additive capacities:

1. $p^{*}_{\alpha,\delta} = \frac{s_2(\alpha,\delta)a_H+(1-s_2(\alpha,\delta))a_L+bc}{2b}$
2. $p^{**}_{\alpha,\delta} = \frac{a_H+bc}{2b}$
3. $p^{***}_{\alpha,\delta} \in [p_1,\infty]$

The parameter $s_2(\alpha,\delta)$ is defined as

$$\begin{aligned}
s_2(\alpha,\delta) &= (1-\delta)\tilde{s} + \delta(\alpha\underline{s} + (1-\alpha)\overline{s}) \\
&= (1-\delta)\sum_{i=1}^{n} q_is_i + \delta(\alpha s_1 + (1-\alpha)s_n)
\end{aligned} \tag{5.9}$$

and can be interpreted as a distorted probability. Which of these price candidates is the global maximizer depends on the Conditions C_1 to C_7 and the threshold value $\hat{s}$ defined in the risk case. Note that it is vital to replace the probability s with $s_2(\alpha, \delta)$ in Conditions C_2 to C_6. The price $p^*_{\alpha,\delta}$ is the only solution of the optimization problem if one of the Conditions C_1, C_2, or C_3 is satisfied. The price $p^{**}_{\alpha,\delta}$ is the only solution of the optimization problem if Condition C_4 or Condition C_5 holds. Both $p^*_{\alpha,\delta}$ and $p^{**}_{\alpha,\delta}$ are the solution of the optimization problem if Requirement C_6 is met. Every $p^{***}_{\alpha,\delta} \in [p_1, \infty]$ is a solution of the optimization problem if C_7 holds.

Proof. The proof is contained in the last section of this chapter.

Corollary 5.8. In cases where the interior solution $p^*_{\alpha,\delta}$ is optimal, the following c.p. comparative static results hold:

1. A higher degree of optimism yields a higher monopoly price.

2. There exists a threshold value $\hat{\alpha}$, which is given by $\hat{\alpha} = 1 - \tilde{s}$, such that an increase in δ yields a higher monopoly price for $\alpha < \hat{\alpha}$ and a lower monopoly price for $\alpha > \hat{\alpha}$. In the special case $\alpha = \hat{\alpha}$, the monopoly price is independent of δ.

Proof. The proof is contained in the last section of this chapter.

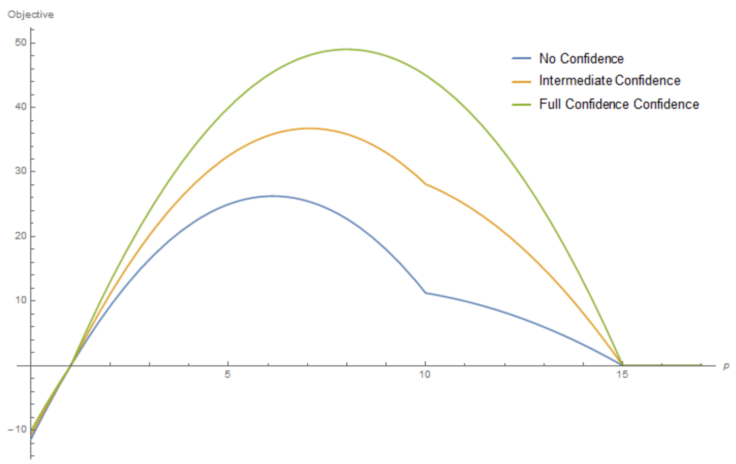


FIGURE 5.4: Objectives for $\delta = 0$, $\delta = 0.5$, and $\delta = 1$ in Case of Extreme Optimism

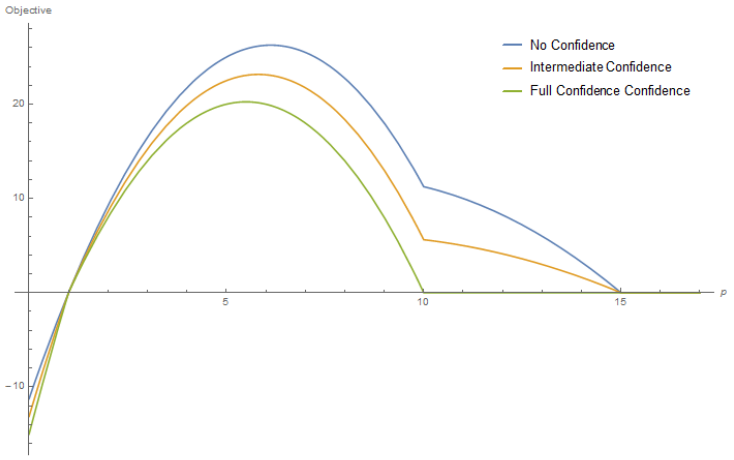


FIGURE 5.5: Objectives for $\delta = 0$, $\delta = 0.5$, and $\delta = 1$ in Case of Extreme Pessimism

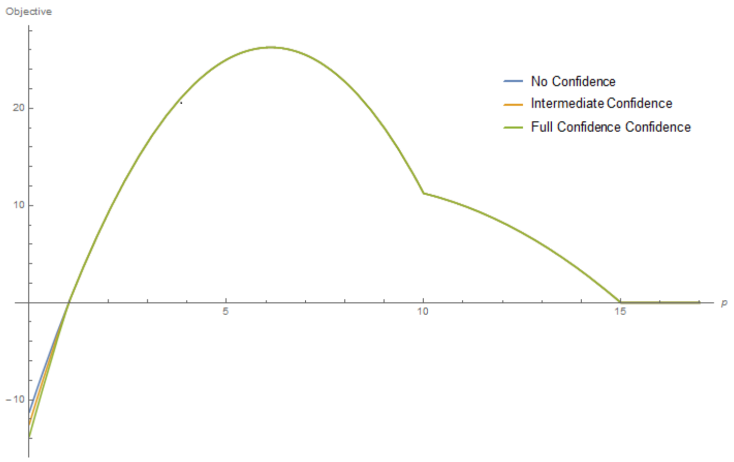


FIGURE 5.6: Objectives for $\delta = 0$, $\delta = 0.5$, and $\delta = 1$ in the Intermediate Case $\alpha = \hat{\alpha}$

The results of Corollary 5.8 can be interpreted in the following way: The fact that a higher degree of optimism yields a higher monopoly price is in line with the results from the MMEU and α-MEU example. As the monpolist expects a lower demand, she adjusts prices downwards. The special feature of the neo-additive model is the parameter δ, which captures how strongly the monopolist incorporates the expected profit $E_s[\Pi]$ into the objective function. When δ increases, the monopolist gives a lower decision weight to the expectation $E_s[\Pi]$. Whether the monopolist adjusts prices up- or downwards depends on the optimism parameter α. If the monopolist is extremely optimistic, then a higher value of δ increases the monopoly price. The converse statement holds if the monopolist is extremely pessimistic. For an intermediate value $\alpha = \hat{\alpha}$ the monopoly price is independent of δ and remains unchanged as δ increases.

An alternative to the Choquet model is the variational model of preference developed by Maccheroni et al. [2006], which incorporates the so-called multiplier preferences developed by Hansen and Sargent [2001], and the MMEU-model as special cases. An important distinguishing feature of the variational model is that it allows preferences to change whenever there are either changes in the variability of outcomes, or outcomes are shifted up- or downwards, see Machina and Siniscalchi [2014], page 32. The variational model applies to situations where the commitment to a particular probabilistic scenario incurs a cost to the decision-maker. In the multiplier model, this cost takes a specific form, which can be interpreted as the cost of deviating from an underlying reference probability function.

Example 5.6 (Monopoly under Multiplier Preferences)**.** Given multiplier preferences, the monopolist's optimization problem is given by

$$\max_{p\geq 0} \min_{s\in[\underline{s},\overline{s}]} \Big\{ E_s[\Pi] + \gamma R(s, s^*) \Big\}$$

where $R(s, s^*)$ denotes the relative entropy, or Kullback-Leibler divergence, of s^* from s; $\gamma \geq 0$ is a parameter and $\mathcal{P} = [\underline{s}, \overline{s}]$ is a compact and convex set of priors with $0 \leq \underline{s} < \overline{s} \leq 1$. In cases where γ equals zero, we obtain the MMEU objective. Consequently, the solution for this special case is already known. For $\gamma > 0$, the monopolist gives positive weight to the relative entropy. The relative entropy is always non-negative due to Gibb's inequality, see Falk [1970] for a proof. Moreover, we can infer that $R(s, s^*) = 0$ if the distributions of s and s^* coincide almost everywhere. The Kullback-Leibler divergence is a distance measure for probability distributions. Hence, if γ is positive, the monopolist faces two countervailing effects as he determines his worst-case prior. One of these effects arises from minimizing the expected profit $E_s[\Pi]$. As a consequence of Example 5.3, the worst-case prior of $E_s[\Pi]$ is given by the upper or lower boundary of the interval $[\underline{s}, \overline{s}]$, depending on whether prices are strictly smaller or larger than marginal costs. The second effect arises from minimizing the relative entropy $R(s, s^*)$. Since R attains the minimum for $s = s^*$, the monopolist faces a cost for not selecting a prior close to s^*. Throughout this example, I assume that $s^* \in [\underline{s}, \overline{s}]$. This restriction makes sense for the following reason: The monopolist considers s^* as a reference probability; if γ is strictly larger than zero, the monopolist takes into account the distance between the worst-case prior $\underline{s}$ and s^*, large distances from s^* are "punished" by a larger relative entropy. If that is the case, why should the monopolist discard s^* as a possible probabilistic scenario? An exclusion of s^* could therefore be considered as logically inconsistent.

In order to solve the optimization problem, I proceed by solving the prior-minimization-problem first.

Proposition 5.2. The minimizing prior

$$s_{min} := \arg \min_{s \in [\underline{s}, \overline{s}]} \Big\{ E_s[\Pi] + \gamma R(s, s^*) \Big\}$$

in the monopoly model with multiplier preferences is given by

$$s_{min} = \begin{cases} s_{worst} & \text{for} \quad \gamma = 0 \\ \overline{s} & \text{for} \quad \gamma > 0 \ \wedge \ p \in A_1 = [0, \min\{c, \hat{p}_1\}] \\ \frac{s^* f_1}{s^* f_1 + (1-s^*) f_2} & \text{for} \quad \gamma > 0 \ \wedge \ p \in A_2 = \big[\max\{0, \hat{p}_1\}, \min\{\frac{a_L}{b}, \hat{p}_2\}\big] \\ \underline{s} & \text{for} \quad \gamma > 0 \ \wedge \ p \in A_3 = \big[\max\{c, \hat{p}_2\}, \frac{a_L}{b}\big] \\ & \text{or} \quad 0 < \gamma \leq \hat{\gamma} \ \wedge \ p \in A_4 = \big[\max\{\frac{a_L}{b}, \hat{p}_3\}, \min\{\hat{p}_4, \frac{a_H}{b}\}\big] \\ \frac{s^* f_3}{s^* f_3 + (1-s^*) f_4} & \text{for} \quad 0 < \gamma \leq \hat{\gamma} \ \wedge \ p \in A_5 = \big[\frac{a_L}{b}, \min\{\hat{p}_3, \frac{a_H}{b}\}\big] \\ & \text{or} \quad 0 < \gamma \leq \hat{\gamma} \ \wedge \ p \in A_6 = \big[\max\{\frac{a_L}{b}, \hat{p}_4\}, \frac{a_H}{b}\big] \\ & \text{or} \quad \gamma > \hat{\gamma} \ \wedge \ p \in A_7 = \big[\frac{a_L}{b}, \frac{a_H}{b}\big] \\ s^* & \text{for} \quad 0 < \gamma \ \wedge \ p \in A_8 = [\frac{a_H}{b}, \infty) \\ & \text{or} \quad 0 < \gamma \ \wedge \ p \in A_9 = \{c\} \end{cases}$$

where s_{worst} denotes the worst-case prior derived in Corollary 5.2 and

(1) $f_1 = e^{\frac{a_H c + a_L p}{\gamma}}$, $f_2 = e^{\frac{a_L c + a_H p}{\gamma}}$, $f_3 = e^{\frac{a_H c + bp^2}{\gamma}}$, $f_4 = e^{\frac{(a_H + bc)p}{\gamma}}$

(2) $h_1 = \frac{\overline{s}(1-s^*)}{s^*(1-\overline{s})}$, $h_2 = \frac{\underline{s}(1-s^*)}{s^*(1-\underline{s})}$

(3) $\hat{p}_1 = c - \frac{\log(h_1)\gamma}{a_H - a_L}$

(5) $\hat{p}_2 = c - \frac{\log(h_2)\gamma}{a_H - a_L}$

(6) $\hat{p}_3 = \frac{a_H+bc}{2b} - \sqrt{\frac{(a_H+bc)^2}{4b^2} - a_H c + \gamma \log(h_2)}$

(7) $\hat{p}_4 = \frac{a_H+bc}{2b} + \sqrt{\frac{(a_H+bc)^2}{4b^2} - a_H c + \gamma \log(h_2)}$

(8) $\hat{\gamma} = \left[a_H c - \frac{(a_H+bc)^2}{4b^2}\right]$

Proof. The proof is contained in the last section of this chapter.

By making use of Proposition 5.2, we obtain the reduced optimization problem

$$\max_{p \geq 0} \left\{ E_{s_{min}}[\Pi] + \gamma R(s_{min}, s^*) \right\}.$$

Corollary 5.9. The monopoly problem with multiplier preferences has a solution.

Proof. The existence of a solution is guaranteed for $\gamma = 0$, since the objective reduces to the objective of the MMEU case. Hence, we continue by considering cases where $\gamma > 0$. From Proposition 5.2, we can infer that the minimum prior is at least a piecewise continuous function in p. To be more precise, we observe that the minimum prior s_{min} is continuous on each of the intervals A_i for $i = 1, ..., 9$. Since sums and products of continuous functions are continuous, we can deduce that the restriction of the objective to A_i is a continuous function. All intervals A_i are compact, except A_8. By using Weierstrass' theorem, we obtain that there is a nonempty set of maximizers $S_i \neq \emptyset$ for each domain restriction of the objective, except for the interval A_8. On the interval $A_8 = [\frac{a_H}{b}, \infty)$, the objective is zero. Hence, all $p \geq \frac{a_H}{b}$ are maximizers on A_8. As a consequence, $S_8 = A_8$. Since there is only a finite number of sets S_i, and since each set S_i is nonempty, we can conclude that the set of global maximizers is nonempty as well. □

The monopoly problem with multiplier preferences has in general no closed-form solution. This is because the objective's first order condition can only be solved numerically.

Corollary 5.10. The minimum prior s_{min} satisfies the subsequent limit property:

$$\lim_{\gamma \to \infty} s_{min}(\gamma) = s^*.$$

Henceforth, let $p(\gamma)$ denote the solution of the monopoly problem under multiplier preference for a given parameter γ. Then,

$$\lim_{\gamma \to \infty} p(\gamma) = p^{risk}(s^*)$$

where $p^{risk}(s^*)$ denotes the solution in the risk case with the prior $s = s^*$. Moreover, the following comparative static result holds: If $p(\gamma_1)$ and $p(\gamma_1)$ are two interior solutions, which satisfy the requirements

1. $c \leq p(\gamma_i) \leq \frac{a_L}{b}$
2. $0 < \hat{\gamma} < \gamma_1 < \gamma_2 < \infty$,

we can conclude that $p(\gamma_1) \leq p(\gamma_2)$. Note that the parameter $\hat{\gamma}$ denotes the threshold value defined in Corollary 5.2.

Proof. The proof is contained in the last section of this chapter.

Figure 5.7 displays the monopolist's objective for different values of γ. The monopoly model with multiplier preferences incorporates two extreme cases. For $\gamma = 0$, we obtain the MMEU model with a monopolist exhibiting extreme pessimism and a low monopoly price. For $\gamma \to \infty$, the monopolist increasingly assumes that the reference prior s^* is the true underlying distribution. Hence, he or she, adjusts prices upwards. Consequently, the parameter γ can be considered as a measure for the monopolist's attitude towards

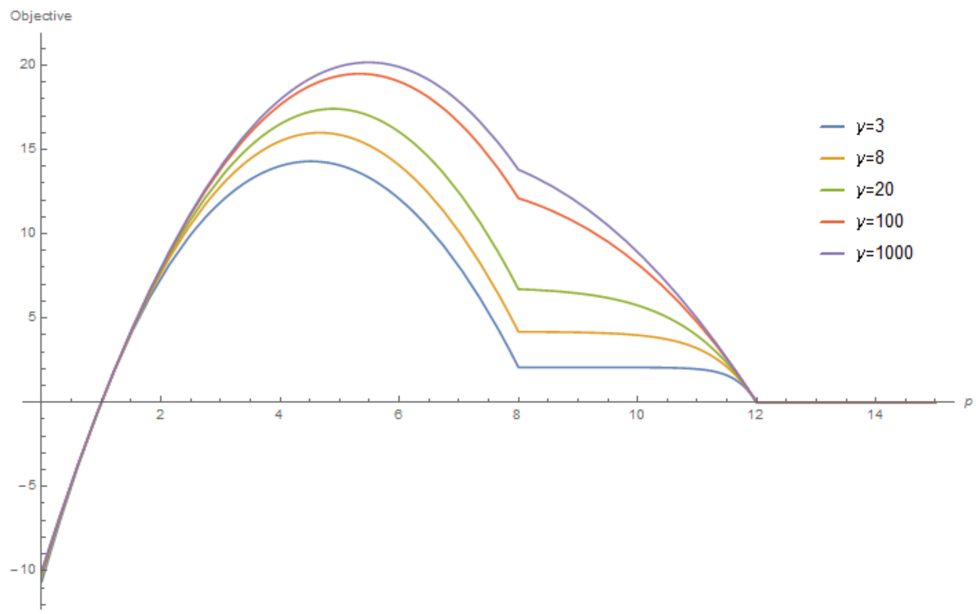


FIGURE 5.7: Objectives for $\gamma = 3$, $\gamma = 20$, $\gamma = 100$, and $\gamma = 1000$

ambiguity. The difference to the α-MEU and Choquet model with neo-additive capacities is that the monopolist compares the worst case with a distance measure for probability distributions, and not with the best-case distribution. A problem of the multiplier approach is that, contrary to the Choquet model with neo-additive capacities, there is no clear separation between the monopolist's attitude towards ambiguity and his confidence into the reference probability s^*. In fact, as γ increases, the monopolist's confidence into the reference probability s^* increases. At the same time, the monopolist exhibits a lower degree of pessimism, since the relative weight of the worst-case scenario decreases. Another drawback of the multiplier model is that it does not generate a tractable closed-form solution for the monopoly price.

A prominent model of decision making under uncertainty is the KMM or Smooth Ambiguity Model. From a practical point of view, the Smooth model has the advantage that it allows for the application of differential calculus, and is therefore frequently considered as a framework which generates tractable results.[11] The KMM model allows for a broad spectrum of different attitudes towards ambiguity which are condensed in a so-called distortion or transformation function. One of the most important features of the KMM

[11] See Machina and Siniscalchi [2014], page 27.

model lies in the fact that it differentiates between first- and second-order beliefs, and that it precludes, in general, the reduction of compound lotteries.

Example 5.7 (Monopoly with KMM preferences)**.** Assume that there is a set of first-stage priors $\mathcal{P}_1$, which consists of two elements $\underline{s}$ and $\overline{s}$ with $0 \leq \underline{s} < \overline{s} \leq 1$. Hence, there are two probabilistic scenarios, an optimistic one where the monopolist assumes that the probability for the high demand scenario is given by $\overline{s}$, and a pessimistic one where probability for the high demand scenario is given by $\underline{s}$. Moreover, suppose the monopolist has a second-stage prior $(q, 1-q)$. The parameter q is the monopolist's subjective probability that the pessimistic scenario $\underline{s}$ is the true probabilistic scenario. Similarly, $1-q$ denotes the monopolist's subjective probability that the optimistic scenario $\overline{s}$ reflects the true underlying distribution. In general, the KMM model allows for more than one second-stage distribution. For the sake of simplicity, I restrict the analysis to cases with a single second-stage probability. Under these assumptions, we obtain the following objective for the monopolist:

$$\Pi_{\Phi}^{KMM} = E_q[\Phi(E_s[\Pi])] = q\Phi(E_{\underline{s}}[\Pi]) + (1-q)\Phi(E_{\overline{s}}[\Pi]) \tag{5.10}$$

A distinguishing feature of the KMM-model is the distortion function Φ. If Φ equals the identity, the decision-maker is termed "ambiguity-neutral", compare Klibanoff et al. [2005] page 1862 for a definition. Similarly, an agent is called "ambiguity-averse" in cases where Φ is concave and "ambiguity-loving" in cases where Φ is convex. An ambiguity-neutral decision-maker is able to reduce compound lotteries to a simple lottery. This special case can be illustrated by means of the monopolist's objective. We obtain

$$\Pi_{Id}^{KMM} = qE_{\underline{s}}[\Pi] + (1-q)E_{\overline{s}}[\Pi] = E_{q\underline{s}+(1-q)\overline{s}}[\Pi].$$

As a consequence, the solution of the monopolist's optimization problem can be derived from Proposition 5.1 by replacing s with the prior $q\underline{s} + (1-q)\overline{s}$. For the general case, there is no closed-form solution for the monopoly price. Nevertheless, we can compare the monopoly price under "neutrality" with situations where the monopolist is ambiguity-averse or ambiguity-loving.

Proposition 5.3. Consider two twice continuously differentiable transformation functions Φ_1 and Φ_2 where Φ_1 is strictly concave and Φ_2 is strictly convex. Moreover, I denote with p_{Φ_i} the respective monopoly prices under KMM-preferences. Then,

$$p_{\Phi_1} < p_{Id} < p_{\Phi_2}$$

if both prices are interior solutions with $D_i(p_{\Phi_j}) > 0$ for all i, $j = 1, 2$. In all other cases, we can infer that $p_{\Phi_1} = p_{\Phi_2}$.

Proof. The proof is contained in the last section of this chapter.

An important class of distortion functions is the so-called class of transformations displaying constant ambiguity aversion, see Klibanoff et al. [2005], which is defined by

$$\Phi_a(x) = \begin{cases} \frac{1-e^{-ax}}{1-e^{-a}} & \text{for} \quad a \neq 0 \\ x & \text{for} \quad a \neq 0. \end{cases} \tag{5.11}$$

Obviously, this definition of constant ambiguity aversion is adopted from the constant absolute risk aversion utility function, see Pratt [1964]. The parameter a captures the monopolist's attitude towards ambiguity. If a is negative, the monopolist is called ambiguity loving, if $a = 0$, the monopolist is called ambiguity neutral, and if $a > 0$, the monopolist is called ambiguity-averse. Surprisingly, it is not true that a higher c.p. degree of absolute ambiguity aversion translates into a lower monopoly price. This is demonstrated by means of Table 5.1.

The reason for this behavior lies in the curvature of the constant absolute ambiguity transformation. It is

$$\frac{\partial^2 \Phi_a(x)}{\partial x^2} = -\frac{a^2 e^{-ax}}{1 - e^{-a}}.$$

Parameter a	**Optimal Price**
0.1	9.06
1.1	5.99
2.1	5.75
3.1	4.72
4.1	4.96
5.1	4.89
6.1	5.69

TABLE 5.1: Monopoly Price for Different Degrees of Absolute Ambiguity Aversion Rounded to Two Decimal Places

Taking the derivative of this function with respect to a yields

$$\frac{\partial^3 \Phi_a(x)}{\partial x^2 \partial a} = \frac{a e^{a-ax} \left(a(-x) + e^a(ax-2) + a + 2\right)}{\left(e^a - 1\right)^2}.$$

We can readily see that the sign of this function is dependent on x. In fact, one can show for positive values of a that there is a threshold value $\hat{x}$ such that $\frac{\partial^3 \Phi_a(x)}{\partial x^2 \partial a}$ is positive for $x > \hat{x}$, zero for $x = \hat{x}$, and negative for $x < \hat{x}$. This implies that an increase in the parameter a can reduce or increase the concavity of Φ_a depending on the value of x. Note that this observation is a direct consequence of a well-known property of the constant absolute risk aversion utility, since a decision-maker's risk attitude is dependent on the initial wealth level. Naturally, the same holds for the KMM-model under ambiguity where the initial "wealth levels" are given by the expectations $E_{\underline{s}}[\Pi]$ and $E_{\overline{s}}[\Pi]$.

5.4 Conclusion

This chapter provides a comparison of the most prominent models of decision making under uncertainty by means of a simple baseline model, the monopoly market with linear demand. The monopolist is assumed to be risk-neutral. Ambiguity is introduced in the form of demand uncertainty. In particular, I consider a scenario with two possible demand realizations. Both demand realizations have the same slope but different intercepts. As it turns out, there is a closed form solution for the monopoly price for the MEU

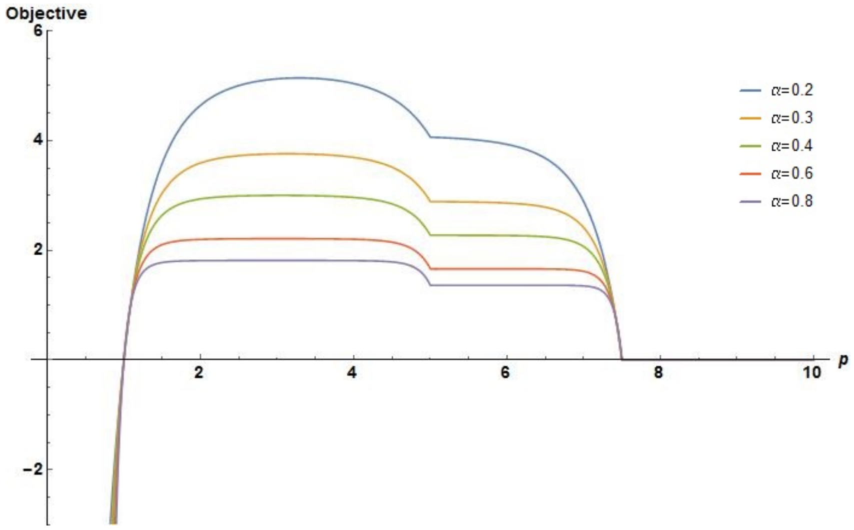


FIGURE 5.8: Objective for Different Values of the Absolute Ambiguity Aversion Parameter a.

model, the α-MEU model and the Choquet model with neo-additive capacities. There is no-closed-form solution for the multiplier model and the Smooth Ambiguity Model. Depending on the underlying parameter constellations, three types of solutions can occur. The first one is that there is a unique monopoly price where demand is positive in both states of the world. The second class of solutions occurs when the monopoly price exceeds the threshold value $\frac{a_L}{b}$ but remains smaller than the threshold value $\frac{a_H}{b}$. In this case, demand for the low consumer case becomes zero and remains positive for the high number of consumers case. Again, there is a unique monopoly price. The third class of solutions emerges if the monopoly price exceeds the threshold value $\frac{a_H}{b}$. In this case, demand is zero in both states of the world. Hence, all prices $p \geq \frac{a_H}{b}$ are optimal.

The monopoly price in the MEU preference specification is smaller than the monopoly price under certainty. Hence, extreme pessimism yields a lower monopoly price. A common feature between the α-MEU model and the Choquet model with neo-additive capacities is that a higher degree of optimism yields a higher monopoly price. Additionally, the Choquet model with neo-additive capacities allows for a ceteris paribus analysis with respect to the confidence parameter δ. If the monopolist is sufficiently pessimistic, we can conclude that a higher degree of confidence (lower value of δ) yields a higher monopoly price. On the other hand, if the monopolist is sufficiently optimistic, an increase in confidence translates into a lower monopoly price. For an intermediate pessimism value $\hat{\alpha}$, the monopoly price remains unchanged as δ decreases. In cases where preferences are modeled within the multiplier framework, we observe that a higher value of γ yields a higher monopoly price. In contrast to the Choquet model with neo-additive capacities, the multiplier model cannot clearly differentiate between optimism and confidence. In particular, an increase in γ can be interpreted as an increase in confidence and optimism at the same time. We obtain that a higher value of γ yields an increasing monopoly price. This is in line with the results of the α-MEU model and the Choquet model with neo-additive capacities. Under the KMM model, we obtain that the monopoly price under ambiguity aversion is lower than the monopoly price under neutrality. Similarly, the monopoly price exceeds the monopoly price under neutrality in cases where the monopolist is ambiguity-loving. A widely-used class of distortion functions for the KMM model is the class of transformations displaying constant ambiguity aversion. Contrary to the optimism parameter α of the α-MEU and the Choquet model with neo-additive capacities, it is not necessarily true that a higher absolute ambiguity aversion parameter a entails a lower monopoly price under KMM preferences.

5.5 Mathematical Proofs

Proof of Corollary 5.1. The optimization problem is solved in two steps:

(1) *Step 1*: Identify the local maxima of $\Pi(p)$ for $0 \leq p \leq p_0$ and $p \geq p_0$.

(2) *Step 2*: Determine the global maximum.

By solving the first order condition $\Pi'(p) = 0$ of the unconstrained profit function $\Pi(p) = (p-c)(a-bp)$ for p, we obtain the well-known monopoly price

$$p^M = \frac{a+bc}{2b}$$

The profit function evaluated at p^M yields

$$\Pi\left(p^M\right) = \frac{(a-bc)^2}{4}.$$

As Π is strictly concave $\Pi''(p) = -2b < 0$ and non-negative at the point p^M, we can infer that p^M is the unique solution of the firm's optimization problem if $p^M \in (0, p_0)$. Since a and b are assumed to be strictly positive, we can deduce that $p^M > 0$. In cases where $p^M \geq p_0$, we can conclude that the unconstrained profit function is monotonically increasing in p. This is because $(p-c)(a-bp)$ is a parabola opening downwards with a unique global maximum at p^M. Since $p^M > p_0$, the profit function $\Pi(p)$ is strictly increasing as long as $p \leq p_0$, otherwise p^M would not be the global maximum of the unconstrained profit function. Therefore, the local maximum of $\Pi(p)$ on $[0, p_0]$ is attained for $p = p_0$. The profit function equals zero for $p > p_0$. Thus, $\Pi(p) = 0$ for $p \geq p_0$. Moreover, due to the fact that Π is strictly increasing for $p < p_0$, we have $\Pi(p) < 0$ for $p < p_0$. To sum up, the monopolist can secure a profit of zero for all prices $p \geq p^*$, which are therefore optimal. Finally, the condition $p^M < p_0$ is equivalent to the condition $a > bc$. □

Proof of Proposition 5.1. The optimization problem is solved in two steps: First, the local maximizers of the objective are determined for the subsequent cases:

(1) $0 \leq p \leq p_0$

(2) $p_0 \leq p \leq p_1$

(3) $p \geq p_1$

Secondly, the overall global optimum is identified by comparing all local optima.

Case 1: $0 \leq p \leq p_0$

Solving the first order condition for p, we obtain the price

$$p^* = \frac{sa_H + (1-s)a_L + bc}{2b}.$$

The profit function evaluated at $p = p^*$ yields

$$\frac{(a_L\, s - a_H\, s + b\, c - a_L)^2}{4\, b}$$

The objective function is strictly concave with $E_s[\Pi]'' = -2b < 0$. Therefore, p^* is a local maximizer if $0 \leq p^* \leq p_0$. The case $p^* < 0$ is obsolete, since $a_L > 0$, $a_H > 0$, $b > 0$, and $c > 0$ by assumption. It follows that $p^* \leq p_0$ if and only if

$$2a_L - sa_H - (1-s)a_L \geq bc. \tag{5.12}$$

If $p^* > p_0$, we can deduce that the objective is strictly increasing for $0 \leq p \leq p_0$. Thus, the maximum is attained for $p^* = p_0$. The results of this case can be summarized as follows:

$$\arg\max_{0 \leq p \leq p_0} E_s[\Pi](p) = \begin{cases} p^* & \text{for} \quad 2a_L - sa_H - (1-s)a_L \geq bc \\ p_0 & \text{for} \quad 2a_L - sa_H - (1-s)a_L < bc \end{cases}$$

Case 2: $p_0 \leq p \leq p_1$

In the second case, $p_0 \leq p \leq p_1$, we can infer that $D_2(p) = 0$. The objective reduces to

$$s(p-c)D_1(p).$$

Solving the first order condition for p yields

$$p^{**} = \frac{a_H + bc}{2b}.$$

Since the reduced objective function is strictly concave with $E_s[\Pi]'' = -2sb < 0$, we can deduce that p^{**} is the global maximizer of $s(p-c)D_1(p)$. In order to fully solve the second case, we must distinguish between the following subcases:

(2a) $p^{**} < p_0$

(2b) $p^{**} \in [p_0, p_1]$

(2c) $p^{**} > p_1$

Observe that in Subcase (a), the global maximizer of $s(p-c)D_1(p)$ is smaller than p_0. The case defining condition $p^{**} < p_0$ is equivalent to the condition $bc < 2a_L - a_H$. Since $s(p-c)D_1(p)$ is a parabola opening downwards, the local maximum is attained at $p = p_0$. In Case (b), the interior solution $p = p^{**}$ is the local maximum. The case defining conditions $p^{**} \geq p_0$ and $p^{**} \leq p_1$ are equivalent to $a_H \geq bc$ and $a_H \geq 2a_L - bc$. Consequently, $a_H \geq \max\{bc, a_L - bc\}$.

In Subcase (c), the global maximizer of $s(p-c)D_1(p)$ is larger than p_1. The case defining condition $p^{**} > p_1$ is equivalent to the condition $a_H < bc$. Again, since $s(p-c)D_1(p)$ is a parabola opening downward, we can infer that the maximum is attained at $p = p_1$. The results for Case 2 can be summarized in the following way:

$$\arg\max_{p_0 \leq p \leq p_1} E_s[\Pi](p) = \begin{cases} p_0 & \text{for} \quad bc < 2a_L - a_H \\ p^{**} & \text{for} \quad a_H \geq \max\{bc, 2a_L - bc\} \\ p_1 & \text{for} \quad a_H < bc \end{cases}$$

Case 3: $p \geq p_1$

In the third case $p \geq p_1$, the objective is constantly zero. Hence, all points $p \geq p_1$ are optimal.

$$\arg\max_{p \geq p_1} E_s[\Pi](p) = [p_1, \infty)$$

Determine the Overall Solution

The second step of the analysis consists in comparing the local optima derived for the different cases discussed above. The subsequent lemma reduces substantially the number of cases to be considered.

Lemma 5.2. *In the monopoly model under risk the subsequent statements hold:*

(I) $(2a)$ *implies inequality* (5.12)*.*

(II) $(2c)$ *implies the negation of inequality* (5.12)*.*

Proof of Lemma 5.2. The lemma is verified by proving statements (I) and (III) separately.

Statement (I):

The case defining condition of Subcase $(2a)$ is given by $2a_L - a_H > bc$. This implies $2a_L - sa_H - (1-s)a_L \geq bc$, since $sa_H + (1-s)a_L < a_H$.

Statement (II):

The case defining condition of Subcase $(2c)$ is given by $a_H < bc$. This implies $a_L < bc$, since $a_L < a_H$ by assumption. Hence, $2a_L - sa_H - (1-s)a_L < 2a_L - a_L = a_L < bc$ which is the negation of (1). □

Taking into account the results of Lemma 5.2, we can derive the solution of the monopolist's optimization problem for the Subcases (a) and (c).

Subcase (2a):

In Subcase $(2a)$, the optimum is given by p^*. The local maximizer on the interval $[0, p_0]$ is p^*. Similarly, we obtain p_0 as the local maximizer on the interval $[p_0, p_1]$. Besides, the set of local maximizers on the interval $[p_1, \infty)$ is given by the complete interval $[p_1, \infty)$. Since $E_s[\Pi](p)$ is a continuous function, and since p^* is the global maximum of the unconstrained profit function

$$(p - c)(s(a_H - bp) + (1 - s)(a_L - bp)),$$

we can conclude that

$$E_s[\Pi](p^*) \geq E_s[\Pi](p_0).$$

Furthermore, as the profit function evaluated at p^* is non-negative, we can deduce that

$$E_s[\Pi](p^*) \geq E_s[\Pi](p_0) \geq 0 = \max_{p \geq p_1} E_s[\Pi](p).$$

Hence, p^* is the global maximizer.

Subcase (2c):

In Subcase $(2c)$, the solution is given by the interval $[p_1, \infty)$. The local maximum on $[0, p_0]$ is given by p_0, on $[p_0, p_1]$ by p_1 and on $[p_1, \infty)$ by the complete interval $[p_1, \infty)$. The objective is strictly increasing on $[0, p_1]$, since it is strictly increasing on each of the sub-intervals $[0, p_0]$ and $[p_0, p_1]$. Furthermore, by continuity

$$E_s[\Pi](p_1) = 0 = \max_{p \geq p_1} E_s[\Pi](p)$$

for all $p > p_1$. Hence, the solution of the monopoly problem is given by $[p_1, \infty)$.

Subcase (2b):

Remember, that the condition $a_H \geq \max\{bc, a_L - bc\}$ holds throughout the case. There are three candidates for the global optimum, p^*, p_0, and p^{**}.[12] Contrary to the Subcases $(2a)$ and $(2c)$, we cannot prove or disprove inequality (5.12). Consequently, we must differentiate between two cases. Under the assumption that (5.12) holds, we can infer that $p^* \in [0, p_0]$. Hence, p_0 can only be a solution of the monopolist's optimization problem if $p^* = p_0$. Thus, we can restrict the analysis to the following instances:

- p^* is the unique global maximizer
- p^{**} is the unique global maximizer
- Both p^* and p^{**} are global maximizers

The price p^* is the only solution of the monopolist's problem if

$$E_s[\Pi](p^*) > E_s[\Pi](p^{**}). \tag{5.13}$$

Both p^* and p^{**} are global maximizers if

$$E_s[\Pi](p^*) = E_s[\Pi](p^{**}). \tag{5.14}$$

The price p^{**} is a unique solution if

$$E_s[\Pi](p^*) < E_s[\Pi](p^{**}). \tag{5.15}$$

Solving the quadratic equation (5.14) for s, we obtain the subsequent pair of solutions: $s_1 = \frac{(b\,c - a_L)^2}{(a_L - a_H)^2}$ and $s_2 = 1$. Thus, we derived the threshold value $\hat{s} := s_1$. More specifically, we can conclude that inequality (5.13) holds for $s < s_1 < 1$. Observe that $s < s_1 < 1$ implies the inequality $s < \frac{bc - a_L}{a_L - a_H}$, which is equivalent to inequality (5.12). This demonstrates

[12]The interval $[p_1, \infty)$ can be excluded, since $E_s[\Pi](p_1)$ and $E_s[\Pi](p_0)$ are both non-negative.

that p^* is the global maximizer. Moreover, inequality (5.15) holds for $s_1 < s < 1$ implying that p^{**} is the global maximizer. Besides, we can infer that $E_s[\Pi](p^*) = E_s[\Pi](p^{**})$ for $s \in \{s_1, 1\}$, in which case both p^* and p^{**} are global maximizers. If the threshold value s_1 is larger than one, inequality (5.13) holds for all $0 \leq s \leq 1$. Moreover, the only solution of equation (5.14) is given by $s = 1$.

The case where the negation of inequality (5.12) holds can be analyzed in a similar fashion. Clearly, p^* can be ruled out as global maximizer, since $p^* \notin [0, p_0]$. Consequently, there are only two candidates, p_0 and p^{**}, for the global optimum. We can distinguish the following cases:

- p_0 is the unique global maximizer
- Both p^{**} and p_0 are global maximizers
- p^{**} is the unique global maximizer

The first and the second case can be excluded. By Condition ($2b$), we can infer that $p^{**} \in [p_0, p_1]$. This implies that $E_s[\Pi]$ is strictly increasing on $[p_0, p^{**}]$ Moreover, we know that $E_s[\Pi]$ is strictly increasing on the interval $[0, p_0]$. Due to the fact that $E_s[\Pi]$ is continuous, we can conclude that $E_s[\Pi](p_0) < E_s[\Pi](p^{**})$. Hence, p^{**} is the only maximizer when the negation of inequality (5.12) holds. □

Proof of Corollary 5.2. Since $D_1(p) > D_2(p)$ for all $c < p < \frac{a_H}{b}$, we can infer that the monopolist's profit in scenario a_H is always larger than the profit in scenario a_L for a given price $c < p < \frac{a_H}{b}$. Formally,

$$\Pi_1(p) = (p - c)D_1(p) > \Pi_2(p) = (p - c)D_2(p)$$

for all $c < p < \frac{a_H}{b}$. Consequently, every prior that gives a smaller weight to the first scenario induces smaller expected profits. This can be expressed formally in the following way. For every $1 \geq s_1 \geq s_2 \geq 0$, we have

$$(p-c)(s_1 D_1 + (1-s_1)D_2) \geq (p-c)(s_2 D_1 + (1-s_2)D_2),$$

for all $c \leq p \leq \frac{a_H}{b}$. Thus, the minimizing prior is given by $s_{worst} = \underline{s}$. If p is strictly smaller than c, profits are negative in both scenarios. Hence, $\Pi_1(p) < \Pi_2(p)$ for all $0 \leq p < c$ and the prior inducing minimum expected profits is given by $s_{worst} = \overline{s}$. If $p = c$ or $p \geq \frac{a_H}{b}$, the objective is zero. Therefore, all priors $p \in [\underline{s}, \overline{s}]$ are minimizing priors. Taking these results, we obtain

$$s_{worst} := \begin{cases} \overline{s} & \text{for} \quad p < c \\ [\underline{s}, \overline{s}] & \text{for} \quad p = c \\ \underline{s} & \text{for} \quad c < p < \frac{a_H}{b} \\ [\underline{s}, \overline{s}] & \text{for} \quad p \geq \frac{a_H}{b}. \end{cases}$$

□

Proof of Corollary 5.3. The reduced optimization problem (5.2) is structurally the same problem as the optimization problem in the risk case. This is because a price smaller than c induces negative profits for every $s \in [0,1]$. Thus, a price $p < c$ can never be a global optimum, since the monopolist can secure zero profits for price $p \geq p_1$. Note furthermore that

$$\Psi(p) = E_{\underline{s}}[\Pi] = 0$$

for $p \geq \frac{a_H}{b}$. As a consequence, the global optimum remains unchanged if we replace the objective $\Psi(p)$ with $E_{\underline{s}}[\Pi]$. In formal terms,

$$\max_{p \geq 0} \Psi(p) = \max_{p \geq c} \Psi(p) = \max_{p \geq c} E_{\underline{s}}[\Pi] = \max_{p \geq 0} E_{\underline{s}}[\Pi]$$

The last equality holds because a price $0 \leq p < c$ cannot be a global maximizer of $E_{\underline{s}}[\Pi]$ This is because the monopolist can secure a profit of zero. Thus, we obtain the solution under MMEU by replacing the prior s defined in Proposition 5.1 by the worst-case prior $\underline{s}$. □

Proof of Corollary 5.4. The proof is similar to the proof of Corollary 5.2. Since $D_1(p) > D_2(p)$ for all $c < p < \frac{a_H}{b}$, we can conclude that the monopolist's profit in scenario a_H is always larger than the profit in scenario a_L for a given price $c < p < \frac{a_H}{b}$. Formally,

$$\Pi_1(p) = (p-c)D_1(p) > \Pi_2(p) = (p-c)D_2(p)$$

for all $c < p < \frac{a_H}{b}$. Consequently, every prior that gives a larger weight to the first scenario induces higher expected profits. This can be expressed formally in the following way. For every $1 \geq s_1 \geq s_2 \geq 0$, we have

$$(p-c)(s_1 D_1 + (1-s_1)D_2) \geq (p-c)(s_2 D_1 + (1-s_2)D_2),$$

for all $c \leq p \leq \frac{a_H}{b}$. The maximizing prior is given by $s_{best} = \overline{s}$. If $p < c$ profits are negative in both scenarios. Hence, $\Pi_1(p) < \Pi_2(p)$ for all $0 \leq p < c$ and the prior inducing maximum expected profits is given by $s_{best} = \underline{s}$. If $p = c$, or $p \geq \frac{a_H}{b}$, the objective equals zero. Therefore, all priors $p \in [\underline{s}, \overline{s}]$ are maximizing priors. Taking these results, we obtain

$$s_{best} := \begin{cases} \underline{s} & \text{for} \quad p < c \\ [\underline{s}, \overline{s}] & \text{for} \quad p = c \\ \overline{s} & \text{for} \quad c < p < \frac{a_H}{b} \\ [\underline{s}, \overline{s}] & \text{for} \quad p \geq \frac{a_H}{b}. \end{cases}$$

□

Proof of Corollary 5.5. Assume, without loss of generality, that $p \geq c$. In this case, the objective is given by

$$\begin{aligned}
\Psi(p) &= \alpha E_{\underline{s}}[\Pi] + (1-\alpha)E_{\overline{s}}[\Pi] \\
&= (p-c)\alpha[\underline{s}D_1(p) + (1-\underline{s})D_2(p)] + (1-\alpha)(p-c)[\overline{s}D_1(p) + (1-\overline{s})D_2(p)] \\
&= (p-c)\Big[\alpha(\underline{s}D_1(p) + (1-\underline{s})D_2(p)) + (1-\alpha)(\overline{s}D_1(p) + (1-\overline{s})D_2(p))\Big] \\
&= (p-c)[(\alpha\underline{s} + (1-\alpha)\overline{s})D_1(p) + (\alpha(1-\underline{s}) + (1-\alpha)(1-\overline{s}))D_2(p)] \\
&= (p-c)[s_2(\alpha)D_1(p) + (1-s_2(\alpha))D_2(p)]
\end{aligned}$$

where

$$s_2(\alpha) := \alpha\underline{s} + (1-\alpha)\overline{s}.$$

The prior $s_2(\alpha)$ is non-negative, since $\underline{s} \geq 0$, $\alpha \geq 0$, and $\overline{s} \geq 0$. Moreover,

$$s_2(\alpha) \leq \overline{s} \leq 1.$$

A similar proof holds for $0 \leq p < c$ with the difference that $\Psi(p)$ equals

$$\alpha E_{\overline{s}}[\Pi] + (1-\alpha)E_{\underline{s}}[\Pi].$$

In this case, we can rewrite the objective by $\Psi(p) = E_{s_1(\alpha)}[\Pi]$ where

$$s_1(\alpha) := \alpha\overline{s} + (1-\alpha)\underline{s}.$$

Using the same reasoning as in the case $p \geq c$, we can conclude that $0 \leq s_1(\alpha) \leq 1$. $\square$

Proof of Corollary 5.6. The simplified objective $E_{s(\alpha)}[\Pi]$ is structurally equivalent to the objective of the risk case with the prior $s = s_2(\alpha)$. This is the case because the monopolist can secure a profit of zero for prices $p \geq p_1$. A price $p < c$ induces negative profits and is therefore never optimal. Hence, we can rewrite the optimization problem in the following way:

$$\max_{p\geq 0}\Psi(p)=\max_{p\geq c}\Psi(p)=\max_{p\geq c}E_{s_2(\alpha)}[\Pi]=\max_{p\geq 0}E_{s_2(\alpha)}[\Pi]$$

The last equality holds since prices smaller than the marginal cost parameter cannot be optimal for $s_2(\alpha)\in[0,1]$. Consequently, we obtain the solution under α-MEU by replacing the prior s defined in Proposition 5.1 by the prior $s_2(\alpha)$. □

Proof of Corollary 5.7. In a first step, I demonstrate that the monopolist's objective is equivalent to

$$\Psi_{\alpha,\delta}(p)=\begin{cases}E_{s_1(\alpha,\delta)}[\Pi] & \text{for}\quad 0\leq p<c\\ E_{s_2(\alpha,\delta)}[\Pi] & \text{for}\quad p\geq c\end{cases}$$

where

$$s_1(\alpha,\delta)=(1-\delta)\tilde{s}+s_1(\alpha)\delta\quad\text{and}\quad s_2(\alpha,\delta)=(1-\delta)\tilde{s}+s_2(\alpha)\delta$$

and

$$s_1(\alpha)=\alpha\overline{s}+(1-\alpha)\underline{s}\quad\text{and}\quad s_2(\alpha)=\alpha\underline{s}+(1-\alpha)\overline{s}.$$

By making use of Corollary 5.5, we can rewrite the monopolist's objective as

$$\Psi_{\alpha,\delta}(p)=\begin{cases}(1-\delta)E_{\tilde{s}}[\Pi]+\delta E_{s_1(\alpha)}[\Pi] & \text{for}\quad 0\leq p<c\\ (1-\delta)E_{\tilde{s}}[\Pi]+\delta E_{s_2(\alpha)}[\Pi] & \text{for}\quad p\geq c.\end{cases}$$

The objective can be further simplified. It is

$$\begin{aligned}\Psi_{\alpha,\delta}(p)=(p-c)[&((1-\delta)\tilde{s}+s_1(\alpha)\delta)D_1(p)\\ &+((1-\delta)(1-\tilde{s})+\delta(1-s_1(\alpha)))D_2(p)].\end{aligned}$$

for $0\leq p<c$. Defining

$$s_1(\alpha,\delta)=(1-\delta)\tilde{s}+s_1(\alpha)\delta,$$

we obtain

$$\begin{aligned}\Psi_{\alpha,\delta}(p) &= (p-c)[s_1(\alpha,\delta)D_1(p) + (1-s_1(\alpha,\delta))D_2(p)]\\ &= E_{s_1(\alpha,\delta)}[\Pi]\end{aligned}$$

for $0 \leq p < c$. By a similar line of arguments, we can define

$$s_2(\alpha,\delta) = (1-\delta)\tilde{s} + s_2(\alpha)\delta$$

and obtain

$$\begin{aligned}\Psi_{\alpha,\delta}(p) &= (p-c)[s_2(\alpha,\delta)\ D_1(p) + (1-s_2(\alpha,\delta))\ D_2(p)]\\ &= E_{s_2(\alpha,\delta)}[\Pi]\end{aligned}$$

for $p \geq c$. This demonstrates that

$$\Psi_{\alpha,\delta}(p) = \begin{cases} E_{s_1(\alpha,\delta)}[\Pi] & \text{for} \quad 0 \leq p < c \\ E_{s_2(\alpha,\delta)}[\Pi] & \text{for} \quad p \geq c. \end{cases}$$

The monopolist can secure a profit of zero by setting a price $p \geq p_1$. Consequently, a price below marginal costs cannot be a global maximizer, since it would induce strictly negative profits. Thus, the optimization problem can be rewritten in the following way:

$$\max_{p\geq 0} \Psi_{\alpha,\delta}(p) = \max_{p\geq c} \Psi_{\alpha,\delta}(p) = \max_{p\geq c} E_{s_2(\alpha,\delta)}[\Pi]$$

Since the expectation $E_{s_2(\alpha,\delta)}[\Pi]$ is negative for prices smaller than marginal costs, it follows that the global maximizer remains unchanged when the domain of the optimization problem is increased to all prices $p \geq 0$. Thus,

$$\max_{p\geq 0} \Psi_{\alpha,\delta}(p) = \max_{p\geq c} E_{s_2(\alpha,\delta)}[\Pi] = \max_{p\geq c} E_{s_2(\alpha,\delta)}[\Pi] = \max_{p\geq 0} E_{s_2(\alpha,\delta)}[\Pi].$$

Since $0 \leq s_2(\alpha, \delta) \leq 1$, we can proceed with the analysis of the risk case by replacing the prior s with $s_2(\alpha, \delta)$. □

Proof of Corollary 5.8. The derivative of $p^*_{\alpha,\delta}$ with respect to α is given by

$$\frac{\partial p^*_{\alpha,\delta}}{\partial \alpha} = \frac{a_L\,\delta - a_H\,\delta}{2\,b} < 0.$$

The derivative of $p^*_{\alpha,\delta}$ with respect to δ is

$$\begin{aligned}\frac{\partial p^*_{\alpha,\delta}}{\partial \delta} &= \frac{a_L\,(\tilde{s} + \alpha - 1) + a_H\,(-\tilde{s} - \alpha + 1)}{2\,b} \\ &= \frac{(a_H - a_L)\,(-\tilde{s} - \alpha + 1)}{2\,b}.\end{aligned}$$

Hence, the sign of the derivative depends on the sign of $-\tilde{s} - \alpha + 1$. Define $\hat{\alpha} := 1 - \tilde{s}$. Then, $\frac{\partial p^*_{\alpha,\delta}}{\partial \delta} > 0$ for $\alpha < \hat{\alpha}$, $\frac{\partial p^*_{\alpha,\delta}}{\partial \delta} = 0$ for $\alpha = \hat{\alpha}$, and $\frac{\partial p^*_{\alpha,\delta}}{\partial \delta} < 0$ for $\alpha > \hat{\alpha}$. □

Proof of Proposition 5.2. Note first that the objective reduces to the objective of the MMEU case for $\gamma = 0$. Hence, $s_{min} = s_{worst}$ for $\gamma = 0$. Let $\gamma > 0$ for the rest of the analysis. By using the definition of the Kullback-Leibler divergence, we can rewrite the objective in the following way:

$$\begin{aligned}E_s[\Pi] + \gamma R(s, s^*) = &(p - c)(sD_1(p) + (1 - s)D_2(p)) \\ &+ \gamma \left\{ s \log\left(\frac{s}{s^*}\right) + (1 - s)\log\left(\frac{1 - s}{1 - s^*}\right)\right\}\end{aligned} \tag{5.16}$$

The objective function is continuous in s, since sums and products of continuous functions are continuous. Furthermore, the constraint set $\mathcal{P}$ is compact. Using Weierstrass' theorem, we can conclude that a minimum prior s_{min} exists. The objective function is piecewise continuously differentiable with kinks at $p_0 := \frac{a_L}{b}$ and $p_1 := \frac{a_H}{b}$. Hence, it is necessary to differentiate between the cases $0 \leq p \leq p_0$, $p_0 \leq p \leq p_1$, and $p \geq p_1$.

Case 1: $p \geq p_1$

In cases where $p \geq p_1$, the expected profit part of the objective equals zero for all $s \in [0, 1]$. Consequently, the objective is given by $\gamma R(s, s^*)$. Due to Gibb's inequality, the Kullback-Leibler divergence takes the value zero if and only if $s = s^*$. Thus, we can conclude that $s_{min} = s^*$ is the minimizing prior. The same holds if $p = c$, since the objective reduces to $\gamma R(s, s^*)$.

Case 2: $0 \leq p \leq p_0$

In the second case, the objective is twice continuously differentiable in s for given $0 \leq p \leq \frac{a_L}{b}$. Taking the second-order derivative with respect to s yields

$$\frac{\partial^2 (E_s[\Pi] + \gamma R(s, s^*))}{\partial s^2} = \gamma \left(\frac{1}{1-s} + \frac{1}{s} \right) > 0.$$

Hence, the objective is strictly convex in s if $\gamma > 0$ and $s \in (0, 1)$. Two major cases can occur:

(a) There is a unique interior minimizing prior s_{min}.

(b) The minimizing prior is located at the boundary of $[\underline{s}, \overline{s}]$.

In Case (a), there is a closed form solution for s_{min}. This can be seen by looking at the first order condition. It is

$$\begin{aligned} \frac{\partial (E_s[\Pi] + \gamma R(s, s^*))}{\partial s} &= (p - c)(a_H - a_L) + \gamma \frac{\partial R(s, s^*)}{\partial s} \\ &= (p - c)(a_H - a_L) + \gamma \left(\log \left(\frac{s}{s^*} \right) + \log \left(\frac{1 - s^*}{1 - s} \right) \right). \end{aligned}$$

Solving the first order condition for s, we obtain

$$s_{min} = \frac{s^* f_1}{s^* f_1 + (1 - s^*) f_2} \tag{5.17}$$

where

$$f_1 = e^{\frac{a_H c + a_L p}{\gamma}} \quad \text{and} \quad f_2 = e^{\frac{a_L c + a_H p}{\gamma}}.$$

Hence, a unique interior solution exists. Case (b) can only occur if the interior solution is not located in the interval $[\underline{s}, \overline{s}]$. It is

$$\begin{aligned} \frac{s^* f_1}{s^* f_1 + (1-s^*) f_2} < s^* \quad &\Leftrightarrow \\ s^* f_1 < s^*(s^* f_1 + (1-s^*) f_2) \quad &\Leftrightarrow \\ f_1 < s^* f_1 + (1-s^*) f_2 \quad &\Leftrightarrow \\ f_1 < f_2. \quad & \end{aligned} \tag{5.18}$$

Plugging the definitions of f_1 and f_2 into the last inequality, we obtain

$$\begin{aligned} e^{\frac{a_H c + a_L p}{\gamma}} < e^{\frac{a_L c + a_H p}{\gamma}} \quad &\Leftrightarrow \\ a_H c + a_L p < a_L c + a_H p \quad &\Leftrightarrow \\ (a_H - a_L) c < (a_H - a_L) p \quad &\Leftrightarrow \\ p > c. \quad & \end{aligned}$$

In a similar way, we can conclude that $s_{min} > s^*$ for $p < c$, and $s_{min} = s^*$ for $p = c$. Hence, we know that $s_{min} \in (s^*, 1]$ for $p < c$, $s_{min} = s^*$ for $p = c$, and $s_{min} \in [0, s^*)$ for $p > c$. In order to make s_{min} a valid solution, it is necessary to verify under which conditions the following statements hold:

(c) $s_{min} < \overline{s}$ for $p < c$

(d) $s_{min} > \underline{s}$ for $p > c$

Statement (c):

It is

$$\begin{aligned} s_{min} < \overline{s} \quad &\Leftrightarrow \\ \frac{s^* f_1}{s^* f_1 + (1-s^*) f_2} < \overline{s} \quad &\Leftrightarrow \\ s^* f_1 < \overline{s}(s^* f_1 + (1-s^*) f_2) \quad &\Leftrightarrow \\ s^* f_1 (1-\overline{s}) < \overline{s}(1-s^*) f_2 \quad &\Leftrightarrow \\ \frac{f_1}{f_2} < \frac{\overline{s}(1-s^*)}{s^*(1-\overline{s})} := h_1. \end{aligned}$$

Plugging the definitions of f_1 and f_2 into the last inequality, we obtain

$$\begin{aligned} e^{\frac{a_H c + a_L p - a_L c - a_H p}{\gamma}} < h_1 \quad &\Leftrightarrow \\ e^{\frac{(a_H - a_L)(c-p)}{\gamma}} < h_1 \quad &\Leftrightarrow \\ \frac{(a_H - a_L)(c-p)}{\gamma} < \log(h_1). \end{aligned} \tag{5.19}$$

It is a well-known fact that $\log(h_1) > 0$ iff $h_1 > 1$.

$$\begin{aligned} \log(h_1) > 0 \quad &\Leftrightarrow \\ \frac{\overline{s}(1-s^*)}{s^*(1-\overline{s})} > 1 &\Leftrightarrow \\ \overline{s} - \overline{s}s^* > s^* - s^*\overline{s} &\Leftrightarrow \\ \overline{s} > s^* \end{aligned}$$

This condition is true by assumption. Solving inequality (5.19) for p, we obtain the equivalence

$$s_{min} < \overline{s} \quad \Leftrightarrow \quad p > c - \frac{\log(h_1)\gamma}{a_H - a_L} =: \hat{p}_1.$$

As a consequence, the minimizing prior is given by $s_{min} = \overline{s}$ for $p \leq \hat{p}_1$ and by

$$s_{min} = \frac{s^* f_1}{s^* f_1 + (1-s^*) f_2}$$

for $p \geq \hat{p}_1$. This case can be summarized as follows:

$$s_{min} = \begin{cases} \overline{s} & \text{for} \quad 0 \leq p \leq c \wedge p \leq \hat{p}_1 \\ \frac{s^* f_1}{s^* f_1 + (1-s^*) f_2} & \text{for} \quad 0 \leq p \leq c \wedge p \geq \hat{p}_1 \end{cases}$$

This expression can be simplified. It is

$$s_{min} = \begin{cases} \overline{s} & \text{for} \quad p \in [0, \min\{c, \hat{p}_1\}] \\ \frac{s^* f_1}{s^* f_1 + (1-s^*) f_2} & \text{for} \quad p \in [\max\{0, \hat{p}_1\}, c]. \end{cases}$$

Statement (d):

Similar to the previous case, we can show that the condition $s_{min} > \underline{s}$ is equivalent to the condition

$$\frac{f_1}{f_2} > \frac{\underline{s}(1 - s^*)}{s^*(1 - \underline{s})} := h_2. \tag{5.20}$$

Plugging f_1 and f_2 into inequality 5.20, we obtain that the initial condition is equivalent to

$$\frac{(a_H - a_L)(c - p)}{\gamma} > \log(h_2).$$

Moreover,

$$\begin{aligned} h_2 = \frac{\underline{s}(1 - s^*)}{s^*(1 - \underline{s})} < 1 \quad &\Leftrightarrow \\ \underline{s} - \underline{s}s^* > s^* - \underline{s}s^* \quad &\Leftrightarrow \\ s^* > \underline{s}, \end{aligned} \tag{5.21}$$

which is an assumption of the case under consideration. Solving inequality 5.20 for p, we obtain the equivalence

$$s_{min} > \underline{s} \quad \Leftrightarrow \quad p < c - \frac{\log(h_2)\gamma}{a_H - a_L} =: \hat{p}_2.$$

Hence, the minimizing prior is given by $s_{min} = \underline{s}$ for $p \geq \hat{p}_2$ and by

$$s_{min} = \frac{s^* f_1}{s^* f_1 + (1 - s^*) f_2}$$

for $p < \hat{p}_2$.

This case can be summarized as follows:

$$s_{min} = \begin{cases} \underline{s} & \text{for} \quad p \geq c \wedge p \geq \hat{p}_2 \wedge p \leq \frac{a_L}{b} \\ \frac{s^* f_1}{s^* f_1 + (1-s^*) f_2} & \text{for} \quad p \geq c \wedge p \leq \hat{p}_2 \wedge p \leq \frac{a_L}{b} \end{cases}$$

The last expression can be simplified. It is

$$s_{min} = \begin{cases} \underline{s} & \text{for} \quad p \in [\max\{c, \hat{p}_2\}, \frac{a_L}{b}] \\ \frac{s^* f_1}{s^* f_1 + (1-s^*) f_2} & \text{for} \quad p \in [c, \min\{\frac{a_L}{b}, \hat{p}_2\}]. \end{cases}$$

Case 3: $p_0 \leq p \leq p_1$

The objective function reduces to

$$\Pi^{red} := s(p - c) D_1(p) + \gamma \left\{ s \log \left(\frac{s}{s^*} \right) + (1 - s) \log \left(\frac{1 - s}{1 - s^*} \right) \right\}. \tag{5.22}$$

The derivative of Π^{red} with respect to s is given by

$$(p - c)(a_H - bp) + \gamma \left(\log \left(\frac{s}{s^*} \right) + \log \left(\frac{1 - s^*}{1 - s} \right) \right).$$

Solving the first order condition, we obtain

$$s_{cand} = \frac{s^* f_3}{s^* f_3 + (1 - s^*) f_4}$$

with

$$f_3 = e^{\frac{a_H c + bp^2}{\gamma}} \quad \text{and} \quad f_4 = e^{\frac{(a_H + bc)p}{\gamma}}$$

as a candidate for the solution of the minimization problem. As in the previous case, the second-order derivative with respect to s is given by

$$\frac{\partial^2 \Pi^{red}}{\partial s^2} = \gamma \left(\frac{1}{1-s} + \frac{1}{s} \right) > 0.$$

Consequently, the objective is strictly convex. The following cases can occur:

(e) The candidate prior s_{cand} is the unique interior minimizer.

(f) The minimizer is either $\underline{s}$ or $\overline{s}$.

In the following, I determine the conditions under which either (e) or (f) holds. The interior prior is not the minimizer if either $s_{cand} > \overline{s}$ or $s_{cand} < \underline{s}$. Note that the condition $s_{cand} < s^*$ is equivalent to the condition $f_3 < f_4$. This can be demonstrated by replacing f_1 by f_3 and f_2 by f_4 in the proof of Statement (d). Hence,

$$\begin{aligned} e^{\frac{a_H c + bp^2}{\gamma}} < e^{\frac{(a_H + bc)p}{\gamma}} \quad &\Leftrightarrow \\ a_H c + bp^2 < a_H p + bcp \quad &\Leftrightarrow \\ bp(p-c) < a_H(p-c) \quad &\Leftrightarrow \\ p < \frac{a_H}{b}, \end{aligned}$$

which is true by assumption. As a consequence, the case $s_{cand} > \overline{s}$ can be excluded, since $s^* \in [\underline{s}, \overline{s}]$ by assumption. The condition $s_{cand} < \underline{s}$ is equivalent to the condition

$$\frac{f_3}{f_4} < h_2.$$

Then,

$$\begin{aligned} s_{cand} < \underline{s} \quad &\Leftrightarrow \\ e^{\frac{a_H c + bp^2 - (a_H + bc)p}{\gamma}} < h_2 \quad &\Leftrightarrow \\ bp^2 - (a_H + bc)p + a_H c - \gamma \log(h_2) < 0. \end{aligned} \tag{5.23}$$

The quadratic equation on the left-hand side has the solution

$$\hat{p}_{3/4} = \frac{a_H + bc}{2b} \mp \sqrt{\frac{(a_H + bc)^2}{4b^2} - a_H c + \gamma \log(h_2)}.$$

Since the function on the left-hand side of the last inequality is a parabola opening upwards, we can infer that the inequality holds for $\hat{p}_3 < p < \hat{p}_4$. Taking all the conditions of this case together, we obtain that s_{cand} is the minimizing prior if one of the following sets of requirements is satisfied:

(h) $p \geq \frac{a_L}{b}$, $p \leq \frac{a_H}{b}$, $p \leq \hat{p}_3$

(i) $p \geq \frac{a_L}{b}$, $p \leq \frac{a_H}{b}$, $p \geq \hat{p}_4$

The conditions in (h) are equivalent to

$$p \in \left[\frac{a_L}{b}, \min\left\{\hat{p}_3, \frac{a_H}{b}\right\}\right].$$

The conditions in (i) are equivalent to

$$p \in \left[\max\left\{\frac{a_L}{b}, \hat{p}_4\right\}, \frac{a_H}{b}\right].$$

The condition $s_{cand} < \underline{s}$ holds if

(j) $p \geq \frac{a_L}{b}$, $p \leq \frac{a_H}{b}$, $p \geq \hat{p}_3$, $p \leq \hat{p}_4$.

The conditions in (j) are equivalent to

$$p \in \left[\max\left\{\frac{a_L}{b}, \hat{p}_3\right\}, \min\left\{\hat{p}_4, \frac{a_H}{b}\right\}\right].$$

A solution of the quadratic equation exists as long as

$$\frac{(a_H + bc)^2}{4b^2} - a_H c + \gamma \log(h_2) \geq 0.$$

Solving the inequality for γ, we obtain

$$\gamma \leq \frac{1}{\log(h_2)} \left[a_H c - \frac{(a_H + bc)^2}{4b^2} \right] := \hat{\gamma}.$$

Thus, $s_{cand} > \underline{s}$ for all $\gamma > \hat{\gamma}$. □

Proof of Corollary 5.10. It is

$$\lim_{\gamma \to \infty} f_1(\gamma) = \lim_{\gamma \to \infty} f_2(\gamma) = \lim_{\gamma \to \infty} f_3(\gamma) = \lim_{\gamma \to \infty} f_4(\gamma) = 1.$$

Consequently,

$$\lim_{\gamma \to \infty} \frac{s^* f_1}{s^* f_1 + (1 - s^*) f_2} = \lim_{\gamma \to \infty} \frac{s^* f_3}{s^* f_3 + (1 - s^*) f_4} = s^*.$$

What remains to be shown is that the corner solutions $\overline{s}$ and $\underline{s}$ vanish in the limit. It is

- $\lim_{\gamma \to \infty} \hat{p}_1 = -\infty$
- $\lim_{\gamma \to \infty} \hat{p}_2 = \infty$

Besides, the limits $\lim_{\gamma \to \infty} \hat{p}_3$ and $\lim_{\gamma \to \infty} \hat{p}_4$ do not exist, since both zeros $\hat{p}_i$ for $i = 3, 4$ are only defined for $\gamma \leq \hat{\gamma} < \infty$. As a consequence,

$$\lim_{\gamma \to \infty} A_4(\gamma) = \lim_{\gamma \to \infty} A_5(\gamma) = \lim_{\gamma \to \infty} A_6(\gamma) = \emptyset.$$

Moreover,

$$\lim_{\gamma \to \infty} A_1(\gamma) = \lim_{\gamma \to \infty} A_3(\gamma) = \emptyset$$

because of the limiting properties of $\hat{p}_1$ and $\hat{p}_2$ and

$$\lim_{\gamma \to \infty} A_2(\gamma) = \left[0, \frac{a_L}{b} \right]$$

for the same reason. Hence,

$$\lim_{\gamma\to\infty} s_{min}(\gamma) = \begin{cases} \lim_{\gamma\to\infty} \frac{s^* f_1}{s^* f_1+(1-s^*)f_2} & \text{for} \quad p \in [0, \frac{a_L}{b}] \\ \lim_{\gamma\to\infty} \frac{s^* f_3}{s^* f_3+(1-s^*)f_4} & \text{for} \quad p \in [\frac{a_L}{b}, \frac{a_H}{b}] \\ s^* & \text{for} \quad p \in [\frac{a_H}{b}, \infty) \end{cases}$$

which proves the claim due to the fact that the limit of the component functions is given by s^* as well. The property

$$\lim_{\gamma\to\infty} p(\gamma) = p^{risk}(s^*)$$

is an immediate consequence of the limit of the minimizing prior. What remains to be demonstrated is that $p(\gamma)$ converges against $p^{risk}(s^*)$ from below for $c \leq p(\gamma) \leq \frac{a_L}{b}$ and $\gamma > \hat{\gamma}$. The minimum prior function $s_{min}(\gamma)$ is piecewise continuously differentiable. In cases where $s_{min}(\gamma)$ equals s^*, we can conclude that the minimum prior is independent of γ and therefore constant. Subsequently, I investigate how the interior solution

$$h = \frac{s^* f_1}{s^* f_1 + (1 - s^*) f_2}$$

reacts to an increase in γ. The derivative of h with respect to γ is given by

$$\frac{\partial h}{\partial \gamma} = \frac{s^*(1-s^*) f_1 f_2 (p-c)(a_H - a_L)}{\gamma^2 (s^* f_1 + (1-s^*) f_2)^2}. \tag{5.24}$$

The sign of (5.24) depends on the sign of $p - c$ only. Hence, h is negative for $p < c$, zero for $p = c$, and positive for $p > c$. This establishes that h is strictly increasing for $p > c$. Consequently, we can infer that $s_{min}(\gamma)$ converges to s^* from below. Let $0 < \hat{\gamma} < \gamma_1 < \gamma_2 < \infty$. Then, $s_{min}(\gamma_1) \leq s_{min}(\gamma_2)$ and $E_{s_{min}(\gamma_1)}[\Pi] \leq E_{s_{min}(\gamma_2)}[\Pi]$. Taking the derivative of h with respect to p yields

$$\frac{\partial h}{\partial p} = -\frac{(a_H - a_L) e^{\frac{(a_H + a_L)(c+p)}{\gamma}} (1 - s^*) s^*}{\gamma (s^* f_1 + (1 - s^*) f_2)^2}.$$

Obviously, this derivative is negative. Consequently, $s_{min}(p)$ moves further away from s^* as p increases. Hence, $R(s_{min(p)}, s^*)$ is increasing in p. Thus, the monopolist always has an incentive to increase prices if the objective was given by the Kullback-Leibler divergence. What remains to be analyzed is the expected profit part $E_s[\Pi]$ of the objective. As the minimum prior s_{min} increases with γ, the monopolist expects the high number of consumers scenario to be more likely. Therefore, he or she, has an incentive to increase prices, see for instance Example 5.2. Since both effects go the same way, the monopolist has an overall incentive to raise prices. □

Proof of Proposition 5.3. In a first step, I demonstrate that the objective function

$$g = qE_{\underline{s}}[\Pi] + (1-q)E_{\overline{s}}[\Pi] \tag{5.25}$$

yields a lower monopoly price than

$$q\Phi(E_{\underline{s}}[\Pi]) + (1-q)\Phi(E_{\overline{s}}[\Pi]), \tag{5.26}$$

if Φ is a concave distortion function. The objectives (5.25) and (5.26) are only piecewise differentiable with kinks at $p_0 = \frac{a_L}{b}$ and $p_1 = \frac{a_H}{b}$. Therefore, we derive the set of global maximizers on a case-by-case basis.

Case 1: $\mathbf{p \geq p_1}$

We define $D_1 = \{p : p \geq p_1\}$ and obtain

$$\begin{aligned}
&\arg\max_{p\in D_1}\Big\{q\Phi(E_{\underline{s}}[\Pi]) + (1-q)\Phi(E_{\overline{s}}[\Pi])\Big\}\\
&= \arg\max_{p\in D_1}\Big\{q\Phi(0) + (1-q)\Phi(0)\Big\}\\
&= \arg\max_{p\in D_1}\Phi(0) = D_1.
\end{aligned}$$

Moreover, we can conclude that

$$\begin{aligned}
&\arg\max_{p\in D_1} qE_{\underline{s}}[\Pi] + (1-q)E_{\overline{s}}[\Pi] \\
&= \arg\max_{p\in D_1} 0 = D_1.
\end{aligned}$$

Hence, the set of local maximizers is the same for (5.25) and (5.26).

Case 2: $\mathbf{p_0 \leq p \leq p_1}$

We define $D_2 = [p_0, p_1]$. This case needs to be subdivided into three subcases.

(a) The maximum of g is located at $p = p_1$. (corner solution 1)

(b) The maximum of g is located at $p = p_0$. (corner solution 2)

(c) The maximum of g is located in the interior of D_2. (interior solution)

In Case 2, the first objective (5.25) reads

$$\begin{aligned}
qE_{\underline{s}}[\Pi] + (1-q)E_{\overline{s}}[\Pi] &= q\underline{s}(p-c)(a_H - bp) + \overline{s}(1-q)(p-c)(a_H - bp) \\
&= (p-c)(a_H - bp)(q\underline{s} + \overline{s}(1-q)).
\end{aligned}$$

Case (2a):

Since g is quadratic in p, we can conclude that $qE_{\underline{s}}[\Pi] + (1-q)E_{\overline{s}}[\Pi]$ is strictly increasing on D_2. Moreover, if $qE_{\underline{s}}[\Pi] + (1-q)E_{\overline{s}}[\Pi]$ is strictly increasing in p, we can infer that both $\underline{s}(p-c)[a_H - bp]$ and $\overline{s}(p-c)[a_H - bp]$ are strictly increasing in p. This is because multiplications with positive constants leave a function's monotonicity properties unaffected. Since Φ is a strictly increasing transformation, it follows that $\Phi(E_{\underline{s}})$ and $\Phi(E_{\overline{s}})$ are both strictly increasing in p. This implies that objective (5.26) is also strictly increasing in p. Consequently, the local maximizer equals p_1 for both functions and is therefore not affected by the transformation Φ.

Case (2b) and Case (2c):

With the same arguments as in Case (2a), we can establish that both objectives have the same monotonicity properties, and therefore the same maximizer.

Case 3: $\mathbf{p < c}$

We define $D_3 = \{p : 0 \leq p < c\}$. This case can be excluded. A price $p \in D_3$ cannot be a global maximizer of objective (5.25), see Example 5.2. Similarly, p cannot be a global maximizer of objective (5.26). This is because p cannot be a maximizer of $E_{\underline{s}}[\Pi]$. Similarly, p cannot be a maximizer of $E_{\overline{s}}[\Pi]$ as well. This implies that $p \in D_3$ is not a maximizer of $\Phi(E_{\underline{s}}[\Pi])$ and $\Phi(E_{\overline{s}}[\Pi])$. Hence, p is not a maximizer of objective (5.26).

Case 4: $\mathbf{c \leq p \leq p_0}$

We define $D_4 = \{p : c \leq p \leq \frac{a_L}{b}\}$. This case needs to be subdivided into three subcases.

(a) The maximum of g is located in the interior of D_3. (interior solution)

(b) The maximum of g is located at the lower boundary $p = c$. (corner solution 1)

(c) The maximum of g is located at the upper boundary $p = p_0$. (corner solution 2)

Case (4a):

The local maximizers of (5.25) and (5.26) are defined by

$$p_1^{KMM} := \arg\max_{p \in D_4} \left\{ qE_{\underline{s}}[\Pi] + (1-q)E_{\overline{s}}[\Pi] \right\}$$

$$p_2^{KMM} := \arg\max_{p \in D_4} \left\{ q\Phi(E_{\underline{s}}[\Pi]) + (1-q)\Phi(E_{\overline{s}}[\Pi]) \right\}.$$

Since g is quadratic in p, we can infer that objective (5.25) is strictly increasing for $p_0 \leq p < p_1^{KMM}$ and decreasing for $p_1^{KMM} < p < p_1$. As a next step, I investigate the monotonicity properties of $E_{\underline{s}}[\Pi]$ and $E_{\overline{s}}[\Pi]$. Therefore, I denote with $p_{\underline{s}}$ the maximizer of $E_{\underline{s}}[\Pi]$, and I denote with $p_{\overline{s}}$ the maximizer of $E_{\overline{s}}[\Pi]$ on D_4. From Proposition 5.1, we can conclude that $p_{\underline{s}} < p_1^{KMM} < p_{\overline{s}}$. Table 5.2 summarizes the monotonicity properties of objective (5.25) on the following partition of D_4:

1. $I_1 := \{p : p_0 \leq p \leq p_{\underline{s}}\}$,
2. $I_2 := \{p : p_{\underline{s}} \leq p \leq p_1^{KMM}\}$,
3. $I_3 := \{p : p_1^{KMM} \leq p \leq p_{\overline{s}}\}$,
4. $I_4 := \{p : p_{\overline{s}} \leq p \leq p_1\}$.

Interval	**Objective 1**	$\mathbf{E_{\underline{s}}[\Pi]}$	$\mathbf{E_{\overline{s}}[\Pi]}$
I_1	$\uparrow$	$\uparrow$	$\uparrow$
I_2	$\uparrow$	$\downarrow$	$\uparrow$
I_3	$\downarrow$	$\downarrow$	$\uparrow$
I_4	$\downarrow$	$\downarrow$	$\downarrow$

TABLE 5.2: Monotonicity Properties

A direct consequence of Table 5.2 is that p_2^{KMM} is contained in the union $I_2 \cup I_3$, since objective (5.26) is strictly increasing on I_1, and strictly decreasing on I_4. This is due to the monotonicity properties of $E_{\underline{s}}[\Pi]$ and $E_{\overline{s}}[\Pi]$. What remains to be shown is that $p_2^{KMM} \notin I_3 \backslash \{p_1^{KMM}\}$. In order to prove this statement, we contemplate the partial derivative of objective (5.26) with respect to p:

$$q\Phi'(E_{\underline{s}}[\Pi])E_{\underline{s}}[\Pi]' + (1-q)\Phi'(E_{\overline{s}})E_{\overline{s}}[\Pi]'$$

Assuming $p \in I_3$, we can infer that $E_{\underline{s}}[\Pi]' < 0$ and $E_{\overline{s}}[\Pi]' > 0$. Moreover, $E_{\underline{s}}[\Pi] \leq E_{\overline{s}}[\Pi]$ for $p \geq c$. Furthermore, since $\Phi'' < 0$, we can deduce that

$$\Phi'(E_{\underline{s}}[\Pi]) \geq \Phi'(E_{\overline{s}}[\Pi]).$$

Consequently,

$$q\Phi'(E_{\underline{s}}[\Pi]) \cdot E_{\underline{s}}[\Pi]' + (1-q)\Phi'(E_{\overline{s}}[\Pi]) \cdot E_{\overline{s}}[\Pi]' \leq \Phi'(E_{\overline{s}}[\Pi]) \left(qE_{\underline{s}}[\Pi]' + (1-q)E_{\overline{s}}[\Pi]'\right).$$

Besides, we can infer that

$$qE_{\underline{s}}[\Pi]' + (1-q)E_{\overline{s}}[\Pi]' = \underline{s}q(a_H - a_L) + \overline{s}(a_H - a_L)(1-q) + a_L - 2bp + bc. \tag{5.27}$$

Expression (5.27) is negative, if and only if,

$$p > \frac{a_H(q\underline{s} + (1-q)\overline{s}) + a_L(1 - q\underline{s} + (1-q)\overline{s})}{2b} = p_1^{KMM}. \tag{5.28}$$

Hence, the p-derivative of objective (5.26) is negative for $p \in I_3$. As a result, the optimum is located in I_2 and $p_1^{KMM} > p_2^{KMM}$.

Case (4b):

In cases where p_1^{KMM} equals the marginal cost parameter c, we can conclude that objective (5.25) is strictly decreasing on D_4. Due to inequality (5.28), objective (5.26) is also strictly decreasing on D_4. Hence, $p_2^{KMM} = c$ is the unique local maximizer of (5.26) on D_4. Consequently, both objectives have the same local maximizer, and this maximizer is unique.

Case (4c):

In cases where p_1^{KMM} equals the upper boundary p_0 of the interval D_3, we can conclude that objective (5.25) is strictly increasing on D_4. Due to inequality (5.28), objective (5.26) is also strictly increasing on D_4. Hence, $p_2^{KMM} = p_0$ is the unique local maximizer of (5.26) on D_4. Consequently, both objectives have the same local maximizer, and this maximizer is unique.

Note that $p_1^{KMM} = p_2^{KMM}$ in all cases, except in Case $(4a)$. The fact that p_1^{KMM} equals p_2^{KMM} in these instances is independent of the curvature of Φ. Hence, even if Φ is assumed to be convex, we can conclude that $p_1^{KMM} = p_2^{KMM}$ for all cases, except in Case $(4a)$. As a result, it is sufficient to reexamine Case $(4a)$ under the assumption that Φ is strictly convex. Remember, that objective (5.26) is strictly increasing on I_1 and strictly decreasing on I_4. This excludes the possibility that p_2^{KMM} is an element of I_1 or I_4. What remains to be shown is that p_2^{KMM} is not an element of $I_2 \backslash \{p_1^{KMM}\}$ either. This is demonstrated by means of a proof by contradiction. Assume that p_2^{KMM} is an element of $I_2 \backslash \{p_1^{KMM}\}$. Then, $E_{\underline{s}}[\Pi]' < 0$ and $E_{\overline{s}}[\Pi]' > 0$. Furthermore, since $\Phi'' > 0$, we can deduce that

$$\Phi'(E_{\underline{s}}[\Pi]) \leq \Phi'(E_{\overline{s}}[\Pi]).$$

Consequently, we obtain the following estimate for the p-derivative of objective (5.26):

$$q\Phi'(E_{\underline{s}}[\Pi]) \cdot E_{\underline{s}}[\Pi]' + (1-q)\Phi'(E_{\overline{s}}[\Pi]) \cdot E_{\overline{s}}[\Pi]' \geq \Phi'(E_{\overline{s}}[\Pi]) \left(qE_{\underline{s}}[\Pi]' + (1-q)E_{\overline{s}}[\Pi]'\right).$$

From Case $(4a)$, we know that $qE_{\underline{s}}[\Pi]' + (1-q)E_{\overline{s}}[\Pi]'$ is positive for $p < p_1^{KMM}$ and negative for $p > p_1^{KMM}$. Consequently, the p-derivative of objective (5.26) is positive on $I_2 \backslash \{p_1^{KMM}\}$. Hence, p_2^{KMM} cannot be located in the interval I_2. This proves the claim. $\square$

Chapter 6

Conclusion

This thesis applies the concept of ambiguity to different fields of research in economics with a focus on Industrial Organization and Health Economics. In the first chapter, I provide a basic overview on the development of decision-theoretic models under uncertainty.

The second chapter focuses on a Hotelling location-then-price duopoly game under demand ambiguity. Using a Choquet model with neo-additive capacities, this chapter provides a unifying framework for the Hotelling model under risk developed by Meagher and Zauner [2004], and the Hotelling model under ambiguity with α-MEU preferences developed by Król [2012]. It turns out that there is a unique subgame-perfect pure strategy Nash equilibrium for firms' location choices in this general framework. Moreover, this equilibrium features interesting comparative static results with respect to the confidence and optimism parameter of the underlying capacity. One obtains the result that a higher degree of pessimism decreases equilibrium differentiation. A higher degree of confidence decreases equilibrium differentiation if firms are rather pessimistic and increases equilibrium differentiation if firms are sufficiently optimistic. For an intermediate optimism value $\hat{\alpha}$ equilibrium differentiation is independent of δ. More important than these comparative static results is that the neo-additive approach provides an additional source of explanation for a variety of observed product design choices. In this sense, we reinterpret the real-world examples provided by Król [2012] within the neo-additive framework. An important example relates to the mutual funds market. The observation here is that fund managers tend to differentiate their products less after the financial crisis. The explanation put forward by Król [2012] is that the ambiguity attitude parameter α has changed due to the financial crisis. Hence, fund managers have become increasingly pessimistic after the financial crisis. In our view, this conclusion is problematic, since it is not clear whether managers became more pessimistic (change in α), or whether they perceived the market environment to be less reliable (change in δ). If fund managers are rather ambiguity-averse, one can conclude with the neo-additive approach that decreasing confidence lowers product differentiation.

The third chapter of this thesis considers ambiguity in the context of Health Economics, and more specifically in the context of primary prevention. The underlying research question of this project is to examine how patients adjust preventive activities in the light of new information when the relationship between effort and disease probabilities is characterized by Knightian uncertainty. Information is modeled by a random signal. After receiving the signal, patients update their prior beliefs and select an optimal level of effort.[1]

In a first step, I present the primitives of the model and proceed by analyzing the underlying optimization problem. In this context, I specify conditions for the existence and uniqueness of interior and corner solutions. Subsequently, I conduct a comparative static analysis with respect to the optimism parameter α and the confidence parameter δ. It turns out that the effect of optimism on prevention is determined by two concurrent effects, which are denoted as "perceived efficacy effect" and "expected marginal utility effect". The perceived efficacy effect captures the fact that optimists and pessimists might differ in their assessment of the preventive regime's capability to reduce the underlying probability of disease. The expected marginal utility effect takes into account that a shift in the perceived disease probability might increase or decrease marginal gains or losses from additional units of prevention. The overall effect of optimism on prevention is the

[1] Effort is interpreted as level of adherence to a preventive regime.

sum of both effects and can be positive, negative or zero. A similar analysis applies to the confidence parameter δ.

Having explored the connection between optimism, confidence and prevention, I continue by looking at the relationship between prevention and information. By using the three updating rules for neo-additive capacities discussed in Eichberger et al. [2010], I derive patients' ex-post optimization problem. The following section analyzes the effect of information on prevention. Therefore, I introduce two benchmark measures for preventive behavior under Knightian uncertainty. The first benchmark is a Bayesian patient whose prior belief is represented by a unique subjective probability. An interesting finding results from the comparison of Bayesian and Non-Bayesian patients when the prior belief of the Bayesian patient corresponds to the reference prior in the Non-Bayesian case. Non-Bayesian patients exhibit a higher ex-post level of prevention, relative to the Bayesian benchmark patient, under the pessimistic updating rule when the worst case relationship π_{min} is less effective than the updated Bayesian relations $\pi_{q^{Bayes}}$. Similarly, Bayesian patients feature a lower level of ex-post prevention under the optimistic updating rule if π_{max} is less effective than $\pi_{q^{Bayes}}$. Under more restrictive requirements, one can show that the generalized Bayesian updating rule induces lower (higher) ex-post preventive activities than the Bayesian benchmark patients when patients are sufficiently optimistic (pessimistic). More interestingly, information does not necessarily close the gap in preventive activities between Bayesian and Non-Bayesian patients. On the contrary, information has the potential to render extreme patients even more extreme. The second benchmark is a Bayesian patient that is aware of the true underlying preventive relationship. In this context, I introduce the terms "excessive preventive behavior" and "preventive inertia". Patients exhibit excessive preventive behavior when their level of effort under Knightian uncertainty exceeds the optimal level of prevention under perfect information. Similarly, patients display preventive inertia when they exert less effort under Knightian uncertainty than in a situation where they know the relationship between effort and the probability of disease. One can show that excessive preventive behavior and preventive inertia vanish for the Bayesian benchmark patient when the correct preventive relationship π_θ can be perfectly inferred.[2] Surprisingly, this is not true for Knightian patients. In the general case, excessive preventive behavior and preventive inertia persist even for the Bayesian benchmark patient. Moreover, one cannot conclude that excessive behavior is attenuated even if the correct signal is observed. Clearly, if the correct signal is sent, patients update their prior belief such that the posterior probability gives a larger weight to the true underlying relationship. When the signal structure is such that this increase of the posterior is strong enough, one obtains that excessive preventive behavior or preventive inertia is reduced. In cases where the posterior of the correct relationship increases only slightly, the result depends on the posterior probabilities for the remaining relationships. For instance, in a situation where patients feature preventive inertia before observing the signal, preventive inertia can be reinforced if the posterior for those preventive relationships increases strongly enough which would "by themselves" induce preventive inertia.

The fourth chapter relates to the contribution made by Çelen [2012], who extends the well-known Blackwell's theorem to MEU-preferences. We observe that the value of information defined in Çelen [2012] entails dynamically inconsistent behavior. The reason is that it is not defined according to the principle of recursively defined utility. In order to account for this observation, we propose an alternative definition for the value of information under MEU-preferences which is, by construction, consistent with the backward induction principle.

The fifth chapter of this thesis should be understood as a guide for those interested in implementing models of decision making under ambiguity to address research problems in economics. By means of a simple baseline model, a static monopoly market with linear demand, I explain which arguments can be used to justify a modeling approach that prescribes models of decision making under ambiguity. In order to do so, I outline the philosophical discussion on probabilities to clearly define the notions of objective and subjective probabilities. In the end, the justification for ambiguity boils down to two necessary requirements. The first one is to provide a rationale why probabilities are not

[2] This is for instance the case when there is a signal s such that the conditional probability to receive this signal given the true parameter is a Dirac measure.

objectively given in the choice situation under consideration. If this claim can be underpinned with credible arguments, one knows that probabilities are either subjective or that decision-makers hold beliefs that violate the notion of subjective probabilities. The second requirement is to refute subjective probabilities. This can be done by referring to Ellsberg's paradox, see Ellsberg [1961], which has been experimentally confirmed by Camerer and Weber [1992]. Since there is a variety of decision theoretic models consistent with Ellsberg's paradox, I demonstrate the implications of the most prominent models of decision-making under ambiguity for the monopoly pricing problem. I assume a simple scenario with two states of the world and a monopolist facing demand ambiguity. The demand functions are assumed to be linear with the same slope parameter but different intercepts. Contrary to the existing literature on monopoly pricing under Knightian uncertainty, I can demonstrate that ambiguity might increase or decrease optimal prices. Besides, in special cases where corner solutions occur, ambiguity has no influence on optimal pricing. Depending on the underlying parameter constellations, three types of solutions can occur: a unique interior solution where demand is positive in both states of the world, a corner solution where demand is only positive in one of the two states, and a third scenario where demand is zero in both states of the world. It follows that extreme pessimism induces lower monopoly prices throughout all model specifications. In particular, one can conclude that a higher degree of pessimism in the α-MEU or Choquet model with neo-additive capacities yields a lower monopoly price in cases where the interior solution is optimal. [3] A comparable statement does not hold for the KMM model with a constant absolute ambiguity transformation function. One can observe that a higher degree of absolute ambiguity aversion does not necessarily translate into lower monopoly prices.

[3] In cases where corner solutions apply, the monopoly price is independent of α.

Bibliography

M. Abdellaoui, A. Baillon, L. Placido, and P. P. Wakker. The rich domain of uncertainty: Source functions and their experimental implementation. *The American Economic Review*, 101(2):695–723, 2011.

D. Alary, C. Gollier, and N. Treich. The effect of ambiguity aversion on insurance and self-protection. *The Economic Journal*, 123(573):1188–1202, 2013.

S.P. Anderson, J.K. Goeree, and R. Ramer. Location, location, location. *Journal of Economic Theory*, 77(1):102–127, 1997.

F. J. Anscombe and R. J. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

K. Arrow and L. Hurwicz. *Uncertainty and Expectations in Economics*, chapter An optimality criterion for decision-making under ignorance, pages 1–11. Basil Blackwell, Oxford, 1972.

T. Asano and A. Shibata. Optimal pricing and quality choice of a monopolist under knightian uncertainty. *International Journal of Industrial Organization*, 29(6):746–754, 2011.

R. Balvers and L. Szerb. Location in the hotelling duopoly model with demand uncertainty. *European Economic Review*, 40(7):1453–1461, 1996.

D. Bergemann and K. Schlag. Robust monopoly pricing. *Journal of Economic Theory*, 146(6):2527–2543, 2011.

L. Berger. The impact of ambiguity prudence on insurance and prevention. ECARES Working Papers, 2014.

T. F. Bewley. Knightian decision theory. part 1. *Decisions in Economics and Finance*, 25(2):79–110, 2002.

D. Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953.

G. E. P. Box and N. R. Draper. *Empirical Model-Building and Response Surfaces*, volume 424. John Wiley & Sons, 1987.

S. L. Brown and K. L. Eisenhardt. Product development: Past research, present findings, and future directions. *Academy of Management Review*, 20(2):343–378, 1995.

C. Camerer and M. Weber. Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, 5(4):325–370, 1992.

R. Carnap. *Logical Foundations of Probability.* University of Chicago Press, Chicago, second edition, 1962.

F. J. Casado-Izaga. Location decisions: the role of uncertainty about consumer tastes. *Journal of economics*, 71(1):31–46, 2000.

B. Çelen. Informativeness of experiments for meu. *Journal of Mathematical Economics*, 48(6):404–406, 2012.

S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio. Uncertainty averse preferences. *Journal of Economic Theory*, 146(4):1275–1330, 2011.

A. Chateauneuf and J. H. Faro. Ambiguity through confidence functions. *Journal of Mathematical Economics*, 45(9):535–558, 2009.

A. Chateauneuf, J. Eichberger, and S. Grant. Choice under uncertainty with the best and worst in mind: Neo-additive capacities. *Journal of Economic Theory*, 137(1):538–567, 2007.

G. Choquet. Theory of capacities. *Annales de L' Institut Fourier*, 5:131–295, 1955.

M. Cohen, J. Etner, and M. Jeleva. Dynamic decision making when risk perception depends on past experience. *Theory and Decision*, 64(2-3):173–192, 2008.

J. Crémer. A simple proof of blackwell's "comparison of experiments" theorem. *Journal of Economic Theory*, 27(2):439–443, 1982.

C. d'Aspremont, J. J. Gabszewicz, and J.-F. Thisse. On hotelling's "stability in competition". *Econometrica*, 47:1145–1150, 1979.

B. De Finetti. La prévision: ses lois logiques, ses sources subjectives. *Annales de L' Institut Henri Poincaré*, 7(1):1–68, 1937.

A. V. Deardorff. *Deardorff's glossary of international economics.* February, 24, 2008 2001.

D. Denneberg. *Non-additive measure and integral.* Kluwer Academic Publishers, Boston, 1994.

L. Eeckhoudt and C. Gollier. The impact of prudence on optimal prevention. *Economic Theory*, 26(4):989–994, 2005.

L. Eeckhoudt, P. Godfroid, and M. Marchand. Risque de santé, médicine préventive et médicine curative. *Revue d' Economie Politique*, 108, 3:321–337, 1998.

L. Eeckhoudt, C. Gollier, and G. Immordino. How diagnostic tests affect prevention: A cost-benefit analysis. *CSEF Working Papers: Université de Salerne*, 2001.

I. Ehrlich and G. S. Becker. Market insurance, self-insurance, and self-protection. *The Journal of Political Economy*, 80:623–648, 1972.

J. Eichberger and D. Kelsey. E-capacities and the ellsberg paradox. *Theory and Decision*, 46(2):107–138, 1999.

J. Eichberger, S. Grant, and D. Kelsey. Ceu preferences and dynamic consistency. *Mathematical Social Sciences*, 49(2):143–151, 2005.

J. Eichberger, D. Kelsey, and B. C. Schipper. Ambiguity and social interaction. *Oxford Economic Papers*, 61:355–379, 2009.

J. Eichberger, S. Grant, and D. Kelsey. Comparing three ways to update choquet beliefs. *Economics Letters*, 107(2):91–94, 2010.

J. Eichberger, S. Grant, D. Kelsey, and G.A. Koshevoy. The α-meu model: A comment. *Journal of Economic Theory*, 146(4):1684–1698, 2011.

D. Ellsberg. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 75:643–669, 1961.

L. G. Epstein. A definition of uncertainty aversion. *Review of Economic Studies*, 66(3): 579–608, 1999.

L. G. Epstein. A paradox for the "smooth ambiguity" model of preference. *Econometrica*, 78(6):2085–2099, 2010.

L. G. Epstein and M. Le Breton. Dynamically consistent beliefs must be bayesian. *Journal of Economic Theory*, 61(1):1–22, 1993.

L. G. Epstein and M. Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113(1):1–31, 2003.

L. G. Epstein and M. Schneider. Learning under ambiguity. *The Review of Economic Studies*, 74(4):1275–1303, 2007.

H. Ergin and F. Gul. A theory of subjective compound lotteries. *Journal of Economic Theory*, 144(3):899–929, 2009.

J. Etner and M. Jeleva. Risk perception, prevention and diagnostic tests. *Health Economics*, 22(2):144–156, 2013.

J. Etner, M. Jeleva, and J-M. Tallon. Decision theory under ambiguity. *Journal of Economic Surveys*, 26(2):234–270, 2012.

H. Falk. Inequalities of jw gibbs. *American Journal of Physics*, 38(7):858–869, 1970.

J. Gabszewicz and J.-F. Thisse. *Handbook of game theory*, volume 1, chapter Location, pages 281–304. Amsterdam: North Holland, 1992.

P. Ghiradato. Revisiting savage in a conditional world. *Economic Theory*, 20(1):83–92, 2002.

P. Ghirardato, F. Maccheroni, and M. Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118(2):133–173, 2004.

I. Gilboa. *Theory of decision under uncertainty*, volume 1. Cambridge University Press, Cambridge, 2009.

I. Gilboa and D. Schmeidler. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

I. Gilboa, F. Maccheroni, M. Marinacci, and D. Schmeidler. Objective and subjective rationality in a multiple prior model. *Econometrica*, 78(2):755–770, 2010.

F. Gul and W. Pesendorfer. Expected uncertain utility theory. *Econometrica*, 82(1):1–39, 2014.

Y. Halevy. Ellsberg revisited: An experimental study. *Econometrica*, 75(2):503–536, 2007.

P. K. J. Han, R. P. Moser, and W. M. P. Klein. Perceived ambiguity about cancer prevention recommendations: Relationship to perceptions of cancer preventability, risk, and worry. *Journal of Health Communication*, 11(S1):51–69, 2006.

P. K. J. Han, S. C. Kobrin, W. M. P. Klein, W. W. Davis, M. Stefanek, and S. H. Taplin. Perceived ambiguity about screening mammography recommendations: Association with future mammography uptake and perceptions. *Journal of Cancer Epidemiology and Prevention*, 16(3):458–466, 2007.

P. K. J. Han, W. M. P. Klein, and N. K. Arora. Varieties of uncertainty in health care: A conceptual taxonomy. *Medical Decision Making*, 31(6):828–838, 2011.

L. P. Hansen and T. J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91(2):60–66, 2001.

J. F. R. Harter. Hotelling's competition with demand location uncertainty. *International Journal of Industrial Organization*, 15(3):327–334, 1996.

T. Hayashi. Intertemporal substitution, risk aversion and ambiguity aversion. *Economic Theory*, 25(4):933–956, 2005.

D. Heyen and B. R. Wiesenfarth. Informativeness of experiments for meu - a recursive definition. *Journal of Mathematical Economics*, 57:28–30, 2015.

Alan Hájek. Interpretations of probability. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Winter 2012 edition, 2012.

H. Hotelling. Stability in competition. *Economic Journal*, 39(153):41–57, 1929.

R. J. Huang. Ambiguity aversion, higher-order risk attitude and optimal effort. *Insurance: Mathematics and Economics*, 50(3):338–345, 2012.

P. Humphreys. Why propensities cannot be probabilities. *The Philosophical Review*, 94: 557–580, 1985.

L. Hurwicz. Optimality criteria for decision making under ignorance. Discussion paper 370, Cowles Comission, 1951.

H. Jeffreys. *The Theory of Probability.* Oxford University Press, Oxford, 1939. Reprint 1998.

W. E. Johnson. *Logic, 3 vols.* Cambridge University Press, Cambridge, 1921.

E. Karni and D. Schmeidler. Atemporal dynamic consistency and expected utility theory. *Journal of Economic Theory*, 54(2):401–408, 1991.

T. F. Kauffeldt and B. R. Wiesenfarth. Confidence, pessimism, and their impact on product differentiation in a hotelling model with demand location uncertainty. *Discussion Paper Series, University of Heidelberg, Department of Economics*, 2014.

D. Kenkel. *Handbook of Health Economics*, chapter Prevention, pages 1675–1720. Elsevier, Amsterdam, North Holland, 2000.

J. M. Keynes. *Treatise on Probability.* Macmillan, London, 1921.

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

P. Klibanoff, M. Marinacci, and S. Mukerji. Recursive smooth ambiguity preferences. *Journal of Economic Theory*, 114(3):930–976, 2009.

F. H. Knight. *Risk, uncertainty and profit.* Houghton Mifflin, Boston, 1921.

A. N. Kolmogorov. Foundations of the theory of probability. New York: Chelsea, 1950.

D. Korczak, G. Steinhauser, and M. Dietl. Effektivität von maßnahmen im rahmen primärer prävention am beispiel kardiovaskulärer erkrankungen und des metabolischen syndroms. Technical report, Deutsches Institut für Medizinische Dokumentation und Information (DIMDI), 2011.

D. Kreps and E.L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46(1):185–200, 1978.

M. Król. Product differentiation decisions under ambiguous consumer demand and pessimistic expectations. *International Journal of Industrial Organization*, 30(6):593–604, 2012.

S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

P. S. Laplace and F. W. Truscott. A philosophical essay on probabilities, 1814. Dover Publications, English Translation 1951.

P.S. Laplace. *Théorie analytique des probabilités*. Courcier, Paris, 1812.

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

M. J. Machina. Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, 27(4):1622–1668, 1989.

M. J. Machina and M. Siniscalchi. *Handbook of the Economics of Uncertainty*, volume 1, chapter Ambiguity and Ambiguity Aversion, pages 729–807. Elsevier B. V., Amsterdam, 2014.

K. J. Meagher and K. G. Zauner. Product differentiation and location decisions under demand uncertainty. *Journal of Economic Theory*, 117(2):201–216, 2004.

K. J. Meagher and K. G. Zauner. Location-then-price competition with uncertain consumer tastes. *Economic Theory*, 25(4):799–818, 2005.

D. Meintrup and S. Schäffler. *Stochastik: Theorie und Anwendungen*. Springer, Berlin Heidelberg, 2006.

K. Nishimura and H. Ozaki. A simple axiomatization of iterated choquet objectives. *CIRJE*, 2003.

M. Peterson. *An Introduction to Decision Theory*. Cambridge University Press, Cambridge, 2009.

M. C. Politi, P. K. J. Han, and N. F. Col. Communicating the uncertainty of harms and benefits of medical intervention. *Medical Decision Making*, 27(5):681–695, 2007.

K. R. Popper. *Observation and Interpretation*, chapter The propensity interpretation of the calculus of probability, and the quantum theory, pages 65–70. Butterworths Scientific Publications, London, 1957.

K. R. Popper. The propensity interpretation of probability. *British Journal for the Philosophy of Science*, 10(37):25–42, 1959.

J. W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136, 1964.

F. P. Ramsey. *The Foundations of Mathematics and Other Logical Essays*, chapter Truth and probability, pages 156–198. Routledge and Kegan Paul, London, 1931.

F. Riedel. Optimal stopping with multiple priors. *Econometrica*, 77(3):857–908, 2009.

C. Y. Robert and P.-E. Therond. Distortion risk measures, ambiguity aversion and optimal effort. *Astin Bulletin*, 44(02):277–302, 2014.

L. J. Savage. *The foundations of statistics*. Wiley, New York, 1954.

D. Schmeidler. Integral representation without additivity. *Proceedings of the American Mathematical Society*, 97(2):255–261, 1986.

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

U. Segal. Two-stage lotteries without the reduction axiom. *Econometrica*, 58(2):349–377, 1990.

M. Siniscalchi. Vector expected utility and attitudes toward variation. *Econometrica*, 77 (3):801–855, 2009.

M. Siniscalchi. Dynamic choice under ambiguity. *Theoretical Economics*, 6:379–421, 2011.

M. A. Smith, D. Paton, and L. V. Williams. Do bookmakers possess superior skills to bettors in predicting outcomes? *Journal of Economic Behavior and Organization*, 71 (2):539–549, 2009.

A. Snow. Ambiguity and the value of information. *Journal of Risk and Uncertainty*, 40 (2):133–145, 2010.

A. Snow. Ambiguity aversion and the propensities for self-insurance and self-protection. *Journal of Risk and Uncertainty*, 42(1):27–43, 2011.

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

J. Venn. *The logic of chance: an essay on the foundations and province of the theory of probability, with especial reference to its logical bearings and its application to moral and social science, and to statistics.* Macmillan, London, third edition, 1888.

J. Von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior.* Princeton University Press, Princeton, 1944.

T. Wang. Conditional preferences and updating. *Journal of Economic Theory*, 108(2): 286–321, 2003.

B. R. Wiesenfarth. How does information affect primary preventive activities when patients face knightian uncertainty? *Unpublished Manuscript*, 2015.

M. E. Yaari. The dual theory of choice under risk. *Econometrica*, 55(1):95–115, 1987.

M. Zheng, W. Chong, and L. Chaozheng. Optimal nonlinear pricing by a monopolist with information ambiguity. *International Journal of Industrial Organization*, 40:60–66, 2015.

P. Zweifel, F. Breyer, and M. Kifmann. *Health Economics.* Springer, Springer Dordrecht Heidelberg London New York, second edition, 2009.