Essays on Decision Making: Intertemporal Choice and Uncertainty

ISBN 978 90 3610 476 0

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **685** of the Tinbergen Institute Research Series, established through cooperation between Rozenberg Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Essays on Decision Making: Intertemporal Choice and Uncertainty

Essays over Besluitvorming: Tijdsvoorkeuren en Onzekerheid

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof. dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board
The public defence shall be held on
Thursday 26 January 2016 at 13.30 hours

by

Umut Keskin
born in Ankara, Turkey

ERASMUS UNIVERSITEIT ROTTERDAM

Doctoral Committee

Promoters:

Prof. dr. H. Bleichrodt

Prof. dr. P.P. Wakker

Other Members:

Prof. dr. A. Baillon

Dr. A. Predtetchinski

Prof. dr. M. v. Veelen

Acknowledgements

I am grateful to my supervisors, Professor Peter Wakker and Han Bleichrodt, for their support and encouragement throughout my research. This thesis would not have been possible to write without their guidance.

I would like to thank the members of the behavioral economics group at the Tinbergen Institute for their insights and pleasant collaboration during my Ph.D.

Umut Keskin

Contents

Chapter 1

Introduction

Being labeled as a social science, much of economics is about understanding human behavior; be it in the face of uncertainty or delayed payoffs through time or strategic situations such as auctions, bargaining, and so on. This thesis will be concerned with the first two, namely uncertainty and time preferences.

The main focus of this thesis is what we can summarize with two broad titles: "irrationalities" in human behavior and an alternative perspective on "rational behavior". My claim requires a clarification of what is meant by rational or irrational behavior. In one of the early discussions of this topic, Richter (1966) defined a rational consumer as someone for whom there exists a total, reflexive, and transitive binary relation on the set of commodities so that his choice data consists of maximal elements of this binary relation. In this respect, Richter (1966) only imposed minimal consistency conditions on behavior for it to be labeled as rational. Although his setting does not involve any uncertainty or time dimension, analogues of these conditions exist for the models we consider here as well. So one can extend the rationality notion of Richter (1966) to our models too. Yet the essence of his approach to rationality is different than the one we take up in this thesis. This minimalistic approach of Richter would leave little space for discussions on rational behavior because much behavior would be rational except for a few cleverly constructed counterexamples. Instead we will consider more widely accepted norms of rationality and analyze them in the framework of uncertainty and time preferences.

The widely accepted norms of rationality mentioned above are understood to be axioms that lead to decision rules describing people's behavior. In the case of decision making under risk and uncertainty the most commonly used decision model is expected utility, and in the case of dynamic decision making, it is the constant discounted utility model. Although there are models that combine both to explain decision making in a dynamic stochastic settings, in this thesis we study them in isolation to assess the nature of the models in more detail.

Uncertainty and Risk

Attempts to study human behavior in a stochastic environment were mostly done through calculating the expected value of a given random variable until the appearance of the famous St. Petersburg Paradox in 18th century. This paradox was first put forward in a letter by Nicolas Bernoulli (de Montmort 1713). The paradox presents a game of chance where a coin is flipped until a tail shows up, at which stage the game ends. If the tail shows up in the first round, the payoff to the player is 2, if tail shows up in the second round for the first time the payoff is 4, if the tail shows up in the third round the payoff is 8, and so on. It is easy to see that the probability of tail showing up in the n^{th} round is $1/2^n$ in which case the payoff is 2^n. Hence the expected value of the game is $1 + 1 + \ldots$ which is infinite. This presents a paradox in that nobody would pay large sums of money to play this game. Therefore it suggests that we need valuations other than expected value for some gambles. Daniel Bernoulli came up with a suggested solution to the paradox (Bernoulli 1738). Since then his suggestion, to use expected utility rather than expected value, has been the most commonly used tool to analyze individual choice in stochastic environments in economics and other social sciences.

To this day expected utility theory has been widely accepted to be normatively superior to many other models of choice under risk and uncertainty. However its descriptive power has been challenged by the Allais paradox (Allais (1953)) and the Ellsberg paradox (Ellsberg (1961)). After these two famous examples, many other counterexamples violating expected utility theory were provided (see Tversky and Kahneman (1979) for a nice set of early examples). After these series of findings questioning the descriptive power of expected utility, a new theory has been developed to explain choice under uncertainty, namely prospect theory (Tversky and Kahneman 1979, 1992).

In the second chapter of this thesis, we analyze an environment in which there is uncertainty in the sense that the probabilities of events are not given. We study expected utility and deviations from it in an experimental setting. In particular we are interested in the way in which people change their behavior upon receipt of new information regarding the possible events. In standard expected utility, it is assumed that people update their beliefs (their subjective probabilities of relevant events) in accordance with Bayes' rule. Yet there is ample empirical evidence showing that although people do update their beliefs, this updating procedure need not comply with Bayes' rule (Grether 1980, El-Gamal and Grether 1995). Further, in expected utility, utilities that arise from outcomes are weighted by the agents' subjective probabilities. However, in decision models such as the aforementioned prospect theory, utilities are weighted by decision weights that may differ from subjective probabilities. Hence updating probabilities and updating decision weights become different concepts.

In Chapter 2, we assume that decision makers' behavior deviates from the expected utility prescription and their attitudes are instead captured by neo-additive weighting functions (Chateauneuf et al. 2007). These functions are transformations of subjective probabilities and are used in determining decision weights. We designed an experiment and employed a simple method to elicit neo-additive decision weights. In the experiment, subjects traded options on the performance of (anonymous) initial public offerings (IPOs) of new stocks. The reason we chose anonymous IPOs is that this way the subjects did not have any prior knowledge about the performance of the stocks. Then we gave them information about the returns of the stocks and observed how they changed their attitudes upon receipt of these pieces of new information. We found that as people receive new information, they became closer to the behavior that one would see from an expected utility maximizer. If we take expected utility as a benchmark for rationality, then this finding suggests that information makes people behave more rationally. However, we still observed that people deviated from expected utility even when they had more and more information.

In the Chapter 3, we focus on one of the axioms of Bayesian expected utility namely the independence of events. We use informational independence in characterizing different models of ambiguity (unknown probabilities). In the Anscombe-Aumann (1963) axiomatization of expected utility, this notion of independence is implicitly used. In their setting, they use two different types of events: unknown probability events (horse lotteries) and events with known probabilities (roulette lotteries).

In the Anscombe-Aumann setting, roulette lotteries are assumed to be independent of horse lotteries. We show that for the separable representation of preferences this has to be reversed; i.e., one had better assume independence of horse lotteries from roulette lotteries or otherwise a separability paradox results.

We also show that symmetry of independence is necessary and sufficient for expected utility to hold. We further characterize some models other than expected utility using this notion of independence.

Intertemporal Choice

If a historian of economic thought were to study the branches of decision making that specialize in behavior under risk and uncertainty on the one hand and intertemporal choice on the other, she would encounter many similarities. And just like expected utility was (and to a great extent still is) the prominent theory for a long time when there is uncertainty involved, constant discounted utility (with or without linear utility) was the prominent theory of dynamic decision making. Just like the second part of the last century witnessed experimental challenges towards expected utility, constant discounted utility was also questioned on empirical grounds. These two theories share the common normative appeal whereas they lack descriptive power. This led researchers to develop new tools for analyzing decision making in dynamic environments.

The functional forms that represent preferences in intertemporal settings also have behavioral foundations. In the Chapter 4 of this thesis, we present a new tool, the subjective present value, that can be used in characterizing six different decision rules that are commonly used in economics, including constant discounted utility.

We define the present value as follows: Suppose that a decision maker is endowed with a stream of payments spread over future dates and also the present. Then suppose that at some future date we add an amount x to the current endowment at that date. The amount that the decision maker would ask for (without the additional x in future) at present time so that she would be indifferent to the stream with the extra x in the future is called the present value of x in our setting.

We characterize six commonly used models of intertemporal choice in finance and economics by using this present value. It is a more natural tool than other axioms that were previously used in characterizing these models because present values are commonly used in economics and finance and decision makers are familiar with the concept. Therefore this method can more easily be tested empirically, which is an important reason why behavioral axioms must exist in the first place.

As stated in the beginning, the commonly used models mentioned above were empirically challenged. Models with more descriptive power were suggested due to these challenges. The last chapter of this thesis studies two such models, namely variation aversion and decreasing impatience, that can accommodate common violations of constant discounted utility. Both models were characterized before using cumbersome axioms. This chapter presents their characterizations in terms of a new preference condition based on the present value.

To summarize, this thesis has studied some irrationalities in human behavior and also presented novel approaches to defining rational behavior. In the second chapter, we set up an experiment that replicates investment behavior in financial markets and verified departures from expected utility and also that subjects become more rational as they receive information related to unknown events. In the third chapter, we used informational independence to characterize some decision models used for uncertainty. We also pointed to a paradox that results from the way informational independence is implicitly used in Anscombe-Aumann setting and provided an alternative that corrects this paradox. The fourth chapter introduced a natural tool, subjective present value, and used it to give novel characterizations of different decision models of intertemporal choice. The last chapter is an extension of the fourth one and uses the concept of present value to characterize some departures from rational models of intertemporal choice. With all these results, this thesis has shed new light on what is meant by rational and irrational behavior. We carried out our study in the domain of uncertainty and intertemporal choice only. Future research may take up a similar approach for different situations as well.

Chapter 2

The Effect of Learning on Ambiguity Attitudes

Aurlien Baillon[a], Han Bleichrodt[a], Umut Keskin[b], Olivier L'Haridon [c], Chen Li[a]

[a]: Erasmus School of Economics, Erasmus University Rotterdam, the Netherlands

[b]: Istanbul Bilgi University, Istanbul, Turkey

[c]: University of Rennes 1-Crem, Greg-HEC, Rennes, France

Abstract

This paper studies the effect of learning new information on peoples beliefs and their attitudes towards ambiguity. We propose a method to separate ambiguity attitudes from subjective probabilities and to decompose ambiguity attitudes into pessimism (capturing ambiguity aversion) and likelihood insensitivity. We apply our method in an experiment where we elicit the ask prices of options with payoffs depending on the returns of initial public offerings (IPOs) on the New York Stock Exchange. IPOs are a natural context in which to study the effect of learning, as no prior information about returns is available. The results indicate that there was significant likelihood insensitivity, which diminished with more information. We found little pessimism, which was largely unaffected as new information became available. Subjective probabilities were well-calibrated and close to true frequencies. Subjects behavior moved towards expected utility with more information, but substantial deviations remained even in the maximum information condition.

Keywords: ambiguity, learning, updating, neo-additive weighting

2.1 Introduction

In many real-world decision problems, objective probabilities are unknown and decisions have to be made under uncertainty. The traditional approach in decision analysis to analyze such decisions is to assume that the decision maker assigns subjective probabilities to events, behaves according to expected utility, and updates his subjective probabilities according to Bayes rule when new information becomes available. All these assumptions are open to debate.

First, while people change their beliefs when more information becomes available and these updated beliefs have predictive value (Hamermesh 1985, Smith et al. 2001), empirical evidence suggests that they systematically deviate from Bayes rule (e.g. Grether 1980, El-Gamal and Grether 1995, Charness and Levin 2005, Hoffman et al. 2011, Poinas et al. 2012, and Gallagher 2014). Psychologists have uncovered many updating biases, including under- and overconfidence (Griffin and Tversky 1992), conservatism (Phillips and Edwards 1966), representativeness (Kahneman and Tversky 1972), availability (Tversky and Kahneman 1973), and confirmatory bias (Rabin and Schrag 1999).

Second, and even more fundamental, Ellsbergs (1961) paradox, which shows that people prefer betting on known rather than unknown events, undermines not only subjective expected utility, but even the existence of subjective probabilities. To account for Ellsbergs paradox, many new models of decision under ambiguity have been proposed (for overviews see Wakker 2010, Gilboa and Marinacci forthcoming). While in expected utility decision weights are equal to subjective probabilities, in these ambiguity models they also reflect the confidence people have in their beliefs and their aversion towards ambiguity. The ambiguity models capture an intuition expressed already in 1921 by Keynes (1921):

'The magnitude of the probability of an argumentdepends upon a balance between what may be termed the favourable and the unfavourable evidence; a new piece of evidence which leaves this balance unchanged also leaves the probability of the argument unchanged. But it seems that there may be another respect in which some kind of quantitative comparison between arguments is possible. This comparison turns upon a balance, not between the favourable and the unfavourable evidence, but between the absolute amounts of relevant knowledge and relevant ignorance respectively [p.71].

In other words, Keynes conjectured that learning new evidence changes both the balance of evidence (peoples beliefs) and the total amount of evidence (the amount of ambiguity). Under expected utility, the amount of ambiguity plays no role and learning only affects beliefs. In the ambiguity models, more information changes both beliefs and ambiguity attitudes and they make it possible to better understand the effects of learning on behavior. This raises the question of how decision weights are updated.[1] While several papers have approached this question from a theoretical angle and different rules have been proposed,[2] there is a dearth of empirical evidence on how decision weights actually change as more information becomes available[3] This

[1]In the literature the expression updating of non-Bayesian beliefs is sometimes used. To emphasize that beliefs may differ from subjective probabilities under non-expected utility we use the term updating of decision weights.

[2]See Gilboa and Schmeidler 1993, Epstein 2006, Eichberger et al. 2007, Epstein and Schneider 2007, Hanany and Klibanoff 2007, Eichberger et al. 2010, Eichberger et al. 2012).

[3]Cohen et al. 2000) and Dominiak et al. 2012) experimentally studied updating under ambiguity but consider situations in which decision makers receive information that an event cannot occur. In our study decision makers accumulate evidence how often a particular event has been observed in the past.

motivated our paper in which we study experimentally how decision makers change their behavior when more information becomes available.

A difficulty in applying the new ambiguity models is that most of these models involve concepts that are difficult to measure empirically. We present a simple method to measure ambiguity attitudes. Our method is based on an insight from Luce (1991) that for binary acts most ambiguity models are equivalent. We use results from Chateauneuf et al. (2007), to disentangle subjective probabilities and ambiguity attitudes. Our method describes a decision makers ambiguity attitude by two indices, one reflecting his pessimism (capturing ambiguity aversion) and the other his sensitivity to changes in likelihood, which, as we show, is closely related to the decision makers ambiguity perception.

The separation of subjective probabilities and ambiguity attitudes makes it possible to study whether people behave more in line with expected utility when they receive more information. This would be compatible with a commonly held view that learning and more information decrease the irrationalities caused by deviations from expected utility (Myagkov and Plott 1997, List 2004, van de Kuilen and Wakker 2006, Ert and Trautmann 2014).

We applied our method in an experiment, where we elicited subjects ask prices for options with payoffs contingent on the returns of (anonymous) initial public offerings (IPOs). IPOs make it possible to study the effect of more information in a natural decision context (rather than in a more contrived context using urns) for which no prior information is available. The results indicated that pessimism was stable, whereas likelihood insensitivity diminished as more information became available. Aggregate subjective probabilities were close to true frequencies after correction for ambiguity attitudes. Subjects behavior moved into the direction of expected utility with more information about the historical performance of the stocks. However, substantial deviations remained even in the maximum information condition.

2.2 Theoretical Framework

Decision Model

A decision maker faces uncertainty about the outcome he will receive at time T. The decision makers uncertainty is modeled through a finite state space S_T where the subscript T denotes that the uncertainty will be resolved at time point T. The state space contains all possible states of the world s, only one of them finally occurring. The decision maker does not know which state will occur. Events are subsets of S_T. The decision maker chooses between acts, mapping S_T to an outcome space X. In our experiment the outcomes were positive money amounts. We considered only *binary acts*, denoted by $x_E y$, giving money amount x if event E occurs at time T and money amount $y \leq x$ otherwise.

The decision makers information about previous resolutions of uncertainty up to time $t < T$ is formalized by his *history set*, $h_t = (s_1, \dots, s_t)$, where $s_j \in S_j$ for all $1 \leq j \leq t$ and S_j denotes the state space representing the uncertainty at time j. Complete absence of information is denoted by h_0. We assume that $S_t = S_T = S$ for all $t = 1, \dots, T$. In other words, the same states are available at different points in time. The decision makers beliefs may change over time as more information becomes available. The decision makers preferences are represented through a *history dependent preference relation* $\succsim_t$, where the subscript t indicates that preferences depend on the history h_t (with $\succ_t$ and $\sim_t$ defined as usual). A real-valued function V_t *represents* the $\succsim_t$ if for all binary acts $x_E y$ and $z_F w$, $x_E y \succsim_t z_F w$ if and only if $V_t(x_E y) \geq V_t(z_F w)$.

The traditional Bayesian approach assumes that preferences $\succsim_t$ are represented by *expected utility*, i.e.,

$$x_E y \mapsto P_t(E)U(x) + (1 - P_t(E))U(y)$$

where $U : S \to \mathbb{R}$ a utility function defined over outcomes and P_t the subjective probability measure given h_t. In expected utility, new information, which expands the history set from h_t to h_v, with $v > t$, affects probabilities but leaves utility unchanged. Updating takes place in the belief (subjective probabilities) part of the representation and tastes (utility) are not influenced by new information about past events. Time-invariant utility is also commonly assumed in thetheoretical literature on the updating of decision weights under non-expected utility (e.g. Epstein 2006, Eichberger et al. 2007, Epstein and Schneider 2007) and we will also assume it in this paper.

To account for deviations from expected utility, we will assume a *binary rank-dependent utility* (RDU) model (Miyamoto 1988, Luce 1991, Ghirardato and Marinacci 2001), which includes many ambiguity models as special cases. Examples are maxmin expected utility (Gilboa and Schmeidler 1989), alpha-maxmin expected utility (Ghirardato et al. 2004), contraction expected utility (Gajdos et al. 2008), Choquet expected utility (Schmeidler 1989), and prospect theory (Tversky and Kahneman 1992). Under binary RDU, $\succsim_t$ can be represented by

$$x_E y \mapsto W_t(E)U(x) + (1 - W_t(E))U(y) \tag{2.1}$$

with U a real valued function that is unique up to an affine transformation and W_t a unique *weighting function*[4], which need not be additive but satisfies

[4]Sometimes the term capacity is used instead of decision weight.

$W_t(\emptyset) = 0$, $W_t(S_t) = 1$ and $W_t(A) \leq W_t(B)$ if $A \subseteq B$. The subscript t in W_t expresses that the decision weight depends on the history h_t just like P_t in the Bayesian approach.

Chateauneuf et al. (2007) used neo-additive decision weights W_t. These are defined as follows, for a probability measure P_t and parameters a_t and b_t that satisfy $a_t \leq 1$ and $a_t - 2 \leq b_t \leq 2 - a_t$:

$$W_t(E) = \begin{cases} \frac{a_t - b_t}{2} + (1 - a_t)P_t(E) & \text{if } 0 < \frac{a_t - b_t}{2} + (1 - a_t)P_t(E) < 1 \\ 0 & \text{if } \frac{a_t - b_t}{2} + (1 - a_t)P_t(E) \leq 0 \\ 1 & \text{if } \frac{a_t - b_t}{2} + (1 - a_t)P_t(E) \geq 1 \end{cases} \quad (2.2)$$

Neo-additive decision weighting assumes that the decision maker is *probabilistically sophisticated* for a given history, meaning that his decisions can be rationalized by a probability measure P_t. Because k_t and c_t may differ across histories, the decision maker may deviate from probabilistic sophistication when comparing acts involving different histories.

Representation in (1) can be written as:

$$\begin{aligned} x_E y \quad \mapsto \quad & (1 - a_t)[P_t(E)U(x) + (1 - P_t(E))U(y)] \\ & + \frac{a_t - b_t}{2}U(x) + \frac{a_t + b_t}{2}U(y) \end{aligned} \quad (2.3)$$

Equation (2.3) is a linear combination of the maximum utility of $x_E y$, the minimum utility of $x_E y$, and its expected utility. We will refer to (2.3) as the *neo-additive model*. Chateauneuf et al. (2007) imposed stronger constraints: $0 \leq a_t \leq 1$ and $-a_t \leq b_t \leq a_t$. This ensures that decision makers are likelihood insensitive and assign positive weights to extreme outcomes ($-a_t \leq b_t \leq a_t$). Our (weaker) constraints ($a_t \leq 1$ and $a_t - 2 \leq b_t \leq 2 - a_t$) also permit likelihood oversensitivity and zero weights for extreme outcomes.

Likelihood Insensitivity

The parameter a_t in (2.3) reflects the weight that the decision maker gives to expected utility in his evaluation of acts. If a_t is equal to 0 then the decision maker gives maximum weight to expected utility. Larger values of a_t imply that the decision maker gives less weight to expected utility and that he concentrates more on the maximum and minimum utility. In other words, the larger a_t the more the decision maker ignores the relative likelihoods of x and y. This can also be seen from (2.2), where larger values of a_t imply that $P_t(E)$ receives less weight.

Figure 2.1: Likelihood insensitivity: The figure shows the neo-additive weighting function with $a_t > 0$ and $b_t = 0$. The decision maker is insufficiently sensitive to changes in likelihood. The diagonal shows the weighting function when expected utility holds.

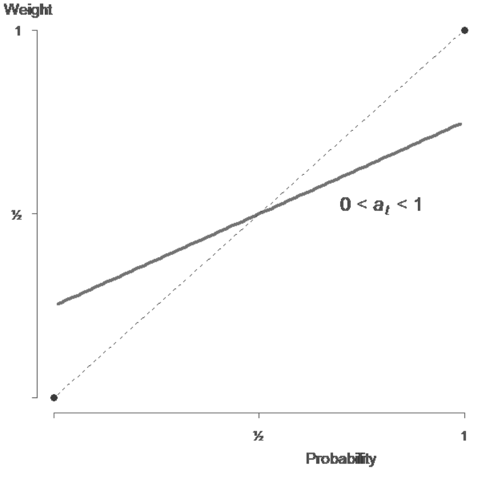


Figure 2.1 shows the effect of changes in a_t when b_t is held constant at 0. When $a_t = 0$, the decision maker behaves according to expected utility (dashed line). When a_t increases, the slope of the decision weighting function becomes flatter and the decision maker is less sensitive to intermediate changes in likelihood. As a result, differences between (non-extreme) decision weights are less than the differences between their underlying probabilities. This is called *likelihood insensitivity.* We take a_t as a *likelihood insensitivity index* with higher values of a_t indicating more likelihood insensitivity.

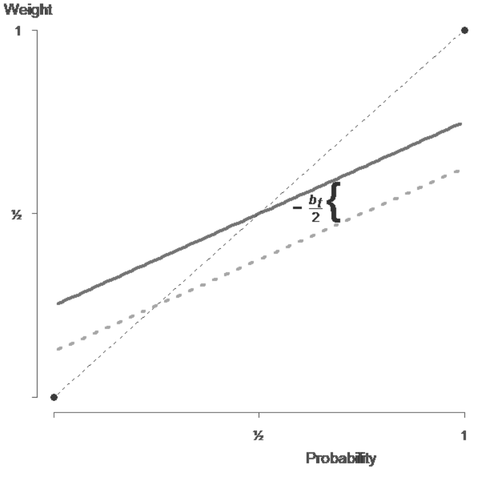
Empirical studies have usually found more likelihood insensitivity for uncertainty than for risk (e.g. Kahneman and Tversky 1979, Kahn and Sarin 1988, Kilka and Weber 2001, Abdellaoui et al. 2005, Wakker 2010, ch. 10). There is also evidence that likelihood insensitivity is stronger for less familiar sources of uncertainty (Kilka and Weber 2001, Abdellaoui et al. 2011). We therefore expect that likelihood insensitivity will diminish with the size of the history set (the amount of information).

Pessimism

Figure 2.2 shows that for a given value of a_t, increases in b_t shift the weighting functions downwards (by $b_t/2$). As can be seen from Eq. (2.1), the decision weights reflect the weight given to the best outcome and, consequently, increases in b_t imply that the decision maker pays more attention to the worst outcome. We will interpret b_t as an index of pessimism with higher values indicating more pessimism, and negative values reflecting optimism. An expected utility maximizer has $a_t = 0$. An extremely pessimistic decision maker, who only considers the worst outcome regardless of its likelihood, has $b_t = 1$ and an extremely optimistic decision maker, who only considers the best outcome, has $b_t = -1$.

Several studies have found that pessimism diminished when the decision maker had more knowledge about a source of uncertainty and, probably, perceived less ambiguity (Heath and Tversky 1991, Kilka and Weber 2001, Fox and Weber 2002, Di Mauro 2008, and Abdellaoui et al. 2011). Pessimism captures the decision makers aversion towards ambiguity. More information may reduce perceived ambiguity, and as the results in the literature suggest, it may also reduce his pessimism.

Figure 2.2: Pessimism. The solid line corresponds to $a_t > 0$ and $b_t = 0$. The parallel dashed line keeps a_t constant and increases b_t. The figure shows that for constant a_t, increases in b_t shift the neo-additive weighting function downwards leading to an increase in pessimism.



The effect of new information on subjective probabilities on the one hand, and on likelihood sensitivity and pessimism on the other hand, illustrates that modern ambiguity theories capture Keynes (1921) intuition about the weight and the balance of evidence. If new information changes the balance of evidence in favor of an event, the decision maker will update his beliefs accordingly. But this new information also changes the balance between the absolute amounts of relevant evidence and relevant ignorance. Our approach reflects this by also allowing changes in the decision makers weighting of subjective probabilities. The new information might make the decision maker rely more on his beliefs and become more sensitive to likelihood, with a_t tending to 0. In section 2.3 we will present a method to disentangle subjective probabilities, pessimism, and likelihood insensitivity and to obtain subjective probabilities that are corrected for ambiguity attitudes. An advantage of our method is that it need not specify an updating rule, because we directly measure subjective probabilities and decision weights from the data.

Multiple Prior Interpretation of The Neo-additive Model

The above analysis is close to Choquet expected utility (Gilboa 1987, Schmeidler 1989) and prospect theory (Tversky and Kahneman 1992) where ambiguity attitudes are modeled through the decision weighting function. The multiple-prior models take a different approach and model ambiguity through a set of priors C_t about the true probability measure P_t. Chateauneuf et al. (2007) showed that the neo-additive model also has a multiple-prior interpretation that can be rewritten as:

$$\begin{aligned} x_E y \quad \mapsto \quad & \alpha_t \min_{\pi \in C_t} [\pi(E)U(x) + (1-\pi(E))U(y)] \\ & + (1-\alpha_t) \max_{\pi \in C_t} [\pi(E)U(x) + (1-\pi(E))U(y)] \end{aligned} \tag{2.4}$$

where $\alpha_t = \frac{a_t + b_t}{2a_t}$ and $C_t = \{\pi | \pi(E) \geq (1-a_t)P_t(E)\}$.

The set of priors, C_t reflects the decision makers perceived ambiguity; the larger the set of priors, the more ambiguity he perceives. Eq. (2.4) shows that the set of priors depends on a_t, which consequently also measures the decision makers ambiguity perception. The parameter α_t reflects the decision makers pessimism. Equation (2.4) is a linear combination of the lowest and the highest expected utility that the decision maker may obtain and higher values of α_t correspond with more weight to the lowest expected utility. Increases in b_t, our measure of pessimism, lead to increases in α_t. However, α_t also depends on a_t and, therefore, it is a different measure of pessimism than b_t.

Equation (2.4) is mathematically equivalent to Eq. (2.3) when a_t is positive. We cannot distinguish these interpretations and the reader can choose the interpretation that he likes best. However, the multiple-prior interpretation only holds in the natural case (with a_t positive). Since several subjects had a negative a_t, we will use only Eq.(2.3) in the individual analyses. On the other hand, as the mean value of a_t was positive we could analyze the aggregate data under both interpretations.

2.3 Measuring Subjective Probabilities and Ambiguity Attitudes

We will now explain how we measured a_t and b_t for different histories h_t. For each history h_t, we considered a three event partition of the state space. The events were defined by the change in the price of stocks on a specific trading day. These events were as follows. Up: the price goes up by at least 0.5%, Middle: the price varies by less than 0.5%; and Down: the price decreases by at least 0.5%.We also considered the event MiddleUp=Middle $\cup$ Up. For given $x > y$, we then elicited four certainty equivalents:

$$CE_{Up} \sim x_{Up}y,\ CE_{Middle} \sim x_{Middle}y,$$
$$CE_{Down} \sim x_{Down}y,\ \text{and } CE_{MiddleUp} \sim x_{MiddleUp}y.$$

With the normalization $U(x) = 1$ and $U(y) = 0$, Eq. (2.1) implies that

$$U(CE_{Up}) = W_t(Up),\quad U(CE_{Middle}) = W_t(Middle),$$
$$U(CE_{Down}) = W_t(Down),\quad U(CE_{MiddleUp}) = W_t(MiddleUp)$$

The decision weights of an expected utility maximizer are equal to his subjective probabilities and, consequently, his subjective probabilities are equal to the utilities of his certainty equivalents. Thus under expected utility,

$$U(CE_{MiddleUp}) + U(CE_{Down}) = P_t(MiddleUp) + P_t(Down) = 1$$

We will refer to this property as *complementarity.* The neo-additive model allows for violations of complementarity:

$$\begin{aligned} &U(CE_{MiddleUp}) + U(CE_{Down}) = \\ &\frac{a_t - b_t}{2} + (1 - a_t)P_t(MiddleUp) + \frac{a_t - b_t}{2} + (1 - a_t)P_t(Down) \\ &= 1 - b_t. \end{aligned} \tag{2.5}$$

Equation (2.5) shows that a neo-additive decision maker violates complementarity if $b_t \neq 0$ and and that more pessimism leads to a lower sum of $U(CE_{MiddleUp}) + U(CE_{Down})$. Hence, studying deviations of this sum allows us to identify the decision makers degree of pessimism.

Under expected utility, the decision maker should also satisfy binary additivity:

$$U(CE_{Up}) + U(CE_{Middle}) - U(CE_{MiddleUp})$$
$$P_t(Up) + P_t(Middle) - P_t(MiddleUp) = 0$$

Under the neo-additive model, we obtain,

$$U(CE_{Up}) + U(CE_{Middle}) - U(CE_{MiddleUp}) = \frac{a_t - b_t}{2} \tag{2.6}$$

Equation (2.6) shows that the neo-additive model predicts violations of binary additivity if $a_t \neq b_t$. The neo-additive model makes it possible to measure pessimism and likelihood insensitivity for any events if utility is known. To measure utility we used the method of Abdellaoui et al. 2008, which we will explain in Section 2.4. Once we know a_t, b_t, and U, we can also determine P_t. If a_t or b_t is unequal to zero, expected utility does not hold and the subjective probabilities that we measure under expected utility will be non-additive: either complementarity or binary additivity will not hold. Our method takes this non-additivity into account and measures subjective probabilities P_t that are corrected for ambiguity attitudes.

2.4 Experiment

2.4.1 Subjects

The experiment was run at Erasmus University in May 2011 with 66 subjects (22 female) with a background in finance. Subjects were either third year undergraduate students with a major in finance or graduate students in finance. Their average age was 24 years, ranging from 21 years to 33 years. We deliberately selected students from finance because the experimental questions involved options and we hoped that finance students would find the experimental tasks easier to understand and would be more motivated to answer the questions. Each subject received a show-up fee of 5 euros and in addition each subject played out one of his choices for real using a procedure described below.

2.4.2 Method

The experiment was computer-run in small group sessions involving at most 3 subjects. Subjects first received instructions and were asked to answer several questions to check their understanding of the experimental tasks. The experimental instructions including the questions to check for subjects understandings are in Appendix 2.B. Subjects could only proceed to the actual experiment after they had answered all test questions correctly.

The source of uncertainty that we used was the variation in the returns on the stocks of IPOs (Initial Public Offerings) traded at the New York Stock Exchange (NYSE). IPOs are new stocks that have just entered the market. We chose IPOs for two reasons. First, the returns on stocks are a natural source of uncertainty unlike, for example, Ellsberg urns. Second, because IPOs are new on the market, there is no price history and learning occurs naturally.

We used data on 328 IPOs in total. All stocks were listed on the NYSE between 1 September 2009 and 25 February 2011. At the start of the experiment, each subject drew 4 numbers which determined the stocks he would trade. Subjects did not know which stocks they traded. The identities of the traded stocks were only revealed after subjects had completed the experiment. We explained subjects how they could verify the stock data on the internet should they wish to do so.

Payoffs were determined by the performance of the stock on the 21$^{\text{st}}$ trading day after their introduction on the NYSE. We defined four events: Up:$(0.5, +\infty)$, i.e. the stock goes up by more than 0.5% on the 21$^{\text{st}}$ trading day, Middle:$[-0.5, 05]$, the stock varies by at most 0.5% on the 21$^{\text{st}}$ trading day, Down:$(-\infty, 0.5)$, the stock goes down by more than 0.5% on the 21$^{\text{st}}$ trading day, and Middle-Up:$(-0.5, +\infty)$, the stock goes up by more than -

0.5% on the 21st trading day. In what follows, we will refer to an option that pays x if event Up obtains as an *Up*(U)*-option*. *Middle-*(M), *Down-*(D), and *Middle-Up*(MU) options are defined similarly. We used the variation in the stock returns rather than the absolute prices of the stocks to make sure subjects had no information about the stocks and to avoid biases. For example, stocks with higher prices might attract more attention leading to biases in the elicited subjective probabilities and ambiguity attitudes.

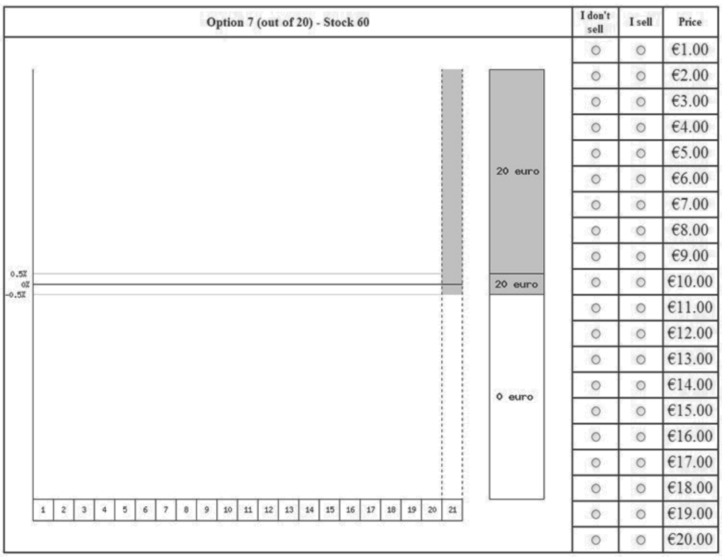
Table 2.1: The 20 Choice Questions

Stk	Cond	y	x	Optn		Stk	Cond	y	x	Optn
1	No Inf	0	10	U		3	1 W	0	20	U
1	No Inf	10	20	U		3	1 W	0	20	M
1	No Inf	5	20	U		3	1 W	0	20	D
1	No Inf	10	15	U		3	1 W	0	20	MU
1	No Inf	0	5	U		3	1 W	0	20	M
1	No Inf	0	20	U		4	1 M	0	20	U
2	No Inf	0	20	U		4	1 M	0	20	M
2	No Inf	0	20	M		4	1 M	0	20	D
2	No Inf	0	20	D		4	1 M	0	20	MU
2	No Inf	0	20	MU		4	1 M	0	20	D

Table 2.1: The columns labeled 'Stk' refer to the four different stocks subjects faced. The questions for stock 1 were used to measure utility. The columns labeled 'Cond' refer to the amount of information subjects received about the historical performance of the stock. 'No Inf', '1 W' and '1 M' correspond to no information, 1 week and 1 month information cases respectively. 'Opt' columns refer to option type and indicate event E. Options were of the type $x_E y$ where the subject received x euros if event E occurred and y euros otherwise.

There were three *informational conditions*, each involving a different history set. In the *no information* condition (history set h_0), subjects had no information about the underlying stock. In the *one week* condition (history set h_5), subjects were informed about the daily returns of the stock in the first 5 trading days following its introduction. Finally, in the *one month* condition (history set h^{20}), subjects were informed about the performance of the stock in the first 20 trading days following its introduction.

Figure 2.3: The choice lists used in the experiment. In this example the option pays 20 euros if event MU occurs on the 21st trading day after the introduction of the stock and 0 otherwise.



We used choice lists to elicit the ask prices of 20 options, summarized in Table 2.1. The ask prices were determined through choice lists. Figure

2.3 gives an example of a choice list for a MiddleUp option. Subjects were told that they owned the option $x_E y$ and they were asked for each price on the choice list whether they wanted to sell the option. The choice lists consisted of 20 prices ranging from $(y + z)$ euros to x euros in increments of $z = \frac{x-y}{20}$ euros. The computer program enforced monotonicity. If a subject indicated for some price that he did not want to sell then the computer automatically selected 'I don't sell for all prices that were lower. Similarly, if a subject indicated for some price that he wanted to sell then the computer automatically selected 'I sell for all prices that were higher. We included two questions to test whether subjects understood the principle of monotonicity and agreed with it. All did.

The 20 choices were divided into four groups (see Table 2.1). Group 1 consisted of six choices to measure utility. The questions in groups 2, 3, and 4 measured the effect of more information on ambiguity attitudes. For groups 3 and 4, we repeated one measurement to test the reliability of our measurements.

The utility questions (group 1) always came first. We counterbalanced the order in which the three information conditions appeared to avoid that a better understanding of the task confounded the effect of more information. We had to use different stocks in each group. If we had used options on the same underlying stock, then subjects who had, for instance, received information on the stocks performance in the first month would have used this information in the no information and in the one week conditions. We also randomized the order of the options within each group.

2.4.3 Incentives

We used a random incentive system. At the end of the experiment, subjects threw a twenty-sided die twice. The first throw selected the choice list and the second throw selected the line of that list to be played out for real. In the selected line, we implemented the choice that the subject had made during the experiment. So if the subject had chosen to sell, we paid him the price. If he had chosen not to sell, we played out the option $x_E y$ and he received x euros if event E had occurred on the 21st trading day and y euros otherwise.

2.4.4 Analysis

To measure utility, we selected history h_0 and elicited the certainty equivalents CE_k of the six binary acts $x_{k_{Up}} y_k$ where $k = 1, \ldots, 6$, the first entries of Table 2.1. By binary RDU:

$$U(CE_k) = W_0(Up)U(x_k) + (1 - W_0(Up))U(y_k)^5 \tag{2.7}$$

We assumed a power utility function, i.e.,

$$U(x) = \begin{cases} x^\beta & \text{if } \beta > 0 \\ \ln x & \text{if } \beta = 0 \\ -x^\beta & \text{if } \beta < 0 \end{cases}$$

The power family is widely used in decision theory and generally fits the data well (Stott 2006). Dividing all money amounts by the maximum payoff 20 euros scales the power utility function such that $U(20) = 1$ and $U(0) = 0$.

We used nonlinear least squares to estimate $W_0(Up)$ and β in (2.7). We then substituted β in Eqs. (2.4) and (2.5) to derive a_t and b_t and the subjective probabilities.

[5] Under subjective expected utility, $W_0(Up) = P_0(Up)$.

As a robustness check we also estimated the parameters β, a_t, b_t and P_t using a non-linear random coefficient model with individual Fechner errors. Rather than estimating parameters for each individual, the random coefficient model estimates the means and standard deviations of the distributions of individual parameters in the population. The parameters a_t, b_t and P_t in the one week and one month conditions were defined as the sum of the random coefficients at history h_0 and a history specific fixed effect. Details of the estimation are in Appendix 2.C.

2.5 Results

2.5.1 Consistency

The consistency of our measurements was good. In both tests, we observed no significant differences between the original and the repeated ask prices and their correlations were substantial (Spearman correlation 0.86 and 0.81, both $p < 0.01$). The mean absolute differences between the ask prices were 1.09 and 1.00 in the two questions, which gives an indication of the average error subjects made.

A comparison between the ask prices of the option $20_{Up}0$ for stocks 1 and 2 (see Table 2.1) gives further information about the quality of the data. In both questions, subjects had no information about the underlying stock and it seems plausible that they treated them similarly. We indeed found no differences between the elicited ask prices and their correlation was substantial (Spearman correlation 0.52, $p < 0.01$), although lower than in the other consistency tests. The mean absolute difference was equal to 1.53.

2.5.2 Subjective Expected Utility

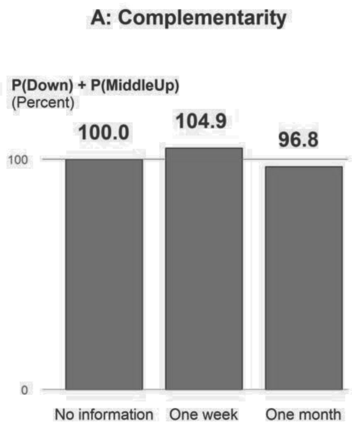
Appendix 2.A shows the median ask prices under the three informational conditions. Under expected utility, the subjective probabilities of the events are $\left(\frac{CE_j}{20}\right)^{\beta}$. Overall, there was little utility curvature both at the aggregate and at the individual level, which is consistent with the hypothesis that utility is about linear for small stakes (Wakker 2010). The median power coefficient was equal to 1 (interquartile range = [0.83,1.28]) and the number of subjects with concave utility (33) did not differ from the number of subjects with convex utility (31).

If expected utility holds then the estimated subjective probabilities $\left(\frac{CE_j}{20}\right)^{\beta}$ should satisfy complementarity and binary additivity. Panel A of Figure 2.4 shows support for complementarity. We could not reject the hypothesis that

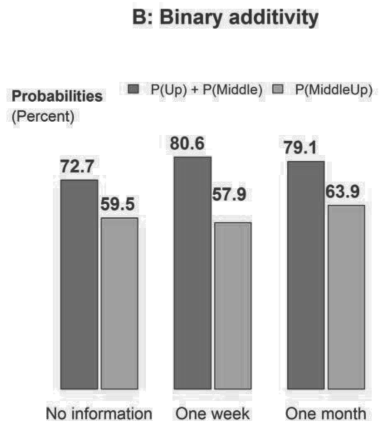
$$P(MiddleUp) + P(Down) = 100\%$$

for all three information conditions. Moreover, we could not reject the hypothesis that the proportion of subjects for whom the sum $P(MiddleUp) + P(Down)$ exceeded 100% and the proportion for whom this sum was less than 100% were the same.

Panel B shows that binary additivity could be rejected and, consequently, that subjective expected utility did not hold. In all three conditions, the sum of $P(Up)$ and $P(Middle)$ exceeded $P(MiddleUp)$ suggesting binary subadditivity instead of binary additivity (Wilcoxon tests, all $p < 0.01$). The difference was significant for all three conditions ($p < 0.01$ in all three tests). There was no difference between the three conditions ($p = 0.13$). At the individual level, we also observed evidence of binary subadditivity: the proportion of subjects who behaved according to binary subadditivity was significantly higher than the proportion of subjects displaying binary superadditivity (bi-



(a) Complementarity



(b) Binary Additivity

Figure 2.4: Tests of complementarity and binary additivity under subjective expected utility. The numbers show the median values. Panel A shows that complementarity held approximately. Panel B shows that binary additivity was violated.

nomial test $p < 0.01$ for all three conditions).

The joint findings of complementarity and binary subadditivity are in line with previous evidence (Tversky and Koehler 1994, Fox and Tversky 1998, Kilka and Weber 2001, Baillon and Bleichrodt forthcoming). They are consistent with support theory, a psychological theory of the formation of subjective probabilities (Tversky and Koehler 1994).

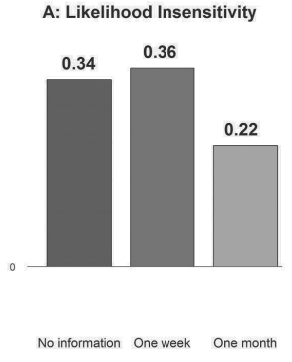
2.5.3 Neo-additive model

Because expected utility was violated, we used Eqs. (2.5) and (2.6) to correct subjective probabilities for ambiguity aversion and to obtain likelihood insensitivity and pessimism indices. Subjects whose (corrected) subjective probabilities were outside the unit interval deviated from the neo-additive model and had to be excluded from the individual analyses. Because these deviations might just reflect error, we included these subjects in the robustness analysis reported in Section 2.5.4.

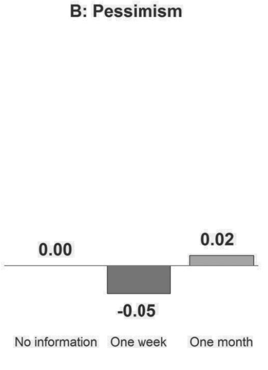
We only excluded subjects for the information condition for which they violated the neo-additive model, but not for the other conditions. This left 56 subjects in the no information condition, 55 subjects in the one week condition, and 52 subjects in the one month condition. To test for robustness and to exclude the possibility of selection bias, we also analyzed the data excluding all subjects who violated the neo-additive model at least once. The results were similar.

Likelihood Insensitivity

Figure 2.5A shows the likelihood insensitivity indices (a_t) for the three conditions. All indices differed from zero suggesting significant likelihood insensitivity (Wilcoxon tests, all $p, 0.01$). Likelihood insensitivity diminished with more information: the median value of a_t fell from 0.34 in the no infor-



(a) Likelihood Insensitivity



(b) Pessimism

Figure 2.5: The likelihood insensitivity and pessimism indices. Panels A and B show the medians of the likelihood insensitivity and pessimism indices for the three information conditions. Likelihood insensitivity falls with more information, but information has no effect on the pessimism indices.

mation condition to 0.22 in the one month condition. Likelihood insensitivity for one month was smaller than for one week (Wilcoxon test, $p = 0.03$), the other indices did not differ.

Figure 2.6: The relations between the individual likelihood insensitivity indices (a_t). If subjects converge to expected utility then the points should lie in the shaded areas. This happens in all panels.

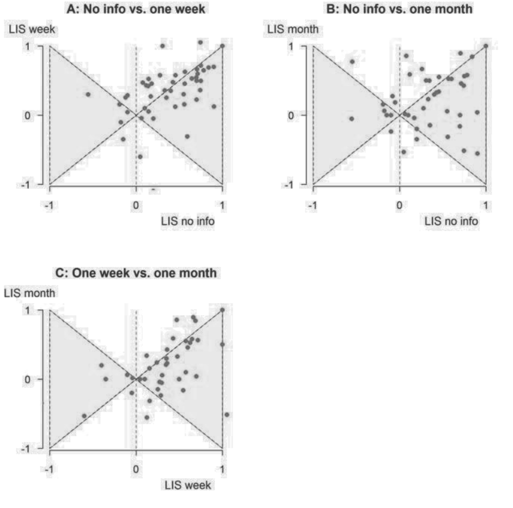


Figure 2.6 displays the individual values of the likelihood insensitivity (LIS) indices for the three information conditions. In each of the panels, the horizontal axis shows the condition in which less information was available. Points on the diagonal represent subjects with the same likelihood insensitivity for the information conditions depicted. If likelihood insensitivity diminished with the amount of information, then the data points should be located below the diagonal.

Figure 2.6 shows that a few subjects had negative likelihood insensitivity indices and were too sensitive to likelihood information. For these subjects, oversensitivity tended to decrease with information. The shaded areas of Figure 2.6 show the subjects who moved in the direction of 'correct sensitivity to likelihood, i.e. to expected utility. The likelihood insensitivity or oversensitity of these subjects decreased but they did not overshoot and went from insensitivity to even larger oversensitivity or from oversensitivity to even larger insensitivity. In all panels, a majority of points (Binomial tests, $p, 0.01$ in all cases) is in the shaded area, which is consistent with convergence towards expected utility with more information.

Pessimism

Figure 2.5B shows the median of the pessimism indices (b_t) for the three information conditions. We could not reject the null of no pessimism in any of the information conditions as none of the pessimism indices differed from zero. There was more optimism in the one week than in the no information condition (Wilcoxon test, p=0.01), the other differences were insignificant.

Figure 2.7 plots the individual pessimism indices for the three information conditions with the condition with less information on the horizontal axis. Points on the diagonal represent individuals with the same pessimism in two information conditions. If pessimism decreases with information then individual points should be in the lower halves of the figures. This was the case for a majority of subjects in Panel A (Binomial test, $p = 0.02$), but not in the other two panels. The shaded areas show the subjects who moved in the direction of expected utility as more information became available. We only observed a significant move to expected utility in the comparison between the one month and the one week conditions (Panel C, Binomial test, $p = 0.01$). However, in interpreting these results it should be kept in mind that most subjects displayed little pessimism in all information conditions.

Figure 2.7: The relations between the individual pessimism indices b_t. If subjects converge to expected utility then the points should lie in the shaded areas. This happens in Panel C.

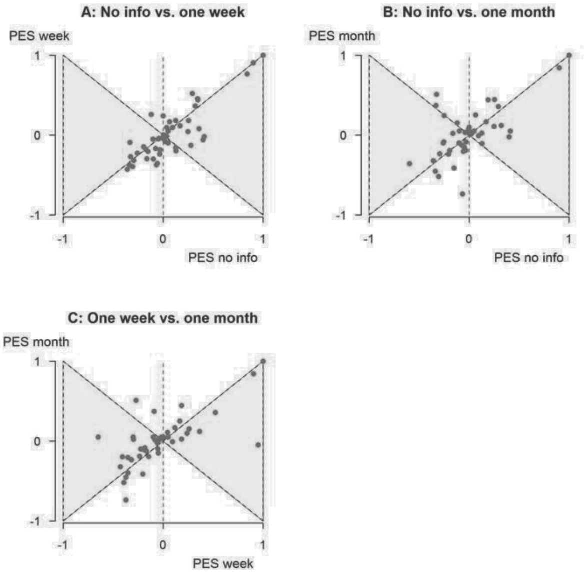


Figure 2.7 suggests that pessimism was a stable trait as the individual data points were clustered around the diagonal. The correlations between the pessimism indices were substantial. The Spearman correlation was 0.76 between the no information and the one week conditions, 0.57 between the no information and one month conditions, and 0.72 between the one week and the one month conditions. They were higher than the correlations between the likelihood insensitivity indices, which varied between 0.19 (no information

and one month) and 0.54 (no information and one week). We conclude that likelihood insensitivity was less stable than pessimism and that it was affected more by information.

2.5.4 Robustness analysis

Table 2.2 summarizes the results of the non-linear random coefficient model. For each parameter of the neo-additive model 1 (Eqn 2.3) and the multiple priors model (Eqn 2.4), we report the estimate of the mean and of the standard deviation. For the likelihood insensitivity and pessimism parameters and the subjective probabilities, we also estimated a fixed effect for the one week and one month conditions. In this estimation, we could include all elicited certainty equivalents, including repeated measurements and the responses of subjects who have violated the neo-additive model.

Table 2.2: Random Coefficients Model

		model 1 Neo-Add	model 2 α-maxmin
LIS	a_0	0.52***	0.51***
		[0.04]	[0.04]
	1w fxd Eff	-0.16**	-0.15***
		[0.05]	[0.05]
	1m fxd Eff	-0.30***	-0.33***
		[0.05]	[0.05]
	σ Rand Eff	0.16***	0.19***
		[0.02]	[0.02]

Table continued on the next page.

PESM	b_0	0.04	
		[0.04]	
	1w fxd Eff	-0.07***	
		[0.02]	
	1m fxd Eff	-0.05**	
		[0.02]	
	σ Rand Eff	0.07***	
		[0.01]	
ALPHA	α_0		0.61***
			[0.01]
	1w fxd Eff		-0.05**
			[0.02]
	1m fxd Eff		0.07**
			[0.05]
	σ Rand Eff		0.00
			[0.01]
$P(Up)$	pup_0	0.41***	0.42***
		[0.02]	[0.01]
	1w fxd Eff	-0.08***	-0.07***
		[0.05]	[0.05]
	1m fxd Eff	-0.07***	-0.08**
		[0.05]	[0.05]
	σ Rand Eff	0.07***	0.07***
		[0.01]	[0.01]

Table continued on the next page.

$P(Middle)$	$pmid_0$	0.30***	0.30***
		[0.02]	[0.02]
	1w fxd Eff	0.004***	0.001***
		[0.02]	[0.02]
	1m fxd Eff	0.05***	0.06***
		[0.02]	[0.02]
	σ Rand Eff	0.13***	0.12***
		[0.01]	[0.01]
Utility	β	1.07	1.15***
		[0.06]	[0.02]
	σ Rand Eff	0.20***	0.26***
		[0.02]	[0.02]
Noise (Fechner Error)	μ_ϵ	-2.33***	-2.39***
		[0.05]	[0.05]
	σ Rand Eff	0.55***	0.59***
		[0.04]	[0.04]
Log-Likelihood		844.32	852.71
N		1280	1280

Table 2.2: Random Coefficients Model. Standard errors in square parentheses. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Neo-additive Model

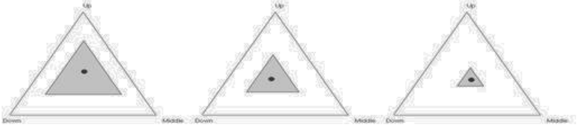
The results of the random coefficients model confirmed most of our conclusions. There was significant likelihood insensitivity in all information conditions, but no pessimism. Likelihood insensitivity diminished as more information became available, suggesting that subjects converged towards expected utility. Our subjects were slightly more optimistic in the one week and one month conditions than in the no information condition. This is consistent with the literature on sources of uncertainty, showing that feeling more knowledgeable reduces ambiguity aversion.

The Multiple Priors Interpretation

Figure 2.8 shows the interpretations of our results in the multiple priors setting. An often-raised objection against these models is that the set of priors is unobservable. Figure 2.8 shows that our method can estimate the set of priors. The black dot shows the estimated subjective probabilities P_t. Together with a_t these determine the set of priors (the light grey area). As the Figure shows, the set of priors decreases with more information and is smallest in the one month condition.

The last column of Table 2.2 shows that the maximum likelihood estimate for α_0 was significantly greater than 0.50 in the no information condition, consistent with ambiguity aversion. The pessimism index α_t decreased in the one week condition but increased in the one month condition. The finding of significant ambiguity aversion in the no information condition is different from what we observed in the neo-additive model. The difference illustrates that α_t and b_t are different measures of pessimism.

Figure 2.8: Sets of priors for the three information conditions (No info, 1 w and 1 m respectively) based on the estimates of Model 2. In each panel, the large triangle is the simplex representing all possible probability measures over the 3 events *Up, Down and Middle*. Each vertex of the simplex denotes an event and corresponds to the measure in which this event is certain. Each opposite side of a vertex represents the probability measures assigning zero probability to the vertex event. The grey triangle is the set of priors and the black dot represents P_t.



Subjective Probabilities

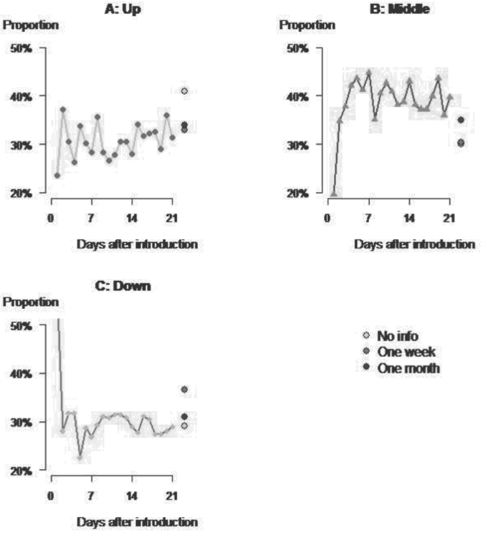
Table 2.2 also shows that $P(Up)$ and $P(Down)$, subjects' subjective probabilities about the events *Up* and *Down*, tended to decrease as more information became available, whereas $P(Middle)$ increased.[6] The elicited probabilities were well-calibrated and close to the true frequencies. For each day from their introduction to the 21st trading day we computed the proportions of the 328 IPOs that went up by more than 0.5% (corresponding with the event *Up*), the proportions that varied by at most 0.5% (corresponding with the event *Middle*), and those that went down by more than 0.5% (corresponding with the event *Down*). A frequentist may interpret these proportions as the

[6]Most differences are significant ($p < 0.01$) except for the differences between $P(Up)$ in the one week and the one month condition, between $P(Middle)$ in the one week and the no information condition, and between $P(Down)$ in the no information and the one month condition.

actual probabilities of the events *Up*,*Middle*, and *Down* at each date t in the history.

Figure 2.9 shows the results of this analysis. Panel A shows the proportions for the event *Up*, Panel B for the event *Middle*, and Panel C for the event *Down*. The figure also shows the estimated probabilities of *Up*,*Middle*, and *Down* for the three information conditions (the dots at the end of the line).

Figure 2.9: Stock history and subjective probabilities. Panel A shows the proportion of the 328 IPOs that went up by more than 0.5% on each trading day from their introduction to the 21st trading day. Panels B and C show the proportions that varied by at most 0.5% and went down by more than 0.5%, respectively. The dots at the end show the estimated probabilities of $P(Up)$ (Panel A), $P(Middle)$ (Panel B), and $P(Down)$ (Panel C) under the three information conditions (in Panel A the points for one week and one month overlap).



All subjective probabilities converged to the actual frequencies in the market. Subjects initially overestimated the probability of the event *Up*. As more information became available, they adjusted their estimate downwards. On the other hand, subjects underestimated the probability of the event *Middle*. This underestimation decreased with information, particularly in the one month condition. Subjects were close to the true frequency of the event *Down* in the no information condition, but then adjusted their estimate upwards in the one week condition, probably because most stocks did not do well in their first five trading days and their returns were highly volatile. In the one month condition, subjects were, again, close to the true frequency.

2.6 Discussion

This paper has studied the effect of learning more information on ambiguity attitudes using a simple method to correct subjective probabilities for likelihood insensitivity and pessimism. The results indicate that there was significant likelihood insensitivity in all information conditions even though we used experienced subjects. Likelihood insensitivity decreased as more information became available and the value of the likelihood insensitivity index fell in the maximum information condition. Subjects went in the direction of correct sensitivity to likelihood information, i.e. they moved towards expected utility. Likelihood insensitivity is often seen as a cognitive bias (Wakker 2010, ch. 7). Our findings suggest that this cognitive bias is reduced with more information.

We found little evidence of pessimism and information had no effect on pessimism. Moreover, the correlations between the pessimism indices were high for the three information conditions. This suggests that pessimism is a stable trait of decision makers preferences and is consistent with the suggestion that pessimism reflects the motivational part of ambiguity attitudes (Wakker 2010, ch. 7). If pessimism is motivational, then more information should not change this inclination.

The finding of little pessimism may be surprising given that most empirical studies have found more pessimism than we did (Trautmann and Van de Kuilen forthcoming). It should be kept in mind that the subjects in our experiment were finance students who were familiar with stocks and options. Empirical evidence suggests that ambiguity aversion decreases when subjects feel competent about the source of uncertainty (Heath and Tversky 1991) and this may have explained why we found little evidence of pessimism.

Another reason for the low amount of pessimism might be the use of ask prices in the elicitation of the certainty equivalents. Ask prices can lead to endowment effects (Kahneman et al. 1990) and, consequently, to an overestimation of certainty equivalents. This would lead to more optimism (Roca et al. 2006, Trautmann et al. 2011). On the other hand, the effect of endowment effects was the same for the three information conditions and, hence, they could not affect our conclusions about the effect of more information on ambiguity attitudes and subjective probabilities.

The joint findings of close-to-zero pessimism and of diminishing likelihood insensitivity as more information became available imply that subjects moved in the direction of expected utility with more information. This agrees with previous findings that experience and learning reduce biases. On the other hand, the likelihood insensitivity index differed significantly from zero even in the one month condition. Moreover, under expected utility, the subjective probabilities violated binary additivity in all information conditions. We conclude that even though more information led to behavior that was more consistent with expected utility, substantial deviations remained.

We made several assumptions in our analysis. First, we assumed that utility did not depend on the information about past events. The utility function reflects preferences over outcomes and more information about the state space has no relevance for these. As we mentioned, this assumption is common in the literature. Abdellaoui et al. (2011) measured utility for different sources of uncertainty and could not reject the null hypothesis that utility was the same across sources.

A more controversial assumption is that probabilistic sophistication held within histories and, hence, that subjective probabilities existed. Different histories can be interpreted as different sources of uncertainty. The notion of sources of uncertainty was first proposed by Amos Tversky in the 1990s (Tversky and Kahneman 1992, Tversky and Fox 1995, Tversky and Wakker 1995). Chew and Sagi (2006, 2008) showed that, if an exchangeability condition holds, subjective probabilities can be defined within sources even when probabilistic sophistication does not hold between sources. Our analysis implicitly assumed this exchangeability condition. Abdellaoui et al. (2011) obtained support for it in all but one of their tests. The only exception was a test involving an unfamiliar source and hypothetical choice. For real incentives, exchangeability always held. Their real incentive system was similar to the one we used. Moreover, because our subjects were finance students, all sources were familiar. Finally, the estimated subjective probabilities were well-calibrated: they were sensitive to more information and they were close to the true frequencies observed in the market. Hence, we are inclined to believe that probabilistic sophistication within histories fitted the preferences of most of our subjects rather well. On the other hand, for a minority of our subjects the estimates did not converge and we found subjective probabilities outside the unit interval, which indicates a poor fit.

We finally assumed that the weighting function could be described by the neo-additive form. This assumption is not very restrictive as the neo-additive weighting function provides a good approximation to more general weighting functions (Diecidue et al. 2009, Abdellaoui et al. 2010). For most subjects the estimated model parameters were plausible and within the range allowed by the model.

2.7 Conclusion

Ambiguity theories are useful for studying the effects of information on decision under ambiguity. Learning new information affects both beliefs and ambiguity attitudes. We have presented a method to separate subjective probabilities from ambiguity attitudes and ambiguity perception. Our method decomposes ambiguity attitudes into likelihood insensitivity and pessimism. We applied our method in an experiment in which we measured the ask prices of options with payoffs depending on the performance of IPOs. The experiment involved three information conditions about historical performance data. The results indicated that there was significant likelihood insensitivity in all three information conditions, but that likelihood insensitivity diminished as more information about the historical performance of the stocks became available. We found little evidence of pessimism and it was not affected by new information. The estimated subjective probabilities, when corrected for ambiguity attitudes, converged to true frequencies. Subjects moved in the direction of subjective expected utility as more information was provided, but substantial deviations remained even in the maximum information condition.

Expected utility is still widely seen as the normative standard for decision under uncertainty. However, it is also well known that people deviate from expected utility and our findings add to the extensive literature on violations of expected utility. The discrepancy between the normative and descriptive status of expected utility makes it desirable to adjust preference measurements for deviations from expected utility. There is a large literature in decision analysis on correcting utility measurements for deviations from expected utility (McCord and de Neufville 1986, Wakker and Deneffe 1996, Delqui 1997, Bleichrodt et al. 2001). Our paper complements this literature by showing how the measurement of subjective probabilities can be corrected for deviations from expected utility. We hope that providing a method to measure (corrected) subjective probabilities and ambiguity attitudes will stimulate the adoption of ambiguity theories in decision analysis practice.

2.8 Appendices

Appendix 2.A: Median Ask Prices

Option	*Up*	*Middle*	*Down*	*MiddleUp*
No Info.	8.50	7.50	7.50	12
1 week	8.50	7.50	8.50	12.50
1 month	7.50	8	7.50	12.50

Appendix 2.B: Experimental Instructions

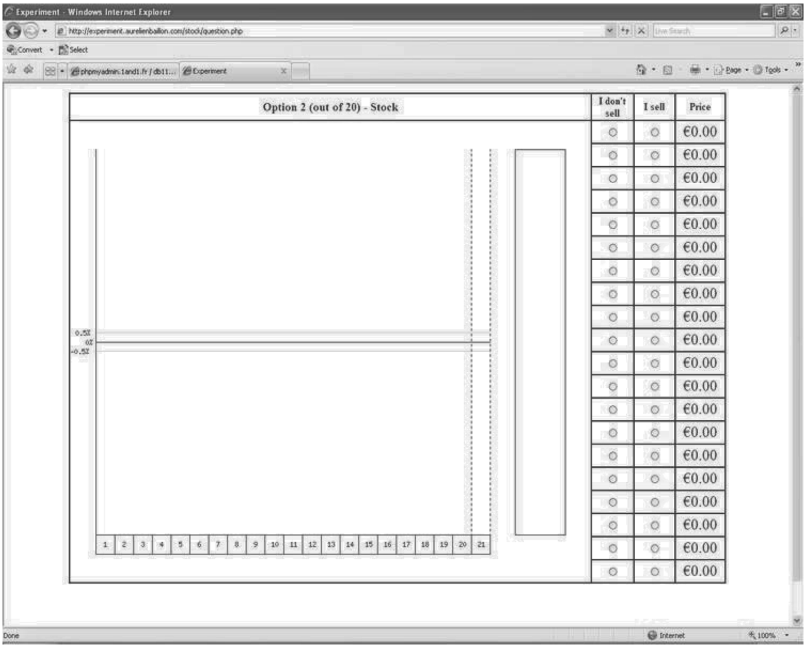
Instructions Thank you for participating in our experiment. For your participation, you will receive a show up fee of 5 euros and an extra payment depending on your choices during the experiment. Please read the instructions carefully. Before starting the experiment, we will ask you several questions to test your understanding of the instructions. If you answer every question correctly, you will proceed to the experiment; otherwise, we will ask you to read the instructions once more and re-answer the questions until all your answers are correct. We want to be sure that you have understood the instructions so that your answers in the experiment reflect your preferences and are not caused by any misunderstandings. If you have any questions, please feel free to ask the experimenter.

During the experiment, you have to answer a series of choice questions. There are **no right or wrong answers** to these questions. We are interested in *your* preferences. Your final payment will be determined by the choices you make during the experiment. Hence it is in your own interest to reveal your true preferences in the choices you will face.

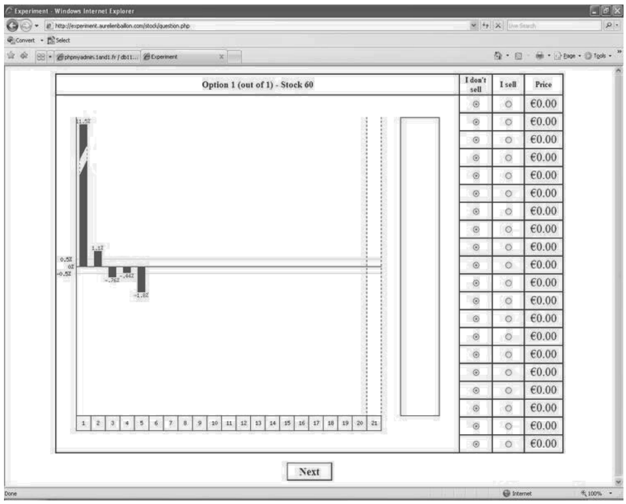
During the experiment, you will be asked to choose between a *digital option* for an underlying stock and a sure money amount. A digital option for an underlying stock pays a pre-specified money amount **H** if a given event occurs and **L** otherwise.

The **underlying stock** is randomly chosen from a database of stocks that were newly-listed on the NYSE between **1 January** 2009 and **25 February** 2011. The stocks in the database are randomly numbered from 1 to 328. At the beginning of the experiment, you will draw 4 numbers from a box, and the 4 corresponding stocks will be used as the underlying stocks of your digital options. At the end of the experiment, the names of the stocks will be revealed, and you can check the historical quotes of the stock prices on Yahoo Finance afterwards. Note that we cannot manipulate the price distribution of the stocks as these are historically given.

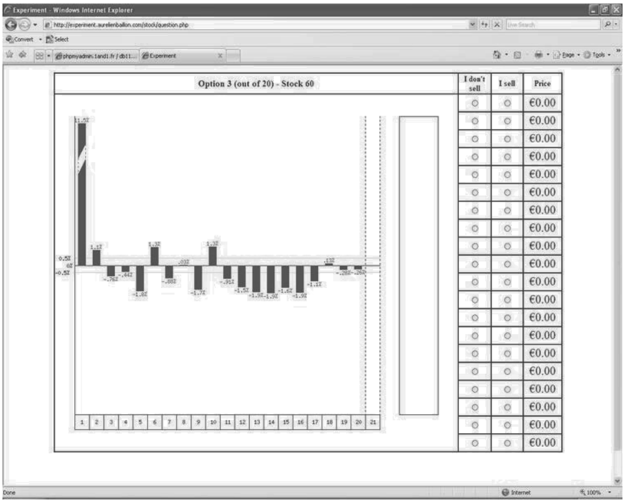
You will face 3 different situations.



Situation 1: You have an option for an underlying stock, which has just been listed on the Stock Exchange. Consequently, you have no quotes of the historical stock price. You know that the expiration date of the option is the 21st trading day of the stock, and the payoff of the option depends on the daily return of the stock on the 21st trading day. (More explanation about the option payoff will be presented later.)



Situation 2: You have an option for an underlying stock, which has been listed on the Stock Exchange for one week. You have 5 quotes of the historical daily return of the stock, which have been depicted by the brown bars. You know that the expiration date of the option is the (same) 21st trading day of the stock, and the payoff of the option depends on the daily return of the stock on the 21st trading day.

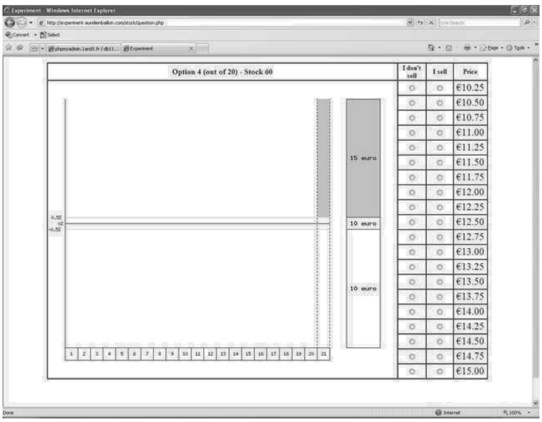


Situation 3: You have an option of an underlying stock, which has been listed on the Stock Exchange for 20 days. You have 20 quotes of the historical daily return of the stock, which have been depicted by the brown bars. You know that the expiration date of the option is the (same) 21st trading day of the stock, and the payoff of the option depends on the daily return of the stock on the 21st trading day.

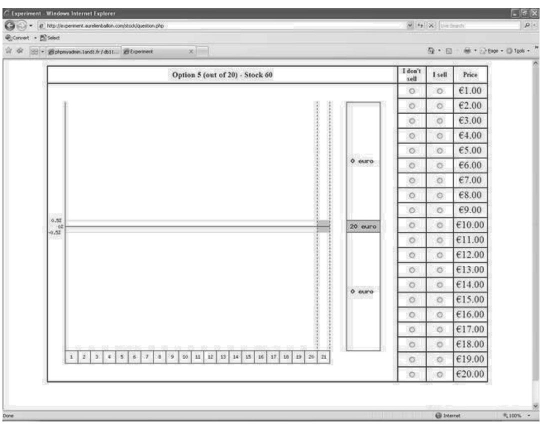
You will face 4 types of *digital options.*

For each situation described above, you may face 4 types of digital options. Here, we use the first situation as an example to illustrate the 4 types of digital options.

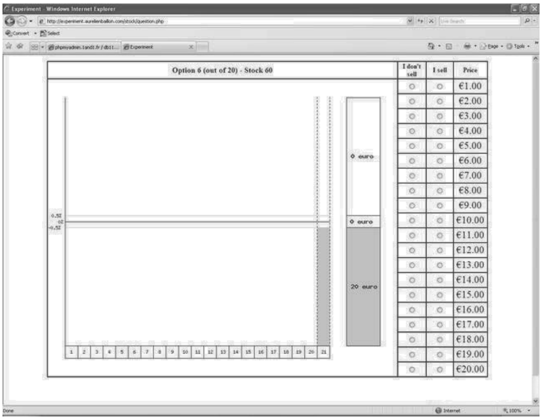
Up-Option



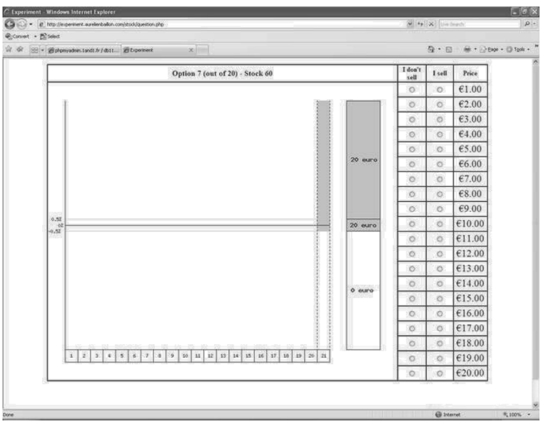
Middle-Option



Down-Option



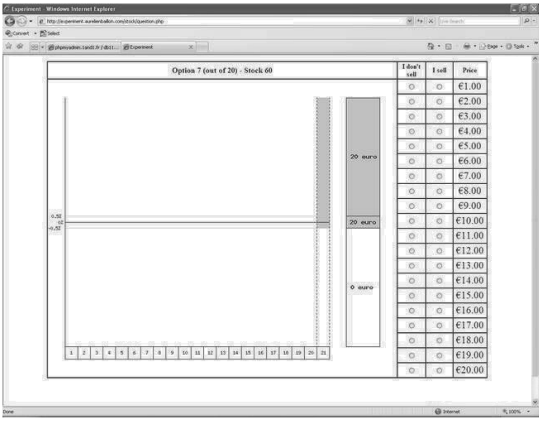
MiddleUp-Option



An **Up-option** pays H euros if the daily return (r) of the underlying stock on its expiration day exceeds +0.5% ($r > +0.5\%$) and L euros otherwise.

A **Middle-option** pays H euros if the daily return (r) of the underlying stock on its expiration day varied between -0.5% and +0.5% ($-0.5\% \leq r \leq +0.5\%$) and L euros otherwise.
A **Down-option** pays H euros if the daily return (r) of the underlying stock on its expiration day is less than -0.5% ($r < +0.5\%$) and L euros otherwise.
A **MiddleUp-option** pays H euros if the daily return (r) of the underlying stock on its expiration day exceeds -0.5% ($r > -0.5\%$) and L euros otherwise.

H and L are pre-specified money amounts. For instance, the figure above displays an Up-option with H=15 and L=10, and the other three types with H=20 and L=0. You may encounter different H and L in the experiment.



We will determine your selling price of 20 different options through a series of choices between the option and a certain money amount. An example is given in the above figure. For each of the 20 prices, you are asked to indicate whether you would like to sell the option or not. The money amount where you switch your choice from I don't sell to I sell is taken as your selling price. All sales will be realized on the 21st day.

If you sell at x euros, do you agree that you also want to sell at prices higher than x euros? Y/N
If you don't sell at y euros, do you agree that you don't want to sell at prices lower than y? Y/N

Payment

As an example, imagine that you throw 7 on your first throw and 6 on your second. Hence the 7th choice will be selected and the price you are offered for the option in the 7th choice is 6 euros. Suppose that option in the 7th choice is a MiddleUp-option with H=20 and L=0, as in the figure above. Suppose further that your selling price for the 7th option was found to be 9 euros. This means that you are not willing to sell the option for a price less than 9 euros and, hence, you do not accept the offered price of 6 euros and thus **you keep the option**;

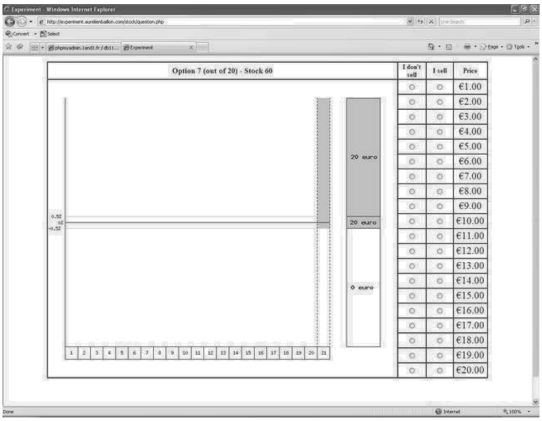
If the daily return on the 21st trading day of the underlying stock is at least -0.5% (e.g. 0.15%), then we pay you 20 euros plus the 5 euros show-up fee. In total you get 25 euros.

If the daily return on the 21st trading day of the underlying stock is smaller than -0.5% (e.g. -1.49%), then we pay you €0 plus the 5 euros show-up fee. In total you get 5 euros.

Now imagine you throw 7 and 10. Then the price offer you are offered is 10 euros. Because you are willing to sell the option if the price is at least 9 euros, you accept the offered price of 10 euros and thus we pay you 10 euros plus the 5 euros show-up fee. In total you get 15 euros.

Note that it is in your best interests to state your selling price truthfully. To see that, suppose your true selling price is 9 euros, but you state a selling price of 11 euros. Then if the price we offer for the option is 10 euros, you keep the option even though it is worth less to you than 10 euros.

Questions



Suppose you are going to play the choice in the picture above for real.

1. What is the minimum selling price?

2. What is the payoff of the plotted option, if the daily return on the 21st trading day is:

   (a) 1.4%

   (b) −0.45%

   (c) −1.4%

3. Suppose that the daily return on the 21st trading day is 1.4%, what is the total payment you get if the second number you throw is 1?

4. Suppose the daily return on the 21st trading day is 1.4%, what is the total payment you get if the second number you throw is 15?

Appendix 2.C: Details of the Random Coefficient Model Estimation

Let $CE_{it}(x_{E_j}y)$ denote subject i's certainty equivalent of option j with history h_t, where $E_j \in \{Up, Middle, Down, MiddleUp\}$ and $t \in \{0, 1w, 1m\}$. To account for errors in subjects reported certainty equivalents, we add a stochastic term ϵ_{ijt} to the certainty equivalent predicted by Equation (2.1) with power utility,

$$CE_{it}(x_{E_j}y = \Big(W_{it}(E_j)x^{\beta_i} + \Big(1 - W_{it}(E_j)\Big)y^{\beta_i}\Big)^{1/\beta_i} \tag{2.8}$$

The individual parameter β_i is normally distributed with mean β, and variance σ^2_β. The history dependent individual weighting function W_{it} is defined according to Equation (2.2) and depends on the parameter vector $\eta_{it} = \{a_{it}, b_{it}, pup_{it}, pmid_{it}\}$ and $pup_{it} = P_{it}(Up)$, $pmid_{it} = P_{it}(Middle)$. The certainty equivalent predicted by Equation (2.1) with utility parameter β_i and the weighting function parameters η_{it} is denoted by $\widehat{CE_{it}}(x_{E_j}y)$.

We assume that η_{i0} follows a multivariate normal distribution with mean η_0 and diagonal variance-covariance matrix $\sigma^2_{\eta_0}$. For $t \in \{0, 1w, 1m\}$, let $\triangle\eta_t = \eta_{it} - \eta_{i0}$. The error term ϵ_{ijt} is assumed to follow a normal distribution with mean 0 and standard deviation σ_i, where σ_i follows a lognormal distribution with parameters μ_ϵ and σ^2_ϵ.

Let $\xi_i = (\beta_i, a_{i0}, b_{i0}, pup_{i0}, pmid_{i0}, \sigma_i)$ be the vector of individual specific random parameters, which are assumed to be independent of each other. Let f denote the density function of ξ and let $\theta = (\beta, \eta_0, \triangle\eta_{1w}, \eta_{1m}, \mu_\epsilon, \sigma_\beta, \sigma_{\eta_0}, \sigma_\epsilon)$ denote the vector of model parameters. For a given θ, the contribution to the likelihood for subject i is therefore:

$$\iota_i(\theta) = \int_{\mathbb{R}^6} \Big[\prod_{j,t} \frac{1}{\sigma_i}\phi\Big(\frac{CE_{it}(x_{E_j}y) - \widehat{CE_{it}}(x_{E_j}y)}{\sigma_i}\Big)\Big] f(\xi|\theta)d\xi, \tag{2.9}$$

where ϕ is the standard normal density function. The log-likelihood is given by the sum of the logarithm of ι_i for all subjects. To approximate the multiple integral in Eq. (2.9), we used simulation techniques, where Halton sequences of length 500 were drawn for each individual (Train 2009, ch. 9). We maximized the log-likelihood function with respect to the vector of model parameters θ using the 'fminunc function in Matlab.

2.9 References

Abdellaoui, M., A. Baillon, L. Placido, P. P. Wakker. 2011. The rich domain of uncertainty: Source functions and their experimental implementation. American Economic Review 101 695-723.

Abdellaoui, M., O. lHaridon, H. Zank. 2010. Separating curvature and elevation: A parametric probability weighting function. Journal of Risk and Uncertainty 41 39-65.

Abdellaoui, M., F. Vossmann, M. Weber. 2005. Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. Management Science 51 1384-1399.

Abdellaoui, M., H. Bleichrodt, O. l'Haridon. 2008. A tractable method to measure utility and loss aversion under prospect theory. Journal of Risk and Uncertainty 36 245-266.

Baillon, A., H. Bleichrodt. forthcoming. Testing ambiguity models through the measurement of probabilities for gains and losses. American Economic Journal: Micro.

Bleichrodt, H., J. L. Pinto, P. P. Wakker. 2001. Making descriptive use of prospect theory to improve the prescriptive use of expected utility. Management Science 47 1498-1514.

Charness, G., D. Levin. 2005. When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity, and affect. American Economic Review 88 933-946.

Chateauneuf, A., J. Eichberger, S. Grant. 2007. Choice under uncertainty with the best and worst in mind: Neo-additive capacities. Journal of Economic Theory 137 538-567.

Chew, S. H., J. S. Sagi. 2006. Event exchangeability: Probabilistic sophistication without continuity or monotonicity. Econometrica 74 771-786.

Chew, S. H., J. S. Sagi. 2008. Small worlds: Modeling attitudes toward sources of uncertainty. Journal of Economic Theory 139 1-24.

Cohen, M., I. Gilboa, J. Jaffray, D. Schmeidler. 2000. An experimental study of updating ambiguous beliefs. Risk, Decision, and Policy 5 123-133.

Delqui, P. 1997. 'Bi-matching': A new preference assessment method to reduce compatibility effects. Management Science 43 640-658.

Di Mauro, C. 2008. Uncertainty aversion vs. competence: An experimental market study. Theory and Decision 64 301-331.

Diecidue, E., U. Schmidt, H. Zank. 2009. Parametric weighting functions. Journal of Economic Theory 144 1102-1118.

Dominiak, A., P. Drsch, J. Lefort. 2012. A dynamic Ellsberg urn experiment. Games and Economic Behavior 75 625-638.

Eichberger, J., S. Grant, D. Kelsey. 2007. Updating Choquet beliefs. Journal of Mathematical Economics 43 888-899.

Eichberger, J., S. Grant, D. Kelsey. 2010. Comparing three ways to update Choquet beliefs. Economics Letters 107 91-94.

Eichberger, J., S. Grant, D. Kelsey. 2012. When is ambiguity attitude constant? Journal of Risk and Uncertainty 45 239-263.

El-Gamal, M. A., D. M. Grether. 1995. Are people Bayesian? uncovering behavioral strategies. Journal of the American Statistical Association 90 1137-1145.

Ellsberg, D. 1961. Risk, ambiguity and the savage axioms. Quarterly Journal of Economics 75 643-669.

Epstein, L. G. 2006. An axiomatic model of non-Bayesian updating. The Review of Economic Studies 73 413-436.

Epstein, L. G., M. Schneider. 2007. Learning under ambiguity. The Review of Economic Studies 74 1275-1303.

Ert, E., S. T. Trautmann. 2014. Sampling experience reverses preferences for ambiguity. Journal of Risk and Uncertainty 49 31-42.

Fox, C. R., M. Weber. 2002. Ambiguity aversion, comparative ignorance, and decision context. Organizational Behavior and Human Decision Processes 88 476-498.

Fox, C. R., A. Tversky. 1998. A belief-based account of decision under uncertainty. Management Science 44 879-895.

Gajdos, T., T. Hayashi, J. M. Tallon, J. C. Vergnaud. 2008. Attitude toward imprecise information. Journal of Economic Theory 140 27-65.

Gallagher, J. 2014. Learning about an infrequent event: Evidence from flood insurance take-up in the united states. American Economic Journal: Applied Economics 6 206-233.

Ghirardato, P., F. Maccheroni, M. Marinacci. 2004. Differentiating ambiguity and ambiguity attitude. Journal of Economic Theory 118 133-173.

Ghirardato, P., M. Marinacci. 2001. Risk, ambiguity, and the separation of utility and beliefs. Mathematics of Operations Research 26 864-890.

Gilboa, I., D. Schmeidler. 1993. Updating ambiguous beliefs. Journal of Economic Theory 59 33-49.

Gilboa, I. & Marinacci, M. (forthcoming). Ambiguity and the Bayesian paradigm. Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society.

Gilboa, I. 1987. Expected utility with purely subjective non-additive probabilities. Journal of Mathematical Economics 16 65-88.

Gilboa, I., D. Schmeidler. 1989. Maxmin expected utility with a non-unique prior. Journal of Mathematical Economics 18 141-153.

Grether, D. M. 1980. Bayes rule as a descriptive model: The representativeness heuristic. The Quarterly Journal of Economics 95 537-557.

Griffin, D., A. Tversky. 1992. The weighing of evidence and the determinants of confidence. Cognitive psychology 24 411-435.

Hamermesh, D. S. 1985. Expectations, life expectancy, and economic behavior. The Quarterly Journal of Economics 100 389-408.

Hanany, E., P. Klibanoff. 2007. Updating preferences with multiple priors. Theoretical Economics 2 261-298.

Heath, C., A. Tversky. 1991. Preference and belief: Ambiguity and competence in choice under uncertainty. Journal of Risk and Uncertainty 4 4-28.

Hoffman, R. M., J. H. Kagel, D. Levin. 2011. Simultaneous versus sequential information processing. Economics Letters 112 16-18.

Kahn, B. E., R. K. Sarin. 1988. Modeling ambiguity in decisions under uncertainty. Journal of Consumer Research 15 265-272.

Kahneman, D., A. Tversky. 1972. Subjective probability: A judgment of representativeness. Cognitive psychology 3 430-454.

Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. Econometrica 47 263-291.

Kahneman, D., J. L. Knetsch, R. H. Thaler. 1990. Experimental test of the endowment effect and the coase theorem. Journal of Political Economy 98 1325-1348.

Keynes, J. M. 1921. A Treatise on Probability. McMillan, London.

Kilka, M., M. Weber. 2001. What determines the shape of the probability weighting function under uncertainty? Management Science 47 1712-1726.

List, J. A. 2004. Neoclassical theory versus prospect theory: Evidence from the marketplace. Econometrica 72 615-625.

Luce, R. D. 1991. Rank-and sign-dependent linear utility models for binary gambles. Journal of Economic Theory 53 75-100.

McCord, M., R. de Neufville. 1986. Lottery equivalents: Reduction of the certainty effect problem in utility assessment. Management Science 32 56-60.

Miyamoto, J. M. 1988. Generic utility theory: Measurement foundations and applications in multiattribute utility theory. Journal of Mathematical Psychology 32 357-404.

Myagkov, M., C. R. Plott. 1997. Exchange economies and loss exposure: Experiments exploring prospect theory and competitive equilibria in market environments. The American Economic Review 87 801-828.

Phillips, L. D., W. Edwards. 1966. Conservatism in a simple probability inference task. Journal of Experimental Psychology; Journal of Experimental Psychology 72 346.

Poinas, F., J. Rosaz, B. Roussillon. 2012. Updating beliefs with imperfect signals: Experimental evidence. Journal of Risk and Uncertainty 44 219-241.

Rabin, M., J. L. Schrag. 1999. First impressions matter: A model of confirmatory bias. The Quarterly Journal of Economics 114 37-82.

Roca, M., R. M. Hogarth, A. J. Maule. 2006. Ambiguity seeking as a result of the status quo bias. Journal of Risk and Uncertainty 32 175-194.

Schmeidler, D. 1989. Subjective probability and expected utility without additivity. Econometrica 57 571-587.

Smith, V. K., D. H. Taylor Jr, F. A. Sloan. 2001. Longevity expectations and death: Can people predict their own demise? American Economic Review 1126-1134.

Stott, H. P. 2006. Cumulative prospect theory's functional menagerie. Journal of Risk and Uncertainty 32 101-130.

Train, K. E. 2009. Discrete Choice Methods with Simulation. Cambridge university press.

Trautmann, S. T., F. M. Vieider, P. P. Wakker. 2011. Preference reversals for ambiguity aversion. Management Science 57 1320-1333.

Trautmann, S. T., Van de Kuilen, G. (forthcoming). Ambiguity attitudes. G. Keren, G. Wu, eds. Blackwell Handbook of Judgment and Decision Making, Blackwell.

Tversky, A., D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. Cognitive psychology 5 207-232.

Tversky, A., D. K. Koehler. 1994. Support theory: A nonexistential representation of subjective probability. Psychological Review 101 547-567.

Tversky, A., C. Fox. 1995. Weighing risk and uncertainty. Psychological Review 102 269-283.

Tversky, A., P. P. Wakker. 1995. Risk attitudes and decision weights. Econometrica 63 297-323.

Tversky, A., D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5 297-323.

van de Kuilen, G., P. P. Wakker. 2006. Learning in the Allais paradox. Journal of Risk and Uncertainty 33 155-164.

Wakker, P. P. 2010. Prospect Theory: For Risk and Ambiguity. Cambridge University Press, Cambridge UK.

Wakker, P. P., D. Deneffe. 1996. Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. Management Science 42 1131-1150.

Chapter 3

Statistical Independence for Axiomatizing Bayesian Expected Utility and Non-Bayesian Ambiguity

Umut Keskin[a], Asli Selim[b], Umut Keskin[b], Peter P. Wakker[c]

[a]: Istanbul Bilgi University, Istanbul, Turkey

[b]: Booking.com, Amsterdam, the Netherlands

[c]: Erasmus School of Economics, Erasmus University Rotterdam, the Netherlands

Abstract

Informational (statistical) independence is an important tool in probability assessments. We extend this concept to decision theory. We use it as a primitive in preference foundations and investigate its implications in modern non-Bayesian ambiguity theories. Symmetry of informational independence is enough to provide a new foundation of Bayesian expected utility. Nonsymmetric versions can be reconciled with ambiguity models, where they give useful implications. They generate a separability paradox for the popular Anscombe-Aumann framework for analyzing ambiguity. The two stages of that framework can better be reversed, as in Jaffray's framework.

Keywords: statistical independence, preference foundation, expected utility, ambiguity, Anscombe-Aumann model

3.1 Introduction

Statistical independence of an event from another (conditioning) event means that the latter is not informative about the former. In probability theory, it means that the probability of the former event is not impacted by conditioning on the latter event. This concept of informational independence is commonly used as a primitive in probability assessments (Smith & von Winterfeldt 2004, p. 565), Bayesian networks (Halpern 2003 Ch. 4; Jensen & Nielsen 2013; Pearl 2000; Williamson 2005), and causal decision theory (Glynn 2011; Harper, Chow, & Murray 2012; Tversky & Kahneman 1980). It appeals to a basic intuition in people.

This paper extends informational independence, or independence for short, into decision theory. Independence requires that a conditioning event not provide relevant information for some other (essential) event and, hence, the conditioning event does not impact preferences on payments contingent on the essential event. As a first topic, we study which decision models can accommodate (informational) independence, and then use it to axiomatize some models. Remarkably, symmetry of independence is enough to imply Bayesian expected utility. Without symmetry, independence can be used in some ambiguity (unknown probability) models that generalize Bayesian expected utility.

The Bayesian statisticians Bernardo, Ferrandiz, & Smith (1985) first introduced statistical independence as a primitive in decision theory. They only considered Bayesian expected utility, and assumed a rich structure comprising both Savage's (1954) and Anscombe-Aumann's (1963) (AA) framework. We simplify and generalize their derivation of expected utility, and extend the analysis to ambiguity models.

Independence is implicitly used in the AA framework, which is nowadays the most popular framework for analyzing ambiguity. In this framework, so-called roulette events (known probabilities) are assumed independent of so-called horse events (ambiguous). We show that this independence assumption leads to a separability paradox. We hence recommend a reversal of stages in the AA framework, as in Jaffray's models.

3.2 Notation, definitions, and well-known representations

We only consider two complementary events E_1 and E_2, and two other complementary events C_a and C_b. Of each pair, exactly one is true and the other is not. For each pair, it is uncertain which event is the true one. The extension of our results to general independent partitions with n events $E_1, \ldots, E_n$ and m events $C_1, \ldots, C_m$, or infinitely many events, is straightforward. This paper does not seek mathematical generality and aims to keep technical details to a minimum, so as to make the conceptual issues maximally clear. We therefore focus on two binary partitions. We call E_1 and E_2 *essential*. In most examples, these events will determine the outcomes. C_a and C_b are *conditioning events*. In most examples, they serve to provide information about E_1 and E_2.

In the AA framework, defined later, C_a (say an American horse) and C_b (say a Belgian horse) refer to two horses, one of which will win the next race, and E_1 and E_2 refer to two roulette events (say odd and even).[1]

[1]Which of these events are best suited to play the role of essential or conditioning event will be a central point of debate in what follows. We will assume later that odd and even have 0.5 probabilities in the roulette example. In the regular roulette game, these probabilities may be slightly different, e.g., because of the 0 number.

The above events generate a *state space* S through their intersections in the usual way: $S = \{E_1C_a, E_1C_b, E_2C_a, E_2C_b\}$. All subsets of S are called *events*.

$x = (x_{1a}, x_{1b}, x_{2a}, x_{2b})$ refers to a *prospect* yielding *outcomes* x_{1a} if E_1 and C_a occur, with the other outcomes defined similarly. Outcomes are monetary. The prospect is displayed in Eq. 3.1.

$$x = \quad \begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & x_{1a} & x_{1b} \\ E_2 & x_{2a} & x_{2b} \end{array} \tag{3.1}$$

The set of prospects is isomorphic to $\mathbb{R}^4$. We often use Greek letters $\alpha, \beta, \gamma, \delta$ to designate outcomes. Then $(\alpha_{1a}, \beta_{1b}, \gamma_{2a}, \delta_{2b})$ denotes the obvious prospect. For example, $(\alpha_{1a}, \alpha_{1b}, \beta_{2a}, \beta_{2b})$ yields α under E_1 and β under E_2. By $\succcurlyeq$ we denote a preference relation over the prospects, with $\succ, \preccurlyeq, \sim$ defined as usual. We assume *weak ordering* (completeness: $x \succcurlyeq y$ or $y \succcurlyeq x$ for all prospects x, y, and transitivity), *monotonicity* (any increase of any outcome is strictly preferred), and continuity throughout. We summarize the assumptions made throughout this paper.

Structural Assumption: Prospects (denoted by $(x_{1a}, x_{1b}, x_{2a}, x_{2b})$) are mappings from the state space $\{E_1C_a, E_1C_b, E_2C_a, E_2C_b\}$ to $\mathbb{R}$, the outcome set. $\succcurlyeq$ is a monotonic continuous weak order on $\mathbb{R}^4$, the set of prospects.

E_1-prospects are prospects whose values depend only on E_1 and its complement E_2. That is, they are prospects of the form $(\alpha_{1a}, \alpha_{1b}, \beta_{2a}, \beta_{2b})$, also denoted $\alpha_{E_1}\beta$, and they yield outcome α under event E_1 and β under event E_2. *E_1-preferences* designate the preference relation $\succcurlyeq$ restricted to E_1-prospects. *C_a-prospects* $(\alpha_{1a}, \beta_{1b}, \alpha_{2a}, \beta_{2b})$ are defined similarly.

The following condition entails, informally, that preferences conditional on C_a are independent of common outcomes (denoted c_j below) outside of

C_a. Event C_a is *separable* if

$$\begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \alpha & c_1 \\ E_2 & \beta & c_2 \end{array} \quad \succcurlyeq \quad \begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \gamma & c_1 \\ E_2 & \delta & c_2 \end{array} \tag{3.2}$$

implies

$$\begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \alpha & c_1' \\ E_2 & \beta & c_2' \end{array} \quad \succcurlyeq \quad \begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \gamma & c_1' \\ E_2 & \delta & c_2' \end{array} \tag{3.3}$$

for all outcomes considered.

Separability of any event other than C_a is defined similarly. It again refers to independence of preferences conditional on that event from the level at which common outcomes outside that event are kept fixed. Monotonicity implies that all "atomic" events $E_1C_a, E_1C_b, E_2C_a, E_2C_b$ are separable. *Separability* without qualification means that all event are separable. It is Savage's (1954) sure-thing principle, being his postulate P2.

A function V *evaluates* prospects if $V : \mathbb{R}^4 \to \mathbb{R}$ and $x \succcurlyeq y \Leftrightarrow V(x) \geq V(y)$. *Expected utility* (*EU*) holds if there exist *probabilities* $p_{1a}, p_{2a}, p_{1b}, p_{2b}$ (positive and summing to 1) and a continuous strictly increasing *utility function* $U : \mathbb{R} \to \mathbb{R}$ such that $x \mapsto \sum_{j\in\{1a,1b,2a,2b\}} p_jU(x_j)$ evaluates prospects. The probabilities and utilities can be derived from preferences and, hence, are often called *subjective*, and so is this EU model.

We will also consider the state-dependent generalization of EU, where U can depend on events. Then, it is well-known that, without further assumptions, utilities and probabilities cannot be separated (Kreps 1988 Eqs. 4.4 and 7.13). We therefore use general functions V_j to generalize the products p_jU. *State-dependent expected utility (EU)* holds if there exist continuous strictly increasing *additive value functions* $V_j : \mathbb{R} \to \mathbb{R}, j = 1, \ldots, n$ such that $x \mapsto \sum_{j \in \{1a,1b,2a,2b\}} V_j(x_j)$ evaluates prospects. An alternative term for the latter function is *additive(ly decomposable) representation*. We have (Debreu 1960):

Theorem 3.1. *State-dependent EU holds if and only if separability holds. The additive value functions are unique up to location (we can add a constant to each) and a joint unit (we can multiply them all by the same positive factor).*

The theorem shows that separability is not enough to give EU (because the V_js need not be proportional). Savage (1954), in a somewhat different set-up, added a likelihood consistency condition (his P4) to obtain EU. Other authors added other conditions for the same purpose.[2]

3.3 Independence

The term independence has been used in many different meanings in decision theory. Our term concerns a preference version of statistical independence, and it can be called informational independence or i-independence. Because no other concepts of independence appear in this paper, and no confusion will arise, we use the short term independence throughout.

3.3.1 Independence; general introduction

Independence of an uncertain event E from a conditioning event C means, informally, that C carries no information about E. In traditional probability theory, it means that conditioning on C does not affect the probability of E.

[2]See Abdellaoui & Wakker (2005), Chew & Karni (1994), Gul (1992), Köbberling & Wakker (2003), Nakamura 1995).

Independence of E is always equivalent to independence of its complement E^c, both in traditional theories and in the generalized theories considered in this paper. In traditional theories, independence is also *symmetric* in E and C: If E is independent from C, then so is C from E. This symmetry will have to be abandoned in generalized theories.

In preference theory, independence means that preferences contingent on E are unaffected by conditioning on C. A particular independence is implicitly assumed in the AA framework, where it causes some problems as we will see later. For obtaining our results, the main mathematical tools are theorems by Gorman (1968), van Daal & Merkies (1988), and Mongin & Pivato (2015). Our paper shows how these mathematical results, and some other results in the literature, are related to the concept of statistical independence.

3.3.2 Independence related informally to dynamic updating

Our formal analysis will define all concepts in terms of static, "present," preferences, and we will investigate the implications and restrictions of independence for static theories. Independence is often interpreted by referring to dynamic settings, involving future updating. This subsection, which plays no role in the formal analysis, explains the relations of such definitions with what we do.

In studies on future updating it is typically assumed, for instance, that the information will be received that C_a is true, ruling out the events E_1C_b and E_2C_b. Then preferences are determined, for instance between a conditional prospect $(E_1C_a : \alpha, E_2C_a : \beta)$, yielding α under E_1C_a and β under E_2C_a, and another conditional prospect $(E_1C_a : \gamma, E_2C_a : \delta)$. It is then assumed that the outcomes under E_1C_b and E_2C_b, counterfactual as they are by then, are no more relevant and can be ignored. The latter assumption is a special case of Machina's (1989) consequentialism (applied only to E_1). For updated preferences to be relevant to present preferences, it is desirable that the mentioned outcomes then also be irrelevant under present preferences. This assumption is a special case of Machina's (1989) dynamic consistency. Separability of the conditioning event C_a follows. Bernardo, Ferrandiz, & Smith (1985 Definition 3 and Axiom 3(iii)) similarly defined conditional preferences, imposing separability.

Separability is a restrictive assumption that has often been discussed. When imposed on one event, as is done here, the condition is not very restrictive. When imposed on all single elements of a partition of the universal event (e.g., $\{C_1, C_2\}$), as will be done later, the condition still is not very restrictive. It then becomes what is called weak separability in consumer theory (Blackorby, Primont & Russell 1978 pp. 42-60), amounting to a kind of monotonicity, and it can then be accommodated by many nonexpected utility theories. The condition becomes restrictive, capturing much of expected utility, only when imposed on overlapping composite events, mainly by Gorman's (1968) powerful result (Theorem 8 in the appendix).

Too much of the spirit of statistical independence is lost if separability of the conditioning event is given up. If the preference in Eq. 3.2 depends on c_1 and c_2 (so, can be different from Eq. 3.3), then there can be no clear relation between updated preference and present preference. Hence all versions of independence considered later will imply this separability. We will investigate to what extent it is possible to still have such independence for static nonEU theories. Thus, our formal analysis does not consider dynamic aspects.

3.3.3 Independence defined formally

The following two versions of independence are equivalent under traditional theories and, therefore, are usually not distinguished. In our general setting, they can be different. The first version, also used by Bernardo, Ferrandiz, & Smith (1985 Definition 4), considers E_1-preferences. E_1 is *weakly (informationally) independent* of C_a if:

$$\begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \alpha & \alpha \\ E_2 & \beta & \beta \end{array} \quad \succcurlyeq \quad \begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \gamma & \gamma \\ E_2 & \delta & \delta \end{array} \tag{3.4}$$

if and only if

$$\begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \alpha & c_1 \\ E_2 & \beta & c_2 \end{array} \quad \succcurlyeq \quad \begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \gamma & c_1 \\ E_2 & \delta & c_2 \end{array} \tag{3.5}$$

for all outcomes considered.

The condition implies that for E_1-preferences it does not matter whether or not the information that C_a is true is received. The condition directly appeals to intuitions of noninformativeness, and can serve well as a primitive in empirical preference assessments. This paper focuses on theoretical results. As explained before, the above condition implies separability of C_a. It adds further restrictions as we will see later.

The second version of independence is stronger. It implies that it does not matter for E_1-preferences whether the received information is that C_a is true or that C_a is not true. E_1 is *complement-independent*, or *independent* for short, of C_a if:

$$\begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \alpha & c_1 \\ E_2 & \beta & c_2 \end{array} \quad \succcurlyeq \quad \begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & \gamma & c_1 \\ E_2 & \delta & c_2 \end{array} \tag{3.6}$$

if and only if

$$\begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & c'_1 & \alpha \\ E_2 & c'_2 & \beta \end{array} \quad \succcurlyeq \quad \begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & c'_1 & \gamma \\ E_2 & c'_2 & \delta \end{array} \tag{3.7}$$

for all outcomes considered.

It is always understood in independence conditions that the conditioning event plays the role of C_a above. Both definitions are symmetric in E_1 and E_2. Thus, (weak) independence of E_1 from C_a is equivalent to (weak) independence of E_2 from C_a. Independence is also symmetric in C_a and C_b. As discussed before, neither version of independence needs to be symmetric in E_1 and C_a.

OBSERVATION 3.1: *Independence implies weak independence, and separability of the conditioning event and its complement. Weak independence implies separability of the conditioning event.*

Example 10 in the appendix shows that weak independence in general is weaker than independence.

3.4 Independence with separability

This section considers some implications of independence, and the possibility to have independence outside of EU. The following theorem, which is our first main result, gives a foundation of EU using independence. From the perspective of using independence outside of EU, this theorem is a negative result because it excludes non-EU. However, from a more positive perspective, giving a foundation of EU entirely in terms of statistical independence is useful. As it turns out, independence with symmetry implies not only separability (Savage's 1954 sure-thing principle), but also Savage's likelihood consistency P4, which is the other main condition that Savage used to derive EU.

Theorem 3.2. [3] *The following two statements are equivalent:*

1. *E_1 is independent of C_a and C_a is independent of E_1.*
2. *Expected utility holds. Writing $P(E_1C_a) = p_{1a}$, $P(E_1) = p_{1a} + p_{1b}$, and $P(C_a) = p_{1a} + p_{2a}$, we further have $P(E_1C_a) = P(E_1) \times P(C_a)$.*[4]

From a positive perspective on Theorem 3.2, the traditional independence conditions, when stated as preference conditions, give not only the multiplicative form of probabilities, but the whole EU model itself. From a negative perspective, some of the traditional properties of independence will have to be abandoned in generalized models. This is the topic of the rest of this paper.

[3]Bernardo, Ferrandiz, & Smith (1985) also derived expected utility from independence but used a rich structure. It did not only comprise Savage's (1954) state space and axioms, but also an AA richness through their Axiom 4. They required existence of independent random events with any objective probability. Mongin & Pivato (2015) also used the theorem of aggregation following from Gorman (1968), and then imposed an equal-ordering condition of conditionals similar to independence. This equal-ordering condition was discussed by van Daal & Merkies (1988) as homogeneity of individuals (= rows). Mongin & Pivato (2015) consider a more general mathematical setup with more than four events and with non-Euclidean outcomes, but do not relate their conditions to statistical independence.

[4]This implies the usual $p_{1b} = P(E_1) \times (1 - P(C_a))$, $p_{2a} = (1 - P(E_1)) \times P(C_a)$, and $p_{2b} = (1 - P(E_1)) \times (1 - P(C_a))$.

In the rest of this section, we maintain separability, that is, we assume the state-dependent generalization of EU. A classical problem in this model is the impossibility of identifying probability and, then, of finding out what plausible assumptions can serve to identify probabilities after all (Drèze 1987; Karni 1996, 2013; Kadane & Winkler 1988; Schervish, Seidenfeld, & Kadane, 1990; Nau, 1995). The following theorem provides a new way, showing that statistical independence delivers probabilities for the conditioning events. The result can be interpreted as negative in the sense that utility can no longer depend on that conditioning event. Thus, having probabilities and generating state dependent utility still does not go together for the same event.

Theorem 3.3 (Independence and separability)**.** *Assume separability. Then the following three statements are equivalent:*

1. *E_1 is weakly independent of C_a.*

2. *E_1 is independent of C_a.*

3. *There exists probabilities $P(C_a)$ and $P(C_b) = 1 - P(C_a)$, and continuous strictly increasing functions V_1 and V_2 such that prospects are evaluated by*

$$\begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & x_1 & y_1 \\ E_2 & x_2 & y_2 \end{array} \mapsto P(C_a)(V_1(x_1) + V_2(x_2)) + P(C_b)(V_1(y_1) + V_2(y_2)). \tag{3.8}$$

Thus, we get an EU representation for C_a-prospects with probabilities $P(C_a)$ and $P(C_b)$, and a utility function $U(\alpha) = V_1(\alpha) + V_2(\alpha)$. Apparently, some space is left for non-EU, but not much, and only for events that can never play a role as conditioning events.

3.5 Independence without separability

We now turn to general nonexpected utility models, where there is no separability other than what is implied by independence. We will provide a representation theorem for rank-dependent utility *(RDU)*. Implications for other nonexpected utility models are left as a topic for future research. In particular, it is an open question to us whether, under general nonexpected utility, weak independence of E_1 from C_a and of C_a from E_1, i.e., symmetric weak independence, is possible, or whether these assumptions are already enough to imply EU. The latter result would generalize our Theorem 3.2. We will now see that under RDU at least, symmetric weak independence is not possible outside of EU.

RDU generalizes EU by using a *weighting function* W defined on events, satisfying $W(\emptyset) = 0, W(S) = 1$, and $A \supset B \Rightarrow W(A) \geq W(B)$. W is allowed to be nonadditive. U is again a continuous strictly increasing *utility function*. To evaluate a prospect $x = (x_{1a}, x_{1b}, x_{2a}, x_{2b})$, we first rank-order the outcomes from best to worst. That is, we take $\rho : \{1, 2, 3, 4\} \to S$ such that $x_{\rho(1)} \geq \cdots \geq x_{\rho(4)}$. Then x is evaluated by

$$RDU(x) = \sum_{j=1}^{4} \pi_{\rho(j)} U(x_{\rho(j)}) \tag{3.9}$$

where the $\pi_{\rho(j)}$s are positive weights adding to 1, defined by

$$\pi_{\rho(j)} = W(\{\rho(1), \ldots, \rho(j)\}) - W(\{\rho(1), \ldots, \rho(j-1)\}). \tag{3.10}$$

Here $\pi_{\rho(1)} = W(\{\rho_1\})$. RDU holds whenever a weighting function W and a utility function U exist that give the evaluation just described. The following theorem[5] is remarkably similar to Theorem 3.3.

[5]Sarin & Wakker (1998 Theorem 3.1) derived a conclusion similar to Theorem 3.4

Theorem 3.4 (Independence and no separability)**.** *Assume RDU. Then the following three statements are equivalent:*

1. *E_1 is weakly independent of C_a.*

2. *E_1 is independent of C_a.*

3. *There exists probabilities $P(C_a) = W(C_a)$ and $P(C_b) = 1 - P(C_a) = W(C_b)$, such that prospects are evaluated by*

$$\begin{array}{c|cc} & C_a & C_b \\ \hline E_1 & x_1 & y_1 \\ E_2 & x_2 & y_2 \end{array} \rightarrow P(C_a)RDU(x_{1_{E_1}}x_2) + P(C_b)RDU(y_{1_{E_1}}y_2). \quad (3.11)$$

3.6 Implications for the Anscombe-Aumann framework

In the AA framework there are two kinds of events. First, there are complementary ambiguous events H_a and H_b called *horse events*. Say they describe the winner of an upcoming horse race between two horses, an American and a Belgian. These events are of most interest to us. Then there are complementary risky events R_1, R_2 called *roulette events*, for which probabilities are given. They may describe the result of a spin of a roulette wheel. They play an auxiliary role, serving to clarify the analysis of the ambiguous events.

R_1-prospects are evaluated by EU. H_a-prospects $\alpha_{H_a}\beta$ are evaluated by an ambiguity functional $V(\alpha, \beta)$ that can take different forms, depending on

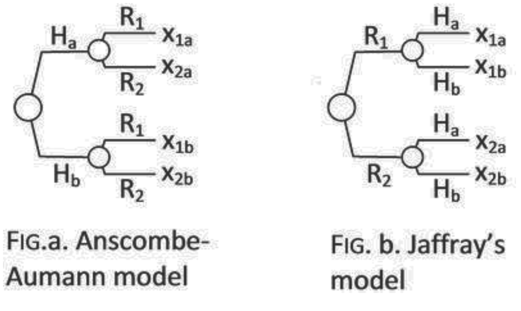
(EU in the first stage and not in the second) from different assumptions. Their main assumption was sequential consistency, implying that general models, such as RDU, used to evaluate general prospects, should also be used for evaluating conditional prospects in subtrees. Independence was implicitly implied by backward induction in their approach. Theorem 3.4 only assumes the overall model and independence.

applications and interests. Schmeidler (1989) considered an RDU functional V, and Gilboa & Schmeidler (1989) considered a maxmin multiple priors functional V. Many other functionals have been considered (surveyed by Etner, Jeleva, & Tallon 2012 and Trautmann & Wakker 2015).

General prospects, depending jointly on H and R events, are evaluated by

$$\begin{array}{c|cc} & H_a & H_b \\ \hline R_1 & x_1 & y_1 \\ R_2 & x_2 & y_2 \end{array} \mapsto V(EU(x_{1_{R_1}} x_2), EU(y_{1_{R_1}} y_2)). \tag{3.12}$$

Figure 3.1: Different orderings of events



That is, we first take EU conditional on H_a and H_b, and then apply the functional V.[6] The evaluation is usually justified by temporal assumptions (Figure 3.1.a). It is assumed that first the uncertainty about the horse race (event H_a or event H_b) is resolved, and then the roulette wheel is spun (event R_1 or R_2). Next backward induction is applied: First the R_1-prospects are

[6] Or we apply $V \circ U^{-1}$ where U^{-1} gives certainty equivalents, if V is to be applied to outcomes.

replaced by their expected utilities, and then the ambiguity functional is applied to the resulting H_a-prospect.

The evaluation in Eq. 3.12 is appropriate if R_1 and R_2 can be taken to be independent of H_a and H_b. Thus, the ambiguous events H_a and H_b play the role of conditioning events here. A comparison with Theorems 3.2 and 3.4 suggests problems. These theorems show that events with probabilities are best suited to play the role of conditioning events. In the AA framework the opposite happens, and the H events, which have no probabilities, play the role of conditioning events. Our theorems suggest that it is preferable to condition on R_1 and R_2, rather than on H_a and H_b. In other words, the approach displayed in Fig. 1b seems to be more suited. Theorem 3.4 shows in particular that in Schmeidler's (1989) derivation of RDU, the order of stages assumed in the AA framework is problematic.

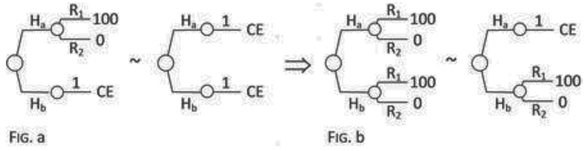
Formally, the above claims were proved only for state-dependent EU and RDU. We next show that the problems of the AA framework hold for general ambiguity models, by using Observation 3.1. This observation shows that the conditioning events—the ambiguous H events in the common AA framework—should be separable. Although, as explained before, separability of the elements of a partition, as with $\{H_1, H_2\}$ in the AA framework, may not be very restrictive, its restrictions still are incompatible with ambiguity. This point is explained by the following paradox. The paradox is a variation on an informal example by Wakker (2010, Figure 10.7.1), modified using the formal concepts of our paper.

Example 3.1: [The separability paradox for the AA framework; see Figure 3.2]

Assume that R_1 and R_2 have probability 0.5, and assume indifference between CE for sure and $100_{R_1}0$. Thus, CE removes the risk in a preference-neutral manner. Under risk aversion, $CE \leq 50$. Readers can substitute their own CE here. The AA framework implies the indifferences in Figures 3.2.a and 3.2.b. R_1 and R_2 are treated independently of H_a and H_b, as in Equations 3.4 and 3.5: conditioning on H_a does not affect R_1-prospects. Events H_a and H_b must be separable, as in Equation 3.3. However, these implied indifferences are not plausible. Ambiguity aversion will generate interactions between H_a and H_b, based on global properties of prospects, as explained next. For ambiguity averse readers who do not apply the AA framework mechanically, but follow their intuitive preferences, we predict the following preferences, deviating from the AA implications.

In Figure 3.2.a, replacing the upper left lottery by its CE removes the risk in a preference-neutral manner. But it does more: it removes all ambiguity. In the left prospect, all outcomes are ambiguous. Under both events R_1 and R_2, there is ambiguity about the resulting outcome. The decision maker observes the disappearance of ambiguity due to the CE substitution by comparing what happens under event H_a with what happens under event H_b, entailing a global observation of the prospects. Such observations go against the separable evaluation of H_a and H_b assumed in the AA framework. The disappearance of not only risk but also ambiguity generates extra preference for the right prospect in Figure 3.2.a, which will be strictly preferred to the left prospect.

Figure 3.2: The separability paradox for the Anscombe-Aumann model



The left prospect in Figure 3.2.b yields a fifty-fifty lottery irrespective of the H events, and the outcomes have known probabilities (0.5). Hence this prospect is unambiguous. Neither under R_1 nor under R_2 is there ambiguity. The CE substitution in the upper branch does not only remove risk, which by itself happens in a preference-neutral manner. It has an extra effect: it generates ambiguity. In the right prospect, all outcomes are ambiguous (exactly as with the left prospect in Figure 3.2.a). Ambiguity aversion generates negative preference for the right prospect, leading to a strict preference for the left prospect.

Formally, the indifference that defines CE together with a strict preference in Figure 3.2.a (and similarly with a strict preference in Fig. 2b) implies that R_1 is not weakly independent of H_a, invalidating backwards induction. The reversal of strict preference from Figure 3.2.a to Figure 3.2.b by itself also directly violates weak independence.

Ambiguity means almost by definition that the separability of EU is violated, with interactions occurring between disjoint events. This was the main lesson of the Ellsberg paradoxes, where it crucially depends on outcomes under other events to decide whether a substitution conditional on some events generates or removes ambiguity. Global inspections of prospects are relevant, as it happens in our example with the disjoint events H_a and H_b.

The preceding example showed that the separable treatment of the H-events in the AA framework is inconsistent with their ambiguity. Jaffray (personal communication related to Jaffray & Wakker 1993), who considered EU to be appropriate for risk but not for ambiguity, emphasized that conditioning should be done only with respect to unambiguous events, a principle satisfied in all his ambiguity papers (Jaffray 1988, 1989a, 1989b, 1990, 1991a, b, 1994). He thus adopted the framework depicted in Figure 3.1.a rather than the one in Figure 3.1.b. It leads to the following evaluation:

$$\begin{array}{c|cc} & H_a & H_b \\ \hline R_1 & x_1 & y_1 \\ R_2 & x_2 & y_2 \end{array} \mapsto P(R_1)V((x_{1_{H_a}} y_1) + P(R_2)V((x_{2_{H_a}} y_2), \tag{3.13}$$

where V can be any ambiguity functional of interest.

Currently, Figure 3.1.a is very popular-to the extent that it is almost the only used way to model ambiguity-because of its mathematical convenience. For example, Machina (2012), when analyzing mixed bets (prospects) with both risky and ambiguous events being outcome-relevant, only considers evaluations using Fig. 1a, even though there is no temporal ordering of the events.

Figure 3.1.b also provides convenient mathematical tools. The mathematics of Eq. 3.13 has been extensively studied in multiattribute utility theory, where the pairs (x_1, y_1) and (x_2, y_2) designate commodity bundles (multiattribute outcomes). Many kinds of interactions between the x and y coordinates have been developed and axiomatize, in Keeney & Raiffa (1976) and numerous follow-up papers. Such techniques can be used to capture ambiguity attitudes. Jaffray's models illustrate this point. Recent papers that used this approach, and that were inspired by Jaffray's work, include Gul & Pesendorfer (2014) and Olszewski (2007).

3.7 Conclusion

We have stated preference conditions capturing independence in a statistical sense, and have examined independence in various decision models. A symmetry condition, routinely assumed in the traditional Bayesian model, turns out to imply this Bayesian model. From a positive perspective, this symmetry gives a new foundation to the Bayesian model, but from a negative perspective, this symmetry cannot be used in other models. Nonsymmetric independencies can be applied to non-Bayesian (ambiguity) models, where we derive their implications. In particular, these implications reveal a problem for the Anscombe-Aumann framework, which is very popular in the ambiguity literature today. This problem can be avoided by reversing the order of stages in this framework.

3.8 Appendices

Appendix 3.A: Preparations for proofs

As a preparation for the proofs, we present Gorman's (1968) powerful theorem.

Theorem 3.5 (Gorman's theorem)**.** *Assume that A and B are two separable events with nontrivial overlap ($A \cap B, A \cap B^c$, and $A^c \cap B$ are nonempty). Then $A \cap B$, $A \cap B^c$, $A^c \cap B$, and all their unions, are also separable.*

The following corollary follows from repeated application of Gorman's result. It is a special case of a well-known implication of Gorman's theorem, sometimes called the theorem of aggregation: If prospects can be represented in a matrix, as in Eq. 3.1, and all rows and all columns are separable, then separability holds. This result has a long history (Nataf 1948; reviewed by Aczl 1997 and van Daal & Merkies 1988). For example, if rows refer to individuals and columns to production inputs, then under separability of rows and columns, first aggregating over individuals (micro) can be equivalent to first aggregating production inputs (macro).

Corollary 3.1. *Separability holds as soon as three of the four events E_1, E_2, H_a, H_b are separable.*

Proof. $s_1 = E_1H_a$, $s_2 = E_1H_b$, $s_3 = E_2H_a$, $s_4 = E_2H_b$. Assume that $\{s_1, s_2\}$, $\{s_3, s_4\}$, and $\{s_1, s_3\}$ are separable. Monotonicity implies separability of all four s_j; this could also be derived from Gorman's theorem applied to $\{s_1, s_2\}$ and $\{s_1, s_3\}$, and then to $\{s_3, s_4\}$ and $\{s_1, s_3\}$. For each newly derived separable set below, we indicate the two preceding sets that, through Gorman's theorem, imply separability of it.

$$\{s_1, s_2\}, \{s_2, s_3\} \Rightarrow \{s_1, s_2, s_3\};$$

$$\{s_3, s_4\}, \{s_2, s_3\} \Rightarrow \{s_2, s_3, s_4\};$$

$$\{s_1, s_2\}, \{s_2, s_3, s_4\} \Rightarrow \{s_1, s_3, s_4\};$$

$$\{s_3, s_4\}, \{s_1, s_2, s_3\} \Rightarrow \{s_1, s_2, s_4\}.$$

Now that we have all three-element events separable, we can get all two-element events from intersections of the proper three-element events. Thus, we have separability of all events. □

The following concepts will be used in several proofs. We call μ the *(unconditional) certainty equivalent* (*CE*) of the E_1-prospect $\alpha_{E_1}\beta$ if $\mu_{E_1}\mu \sim \alpha_{E_1}\beta$. By continuity and monotonicity, this CE always exists and is unique. We next consider $\alpha_{E_1}\beta$ conditioned on C_a, with common (or counterfactual) outcomes c_{1b}, c_{2b}. We call μ the *conditional certainty equivalent* (*CCE*) of (α, β) under C_a if C_a is separable, and $(\alpha_{1a}, c_{1b}, \beta_{2a}, c_{2b}) \sim (\mu_{1a}, c_{1b}, \mu_{2a}, c_{2b})$. By separability of C_a, μ is independent of c_{1b} and c_{2b} and, hence, it is well defined. By continuity and monotonicity, CCEs again always exist and are unique. CCEs under C_b, or under other events, are defined analogously. Because of continuity and monotonicity, CEs of E_1-prospects completely determine preferences over these prospects, and CCEs completely determine conditional preferences. Independence means that CCEs under C_a and C_b are the same. Weak independence means that conditional CEs under C_a agree with unconditional CEs of E_1 prospects.

The following example shows that, for general preferences, weak independence is weaker than independence.

Example 3.2: Assume that prospects $(x_{1a}, x_{1b}, x_{2a}, x_{2b})$ are evaluated by

$$\frac{x_{1a}+x_{2a}+x_{1b}+x_{2b}}{4}+f((x_{1a}+x_{2a})-(x_{1b}+x_{2b}))\times g(x_{1b}), \tag{3.14}$$

where both f and g are 0 on $\mathbb{R}^-$, are strictly increasing on $\mathbb{R}^+$, have both their values and derivatives always below 0.25, and are continuous. For example, on $\mathbb{R}^+$, we may take $f(\alpha) = g(\alpha) = (1 - exp((\alpha)/4))/4$. The preference relation generated by this evaluation function is, by definition, transitive and complete. It is also continuous. The first term in Eq. 3.14 is strictly increasing in all four outcomes with derivatives 0.25, and is also strictly increasing in $x_{1a} + x_{2a}$ with derivative 0.25. The second term in Eq. 3.14 is nondecreasing in x_{1a}, x_{2a}, and $x_{1a} + x_{2a}$. Thus, monotonicity with respect to x_{1a}, x_{2a}, and their sum follows. The latter monotonicity implies separability of C_a. Although the second term in Eq. 3.14 may sometimes be decreasing in x_{1b} and x_{2b}, the derivative is never below $-1/8$ for x_{1b} and never below $-1/4$ for x_{2b}, and the first term ensures monotonicity also in x_{1b} and x_{2b}. For all E_1-prospects, the second term vanishes, and E_1-prospects are evaluated by expected value, as are preferences with x_{1b} and x_{2b} kept fixed due to monotonicity in $x_{1a} + x_{1b}$. Hence weak independence holds. However, $(2, 0, 2, 2) \sim (0, 2, 2, 2)$ and $(2, 2, 2, 0) \succ (2, 2, 0, 2)$ show that independence is violated.

Appendix 3.B: Proofs

PROOF OF OBSERVATION 3.1: Independence implies that the preferences in Eqs. 3.6 and 3.7 are independent of c_1, c_2, c_1' and c_2', implying separability of both C_a and C_b. Independence then implies that the CCE μ of (α, β) is the same under C_a as under C_b. We then have $(\alpha_{1a}, \alpha_{1b}, \beta_{2a}, \beta_{2b}) \sim (\mu_{1a}, \alpha_{1b}, \mu_{2a}, \beta_{2b}) \sim (\mu_{1a}, \mu_{1b}, \mu_{2a}, \mu_{2b})$. That is, μ, the unconditional CE of $(\alpha_{1a}, \alpha_{1b}, \beta_{2a}, \beta_{2b})$, is identical to the two conditional CEs mentioned. This implies weak independence of E_1 with respect to both C_a and C_b.

Weak independence implies that the preference in Eq. 3.5 is independent of c_1 and c_2, which implies separability of the conditioning event C_a. $\square$

PROOF OF THEOREM 3.2: Statement 1. readily follows from Statement 2. Hence we assume 1, and derive 2. The independence assumptions imply separability of E_1, E_2, C_a, C_b. Corollary 3.1 implies separability. Theorem 3.1 implies state-dependent EU:

$$V_{1a}(x_{1a}) + V_{1b}(x_{1b}) + V_{2a}(x_{2a}) + V_{2b}(x_{2b})$$

By the uniqueness results, we may assume that all V functions are 0 at 0. Independence of E_1 from C_a implies that $V_{1a}(\alpha) + V_{2a}(\beta)$ represents the same preference relation over $\mathbb{R}^2$, the set of all pairs (α, β), as does $V_{1b}(\alpha) + V_{2b}(\beta)$ (with the same CCEs). The uniqueness results (that similarly hold for additive representations on $\mathbb{R}^2$) imply that there exists a $\lambda_{ba} > 0$ such that $V_{1b} = \lambda_{ba} V_{1a}$ and $V_{2b} = \lambda_{ba} V_{2a}$.

Independence of C_a from E_1 implies that $V_{1a}(\alpha) + V_{1b}(\beta)$ represents the same preference relation over $\mathbb{R}^2 = \{(\alpha, \beta)\}$ as does $V_{2a}(\alpha) + V_{2b}(\beta)$. By the uniqueness results on $\mathbb{R}^2$, there exists a $\lambda_{21} > 0$ such that $V_{2a} = \lambda_{21} V_{1a}$ and $V_{2b} = \lambda_{21} V_{1b}$. We define $U = (1 + \lambda_{ba} + \lambda_{21} + \lambda_{ba}\lambda_{21}) \times V_1$, $p_{1a} =$

$\frac{1}{1+\lambda_{ba}+\lambda_{21}+\lambda_{ba}\lambda_{21}}$, $p_{1b} = \frac{\lambda_{ba}}{1+\lambda_{ba}+\lambda_{21}+\lambda_{ba}\lambda_{21}}$, $p_{2a} = \frac{\lambda_{21}}{1+\lambda_{ba}+\lambda_{21}+\lambda_{ba}\lambda_{21}}$, and $p_{2b} = \frac{\lambda_{21}\lambda_{ba}}{1+\lambda_{ba}+\lambda_{21}+\lambda_{ba}\lambda_{21}}$. This gives the representation in Statement 2. □

PROOF OF THEOREM 3.3. Assume Statement 1., with weak independence of E_1 from C_a. Let the CCE of (α, β) under C_a be μ. Then so is the unconditional CCE of the E_1 prospect $(\alpha_{1a}, \beta_{1b}, \alpha_{2a}, \beta_{2b})$. We have:

$$(\mu_{1a}, \alpha_{1b}, \mu_{2a}, \beta_{2b}) \sim (\alpha_{1a}, \beta_{1b}, \alpha_{2a}, \beta_{2b}) \sim (\mu_{1a}, \mu_{1b}, \mu_{2a}, \mu_{2b}).$$

Indifference between the first and third prospect shows that the CCE of (α, β) under C_b must also be μ. Statement 2. follows.

Assume Statement 2. As in the proof of Theorem 3.2, independence of E_1 from C_a implies for the additive value functions that $V_{1a}(\alpha) + V_{2a}(\beta)$ represents the same preference relation over $\mathbb{R}^2 = \{(\alpha, \beta)\}$ as does $V_{1b}(\alpha) + V_{2b}(\beta)$. Assuming that all Vs are 0 at 0, this implies existence of $\lambda_{ba} > 0$ such that $V_{1b} = \lambda_{ba} V_{1a}$ and $V_{2b} = \lambda_{ba} V_{2a}$. Define $P(C_a) = \frac{1}{1+\lambda_{ba}}$, $P(C_b) = \frac{\lambda_{ba}}{1+\lambda_{ba}}$, and $V_j = (1 + \lambda_{ba})V_{ja}$ for $j = 1, 2$. Adapting notation, Statement 3. follows.

It readily follows that Statement 3. implies the other two statements, for instance because CCEs under C_a and C_b, and unconditional CCEs, all agree. Thus, all statements are equivalent. □

PROOF OF THEOREM 3.4. Statement 3. readily implies Statement 2., which readily implies Statement 1. Hence we assume Statement 1., with weak independence of E_1 from C_a, and derive Statement 3.

We first show that the decision weights of E_1C_a and E_2C_a are not affected by the ranking positions of these events with respect to E_1C_b and E_2C_b, and vice versa. We focus on the case where $x_{1a} \geq x_{2a}$. By weak independence, the preferences over these pairs of outcomes are the same irrespective of where we fix the outcomes under C_b. This means that the proportion of the decision weights of x_{1a} and x_{2a} are unaffected by the ranking positions of the other two outcomes (Wakker 1993). If we interchange two adjacent ranking positions of one outcome from x_{1a}, x_{2a} and one of the other two outcomes, then only the decision weights of these two outcomes can change, but because of the aforementioned constant proportions, the outcome from x_{1a}, x_{2a} cannot change its weight. Because decision weights of a prospect always add up to 1, neither of the two decision weights can then change. Hence changes in rankings between subevents of C_a and C_b never affect decision weights.

Changes in rankings of x_{1a} and x_{2a} do not affect the decision weights of the other two outcomes and, hence, do not affect the sum of decision weights of the other two outcomes and, hence, of x_{1a} and x_{2a} themselves. Write $P(C_a)$ $(= W(C_a))$ for the former sum, and, similarly, $P(C_b)$ $(= W(C_b))$. The RDU functional decomposes into a weighted sum of two RDU functionals $P(C_a)RDU'(x_{1a}, x_{2a}) + P(C_b)RDU''(x_{1b}, x_{2b})$. By weak independence, $P(C_a)RDU'(\alpha, \beta) + P(C_b)RDU''(\alpha, \beta)$ and $RDU'(\alpha, \beta)$ represent the same preference over pairs (α, β) and, hence, can be taken proportional. It implies that RDU is a weighted sum of the same RDU functionals, as in Statement 3. $\square$

3.9 References

Abdellaoui, Mohammed & Peter P. Wakker 2005. The Likelihood Method for Decision under Uncertainty. Theory and Decision 58 3-76.

Aczél, János, 1997. Bisymmetry and Consistent Aggregation: Historical Review and Recent Results. In. Anthony A.J. Marley (Ed.), Choice, Decision, and Measurement: Essays in Honor of R. Duncan Luce, 225-233, Lawrence Erlbaum Associates, Mahwah, N.J.

Anscombe, Francis J. & Robert J. Aumann 1963. A Definition of Subjective Probability. Annals of Mathematical Statistics. 34, 199-205.

Bernardo, Jose M., Juan R. Ferrandiz, & Adrian F.M. Smith 1985. The Foundations of Decision Theory: An Intuitive, Operational Approach with Mathematical Extensions. Theory and Decision. 19, 127-150.

Blackorby, Charles, Daniel Primont, & Robert R. Russell 1978. Duality, Separability and Functional Structure: Theory and Economic Applications. North-Holland, Amsterdam.

Chew, Soo Hong & Edi Karni 1994. Choquet Expected Utility with a Finite State Space: Commutativity and Act-Independence. Journal of Economic Theory. 62, 469-479.

Debreu, Gérard 1960. Topological Methods in Cardinal Utility Theory. In. Kenneth J. Arrow, S. Karlin, & Patrick Suppes 1959 Mathematical Methods in the Social Sciences., 16-26, Stanford University Press, Stanford, CA.

Drèze, Jacques H. 1987. Essays on Economic Decision under Uncertainty. Cambridge University Press, London.

Etner, Johanna, Meglena Jeleva, & Jean-Marc Tallon 2012. Decision Theory under Ambiguity. Journal of Economic Surveys. 26, 234-270.

Gilboa, Itzhak & David Schmeidler 1989. Maxmin Expected Utility with a Non-Unique Prior. Journal of Mathematical Economics. 18, 141-153.

Glynn, Luke 2011. A Probabilistic Analysis of Causation. The British Journal for the Philosophy of Science. 62, 343-392.

Gorman, William M. 1968. The Structure of Utility Functions. Review of Economic Studies. 35, 367-390.

Gul, Faruk 1992. Savage's Theorem with a Finite Number of States, Journal of Economic Theory. 57, 99-110. (.Erratum, Journal of Economic Theory. 61, 1993, 184.)

Gul, Faruk & Wolfgang Pesendorfer 2014 .Expected Uncertain Utility Theory. Econometrica 59, 1273-1292.

Halpern, Joseph Y. 2003. Reasoning about Uncertainty. The MIT Press, Cambridge, MA.

Harper, William, Sheldon J. Chow, & Gemma Murray 2012. Bayesian Chance, Synthese. 186, 447-474.

Jaffray, Jean-Yves 1988. Applications of Linear Utility Theory to Belief Functions. In. Bernadette Bouchon & Ronald R. Yager (eds.) (eds.) Uncertainty and Intelligent Systems., Springer, Berlin.

Jaffray, Jean-Yves 1989a. Linear Utility Theory for Belief Functions, Operations Research Letters. 8, 107-112.

Jaffray, Jean-Yves 1989b. Coherent Bets under Partially Resolving Uncertainty and Belief Functions. Theory and Decision. 26, 99-105.

Jaffray, Jean-Yves 1990. Bayesian Updating and Belief Functions. Proceedings of the 3rd International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU90, Paris, July 1990), 449-451 (published by ENSTA, Paris).

Jaffray, Jean-Yves 1991a. Belief Functions, Convex Capacities, and Decision Making. In. Jean-Paul Doignon & Jean-Claude Falmagne (eds.) Mathematical Psychology: Current Developments., 127-134, Springer, Berlin.

Jaffray, Jean-Yves 1991b. Linear Utility Theory and Belief Functions: A Discussion. In. Atilla Chikan (ed.) Progress in Decision, Utility and Risk Theory. Kluwer Academic Publishers, Dordrecht.

Jaffray, Jean-Yves 1994. Dynamic Decision Making with Belief Functions. In. Ronald R. Yager, Mario Fedrizzi, & Janus Kacprzyk (eds.) Advances in the Dempster-Shafer Theory of Evidence,. 331-352, Wiley, New York.

Jaffray, Jean-Yves & Peter P. Wakker 1993. Decision Making with Belief Functions: Compatibility and Incompatibility with the Sure-Thing Principle. Journal of Risk and Uncertainty. 7, 255-271.

Jensen, Finn V. & Thomas Dyhre Nielsen 2013. Probabilistic Decision Graphs for Optimization under Uncertainty. Annals of Operations Research. 204, 223-248.

Kadane, Joseph B. & Robert L. Winkler 1988. Separating Probability Elicitation from Utilities. Journal of the American Statistical Association. 83, 357-363.

Karni, Edi 1996. Probabilities and Beliefs. Journal of Risk and Uncertainty. 13, 249-262.

Karni, Edi 2013. Bayesian Decision Making with Action-Dependent Probabilities and Risk Attitudes. Economic Theory. 53, 335-356.

Keeney, Ralph L. & Howard Raiffa 1976. Decisions with Multiple Objectives. Wiley, New York. (Second edition 1993, Cambridge University Press, Cambridge).

Köbberling, Veronika & Peter P. Wakker 2003. Preference Foundations for Nonexpected Utility: A Generalized and Simplified Technique. Mathematics of Operations Research. 28, 395-423.

Kreps, David M. 1988. Notes on the Theory of Choice. Westview Press, Boulder Colorada.

Machina, Mark J. 1989. Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty. Journal of Economic Literature. 27, 1622-1688.

Machina, Mark J. 2012. Ambiguity Aversion with Three or More Outcomes, mimeo.

Mongin, Philippe & Marcus Pivato 2015. Ranking Multidimensional Alternatives and Uncertain Prospects. Journal of Economic Theory. 157, 146-171.

Nakamura, Yutaka 1995. Rank Dependent Utility for Arbitrary Consequence Spaces. Mathematical Social Sciences. 29, 103-129.

Nataf, André 1948. Sur la Possibilité de Construction de Certain Macromodèles. Econometrica. 16, 232-244.

Nau, Robert F. 1995. Coherent Decision Analysis with Inseparable Probabilities and Utilities. Journal of Risk and Uncertainty. 10, 71-91.

Olszewski, Wojciech 2007. Preferences over Sets of Lotteries. Review of Economic Studies. 74, 567-595.

Pearl, Judea 2000. Causality. Models, Reasoning, and Inference. Cambridge University Press, New York.

Sarin, Rakesh K. & Peter P. Wakker 1998. Dynamic Choice and Nonexpected Utility. Journal of Risk and Uncertainty. 17, 87-119.

Savage, Leonard J. 1954. The Foundations of Statistics. Wiley, New York. (Second edition 1972, Dover Publications, New York.)

Schervish, Mark J., Teddy Seidenfeld, & Joseph B. Kadane 1990. State-Dependent Utilities. Journal of the American Statistical Association. 85, 840-847.

Schmeidler, David 1989. Subjective Probability and Expected Utility without Additivity. Econometrica. 57, 571-587.

Smith, James E. & Detlof von Winterfeldt 2004. Decision Analysis in Management Science. Management Science. 50, 561-574.

Trautmann, Stefan & Peter P. Wakker 2015. Making the Anscombe-Aumann Approach to Ambiguity Suited for Descriptive Applications. mimeo.

Tversky, Amos, & Daniel Kahneman 1980. Causal Schemata in Judgments under Uncertainty. In. Martin Fishbein (ed.) Progress in Social Psychology., 49-72, Hillsdale, NJ: Erlbaum.

van Daal, Jan & Arnold H.Q.M. Merkies 1988. The Problem of Aggregation of Individual Economic Relations; Consistency and Representativity in a Historical Perspective. In. Wolfgang Eichhorn (ed.) Measurement in Economics (Theory and Applications of Economic Indices)., 607-637, Physica-Verlag, Heidelberg.

Wakker, Peter P. 1993. Additive Representations on Rank-Ordered Sets II. The Topological Approach. Journal of Mathematical Economics. 22, 1-26.

Wakker, Peter P. 2010. Prospect Theory for Risk and Ambiguity. Cambridge University Press, Cambridge, UK.

Williamson, Jon 2005. Bayesian Nets and Causality. Philosophical and Computational Foundations. Oxford University Press, Oxford.

Chapter 4

Discounted Utility and Present Value: A Close Relation

Han Bleichrodt[a], Umut Keskin[b], Kirsten I.M. Rohde[a],
Vitalie Spinu[a], and Peter P. Wakker[a]

[a]: Erasmus School of Economics, Erasmus University Rotterdam,
the Netherlands
[b]: Istanbul Bilgi University, Istanbul, Turkey

Abstract

We introduce a new type of preference condition for intertemporal choice, which requires present values to be independent of various other variables. The new conditions are more concise and more transparent than traditional ones. They are directly related to applications because present values are widely used tools in intertemporal choice. Our conditions give more general behavioral axiomatizations, which facilitate normative debates and empirical tests of time inconsistencies and related phenomena. Like other preference conditions, our conditions can be tested qualitatively. Unlike other preference conditions, our conditions can also be directly tested quantitatively, and we can verify the required independence of present values from predictors in regressions. We show how similar types of preference conditions, imposing independence conditions between directly observable quantities, can be developed for decision contexts other than intertemporal choice, and can simplify behavioral axiomatizations there. Our preference conditions are especially efficient if several types of aggregation are relevant, because we can handle them in one blow. We thus give an efficient axiomatization of a market pricing system that is (i) arbitrage-free for hedging uncertainties and (ii) time consistent.

Keywords: intertemporal optimization, present value, discounted utility, time inconsistency, arbitrage-free, preference axiomatization

4.1 Introduction

Debates about appropriate models of intertemporal choice, both for the normative purpose of making optimal decisions and for the descriptive purpose of fitting decisions, usually focus on the critical preference conditions that distinguish these models. The two most discussed conditions are time consistency, which plays a role in distinguishing constant and hyperbolic discounting, and intertemporal separability, which pertains to habit formation, satiation, addiction, and sequencing effects.[1] Both time consistency and intertemporal separability are assumed in the classical models but they are usually violated empirically. Their normative status has also been questioned.[2]

To shed new light on the appropriateness of intertemporal decision models, we introduce a new kind of preference conditions to distinguish them, stated directly in terms of present values (PVs). PVs are simple and tractable, and have been widely used in intertemporal choice, both when reflecting the preferences of the financial market[3] and when reflecting subjective preferences of individuals[4]. They relate to the indifferences most commonly en-

[1]See: Attema (2012), Dolan and Kahneman (2008 p. 228), Epper, Fehr-Duda, and Bruhin (2011), Frederick, Loewenstein, and O'Donoghue (2002), Keller and Kirkwood (1999), Loewenstein and Prelec (1993), Tsuchiya and Dolan (2005).

[2]See: Broome (1991), Gold et al. (1996 p. 100), Parfit (1984 Ch. 14), Strotz (1956 p. 178).

[3]See de Wit (1671), Fisher (1930), and Smith and McCardle (1999). Present values are used to compute a company's value when determining stock prices (LeRoy and Porter 1981) and to make investment decisions (Ingersoll and Ross 1992). In such financial decisions at firm or market levels, utility is usually assumed to be linear and discount rates follow market interest rates.

[4]In individual choice experiments, indifferences between a future stream of outcomes and an immediate outcome (the present value) are usually obtained using choice lists.

countered in everyday life. We often have to decide on whether to pay up front for goods consumed later, whether to pay a price now for a financial product with future financial consequences, or whether to choose a savings plan which requires that the money must be delivered now. For these reasons, present values are widely used in experimental measurements of intertemporal preferences.

People can more easily relate to independence conditions imposed on present values than to independence preference conditions. “For your present value of this extra payment on day 10, the payments on the other days do (not) matter” is easier to relate to for most people than the usual preference conditions. In general, PVs can depend on many variables, such as the periods of the receipt of future outcomes, the initial wealth levels at those periods, and the wealth levels at other periods. Our preference conditions will impose independence of PVs from some of those other variables. We show that many models can be characterized by the appropriate independencies.

Like all preference conditions, our conditions can be tested qualitatively. Unlike other preference conditions, our conditions can also be directly tested quantitatively. We can, for instance, carry out regressions with PV as the dependent variable, and the other relevant variables as predictors.[5] We can then test which of those other variables are significantly associated with PV, and whether the variables claimed to be independent in our conditions really are so. Such tests are more widely known and better understood than qualitative tests of preference conditions. To illustrate our new PV

Ahlbrecht and Weber (1997) used both choice lists and direct matching to measure present values. Reviews are in Frederick, Loewenstein, and O’Donoghue (2002) and Soman et al. (2005).

[5]Predictors are often called “independent variables.” We avoid this term so as to avoid confusion

Model	Functional	Pref. conditions	PV condition	PV formula
Non-discounted value	$\sum_{t=0}^{T} x_t$	Additive + Symmetric	$\pi(\varphi)$	φ
Constant discounted value	$\sum_{t=0}^{T} \lambda^t x_t$	Additive + Stationary	$\pi(\varphi, t) = \tau(\varphi, t+1)$	$\lambda^t \varphi$
Time-dependent discounted value	$\sum_{t=0}^{T} \lambda_t x_t$	Additive	$\pi(\varphi, t)$	$\lambda_t \varphi$
Non-discounted utility	$\sum_{t=0}^{T} U(x_t)$	Separable + Symmetric	$\pi(\varphi, e_0, e_t)$	$U^{-1}(U(e_t + \varphi) - U(e_t) + U(e_0)) - e_0$
Constant discounted utility	$\sum_{t=0}^{T} \lambda^t U(x_t)$	Separable + Stationary	$\pi(\varphi, t, e_0, e_t) = \tau(\varphi, t+1, e'_1 = e_0, e'_{t+1} = e_t)$	$U^{-1}(\lambda^t U(e_t + \varphi) - \lambda^t U(e_t) + U(e_0)) - e_0$
Time-dependent discounted utility	$\sum_{t=0}^{T} U_t(x_t)$	Separable	$\pi(\varphi, t, e_0, e_t)$	$U_0^{-1}(U_t(e_t + \varphi) - U_t(e_t) + U_0(e_0)) - e_0$

Table 4.1: The column "Functional" gives the functional forms evaluating $(x_0, \ldots, x_T)$. Here: $0 < \lambda$ is a discount factor; $0 < \lambda_t$ is a period-dependent discount factor with $\lambda_0 = 1$; U is a strictly increasing continuous utility function; U_t is a period-dependent strictly increasing continuous utility function. The column "Pref. conditions" gives preference conditions traditionally used in preference axiomatizations of the functional forms, and defined in Section 4.7. The column "PV condition" gives the PV condition used in our preference axiomatizations indicating how $\pi(\varphi, t, e)$ can be rewritten. Here $\pi(\varphi, t, e)$ denotes the present value of receiving φ extra in period t if the endowment is e. Constant discounting has an extra equality involving τ, tomorrow's value. Note that these cells contain complete definitions. The column "PV formula" gives the formula of PV under each model.

conditions, we apply them to some well-known models. Table 4.1 gives a concise presentation of these models and their representations. Details of the table will be explained in the following sections. Table 4.1 is presented here because it illustrates the organization of the models in the first four sections.

We provide the most concise and most general preference axiomatizations presently available in the literature for: (a) constant discounted value as commonly used by financial markets (Hull 2013); (b) constant discounted utility (Samuelson 1937); (c) general discounted utility, which includes hyperbolic discounting. We also provide results that are relevant to multi-attribute optimization problems other than intertemporal: (d) no-bookmaking and no-arbitrage for uncertainty, which are commonly used for financial markets; (e) additive separability for general multi-attribute aggregations (Debreu 1960; Gorman 1968). In such other contexs we should find a quantitative index that can play a role similar to present value for intertemporal choice. Section 4.6 considers aggregation over two dimensions, for instance, time and uncertainty. Here our technique is particularly efficient because it can handle both aggregations in one blow. We derive the most common pricing model used in finance: as-if risk neutrality together with constant discounting, which avoids arbitrage for both uncertainty and time.

4.2 Preferences and subjective PVs

We derive appropriateness of an intertemporal goal function V from the decisions that it implies, modeled through a binary preference relation $\succcurlyeq$ over *outcome sequences* $x = (x_0, \ldots, x_T) \in \mathbb{R}^{T+1}$. The preferences can, for instance, concern (i) observed consumer choices in descriptive applications, or (ii) pension savings plans or market prices with the financial market taken as decision maker in prescriptive applications. The outcome sequence yields *outcome* x_t in *period* t, for each t; $t = 0$ denotes the present. We assume $T \geq 2$ to avoid trivialities, and keep all other aspects of our analysis as simple as possible (assuming one fixed $T \in \mathbb{N}$) so as to focus on the novelty of our conditions. We use indexed Roman letters x_t to specify the period t of receipt of outcome x_t, and Greek letters $\alpha, \beta, \ldots$ to refer to outcomes (real numbers) when no period of receipt needs to be specified. By $\alpha_t x$ we denote x with x_t replaced by α.

The goal function V *represents* $\succcurlyeq$ if $V : \mathbb{R}^{T+1} \to \mathbb{R}$ and $x \succcurlyeq y \Leftrightarrow V(x) \geq V(y)$ for all $x, y \in \mathbb{R}^{T+1}$. The existence of a representing V implies that $\succcurlyeq$ is a *weak order*; i.e., $\succcurlyeq$ is *complete* ($x \succcurlyeq y$ or $y \succcurlyeq x$ for all x, y) and transitive. We therefore assume throughout that $\succcurlyeq$ is a weak order. Strict preference $\succ$, indifference $\sim$, reversed preference $\preccurlyeq$, and strict reversed preference ($\prec$) are as usual. We also assume *monotonicity* (strictly improving an outcome strictly improves the outcome sequence) and continuity of $\succcurlyeq$ throughout. The conditions imply that all discount weights in this paper are positive, and that all utility functions are continuous and strictly increasing.

The following condition considers sums of outcomes $x_t = \varphi + e_t$. Here we call e_t an *initial endowment*. The specification of the initial endowment only serves to facilitate interpretations and does not refer to any type of reference dependence. Formally, our analysis is entirely in terms of final wealth, and is classical in this respect.

Definition 4.1. $\pi(\varphi, t, e)$ is the *present value* (PV) of outcome φ received in period t with *(initial) endowment* $e = (e_0, \ldots, e_T)$ if $(e_0 + \pi(\varphi, t, e))_0 e \sim (e_t + \varphi)_t e$, i.e.

$$(e_0+\pi(\varphi, t, e), e_1, \ldots, e_{t-1}, e_t, e_{t+1}, \ldots, e_T) \sim (e_0, \ldots, e_{t-1}, e_t+\varphi, e_{t+1}, \ldots, e_T). \tag{4.1}$$

Equation 4.1 means that, with e the current endowment, receiving an additional future outcome φ in period t is exactly offset by receiving an additional present outcome π. That is, the PV of a future outcome φ in period t is π. For simplicity, we assume in this paper that a PV always exists. Generalizations are discussed in Section 4.8.

By monotonicity, the PV is unique. In applications, PVs (denoted π here) are often used for general outcome sequences x with endowment e: $(e_0 + \pi)_0 e \sim e + x$. However, we will not need this general concept in this paper.

The PV, π, can in general depend on all of φ, t, and e, and π is a function $\pi(\varphi, t, e)$. As a convention, if we write π without its arguments, then it designates the general function depending on all its arguments. For $t = 0$, it trivially follows that $\pi(\varphi, 0, e) = \varphi$. Note that π can be subjective, depending on $\succcurlyeq$, and thus it reflects the tastes and attitudes of the decision maker. The preference conditions presented in the following sections all amount to independence of PV, π , from some of the variables (φ, t, e). We express this independence by writing only the arguments that π depends on. For example, if $\pi(\varphi, t, e)$ depends only on φ (i.e., is independent of t and e) then we write

$$\pi = \pi(\varphi). \tag{4.2}$$

Similarly, if π depends on e only through e_0 and e_t, then we write

$$\pi = \pi(\varphi, t, e_0, e_t). \tag{4.3}$$

Preference conditions should be directly verifiable from preferences, the observable primitives in the revealed preference paradigm, without invoking theoretical constructs such as utilities. In Section 4.7, we show how our PV conditions can be rewritten in terms of preferences. They are therefore genuine preference conditions, and our conditions can be tested in the same way as all qualitative preference conditions.

4.3 Linear utility

This section considers models with linear utility, as commonly used in financial markets. Such models can serve as approximations for subjective individual choices if the stakes are moderate (Epper, Fehr-Duda, and Bruhin 2011; Luce 2000 p. 86; Pigou 1920 p. 785). The first model in Table 4.2 maximizes the sum of outcomes:

$$\textit{Non–Discounted Value:} \ \sum_{t=0}^{T} x_t. \tag{4.4}$$

This model does not involve subjective parameters, is directly observable, and does therefore not need a preference axiomatization. But it serves well as a first illustration of the nature of PV conditions.

Theorem 4.1. *The following two statements are equivalent:*

1. *Non-discounted value holds.*

2. $\pi = \pi(\varphi)$.

Throughout this chapter, Condition 4.n.2 refers to Statement 2. of Theorem 4.n. Theorem 4.1 shows that if π depends only on φ (Condition 4.1.2) in whatever general sense one might think of, then it must be through the identity function $\pi(\varphi) = \varphi$. This implication may seem surprising at first, the more so as Condition 4.1.2 in addition implies the summation operation in non-discounted value in Eq. 4.4. To illustrate the strength of Condition 4.1.2, first note that substituting $t = 0$ already implies that φ can only be the identity. The following informal proof further illustrates the condition: According to Condition 4.1.2, the extra value of any extra future outcome is always the same, and can therefore be added to today's wealth. Then all that matters is the sum of all future outcomes, which may as well be received immediately today. The implication $\pi(\varphi) = \varphi$ can also be inferred from the last two columns of the corresponding row of Table 4.2.

Our next model involves a subjective parameter, the discount factor λ:

$$\textit{Constant Discounted Value: } \sum_{t=0}^{T} \lambda^t x_t \quad for\ \lambda > 0. \tag{4.5}$$

By monotonicity, $\lambda > 0$. Under the usual assumption that the decision maker is impatient, we have $\lambda \leq 1$. In PV calculations of cash flows, constant discounted value is commonly used, setting $\lambda = 1/(1 + r)$ with r the interest or discount rate of the market. In this case, if the decision maker is, say, an individual financial trader, the discount factor λ is not a subjective parameter reflecting the attitude of the decision maker but it is a given constant, publicly known and determined by the market. The following theorem then does not apply to the financial trader in the role of decision maker.

The following theorem is still relevant for market pricings if the financial market is the decision maker who determines (rational) PVs. Then λ reflects the market attitude, which may, for instance, be determined by the attitude of a central bank choosing a goal function for its optimal control problem, and which in this sense is subjective. Condition 4.2.2 rationalizes this common evaluation system. In other contexts where the parameter λ reflects the attitude of an individual decision maker, it will probably be influenced by the market interest rate but need not be identical to it, for instance, because it may incorporate extra risks borne (Smith and McCardle 1999, Smith 1998).

In the following theorem, we use tomorrow's value as an analogue to PV, defined as follows:

Definition 4.2. τ is *tomorrow's value* of an outcome φ received in period t ($t \geq 1$) with *endowment* e, denoted $\tau = \tau(\varphi, t, e)$, if $(e_1 + \tau)_1 e \sim (e_t + \varphi)_t e$, i.e.,

$$(e_0, e_1 + \tau, e_2, \ldots, e_t, \ldots, e_T) \sim (e_0, e_1, \ldots, e_t + \varphi, \ldots, e_T). \tag{4.6}$$

Here τ is the extra outcome in period 1 that exactly offsets the extra outcome φ in period t. Thus τ is tomorrow's PV. Such "future" present values are central tools in recursive intertemporal models (Campbell and Shiller 1987; Ju and Miao 2012 and their references; Maccheroni, Marinacci, and Rustichini 2006). Experimental measurements of subjective individual discounting in studies often compare present values with tomorrow's values. To measure the latter, so-called front-end delays are then added (Ahlbrecht and Weber 1997; Luhmann 2013). The main violations of time consistency occur when present value is changed into tomorrow's value (immediacy effect), and this effect is central in the popular quasi-hyperbolic discount model (Laibson 1997). Section 8 discusses preference conditions for the following theorem entirely in terms of present values. We prefer using tomorrow's value here because it leads to the most appealing condition that we have been able to find. As with π, if we write τ without its arguments then it designates the general function depending on all its arguments.

Theorem 4.2. *The following two statements are equivalent:*

1. *Constant discounted value holds.*

2. $\pi = \pi(\varphi, t) = \tau(\varphi, t+1) = \tau$. [6]

In Statement 1., the discount factor λ (λ *as in Eq.* 4.5) *is uniquely determined.*

Condition 4.2.2 entails that π and τ are independent of the endowments, and that tomorrow's perception of future income is the same as today's. The condition implies that PV depends only on stopwatch time (time differences) and not on calendar time (absolute time).

Statement 4.2.2 formulates the common stationarity in a simplified manner for the case of linear utility. Only one future outcome φ and one present value today (π) or tomorrow (τ) are involved, rather than involving general preferences between general outcome sequences as in common formulations. Most tests of stationarity in the literature are, in fact, tests of our simplified condition (see Takeuchi 2010 and his extensive survey), which captures the essence of the condition.

Many studies have shown that constant discounting is violated empirically. Hence the following generalization is of interest:

$$\textit{Time-Dependent-Discounted Value: } \sum_{t=0}^{T} \lambda_t x_t \ (\textit{with } \lambda_0 = 1). \tag{4.7}$$

The weights λ_t are all positive by monotonicity. This model allows for general discount weights with unrestricted time dependence. Many special cases of

[6]The notation here is short for: $\pi = \pi(\varphi, t, e) = \pi(\varphi, t) = \tau(\varphi, t+1) = \tau(\varphi, t+1, e')$. The two endowments e and e' are immaterial, and are allowed to be different. Because of the shift by one period, the condition is imposed only for all $t < T$. Existence of π and the equality in Statement 4.2.2 imply that τ also exists.

such discount weights have been studied in the literature, the best-known being hyperbolic discounting. The representation in Eq. 4.7 is not affected if all λ_ts are multiplied by the same positive factor. The common scaling $\lambda_0 = 1$ is therefore always possible.

Theorem 4.3. The following two statements are equivalent:

1. *Time-dependent discounted value holds.*

2. $\pi = \pi(\varphi, t)$.

In Statement 1., the discount factors λ_t *(*λ_t *as in Eq.* 4.7*) are uniquely determined.*

An implication that can be inferred from the last two columns of Table 4.2 is that if π is any function of φ and t then it must be the function $\lambda_t \times \varphi$.

4.4 Nonlinear utility

The models presented in the preceding section take a weighted or unweighted sum of the outcomes. They assume constant marginal utility in the sense that an extra euro received in a particular period gives the same utility increment regardless of the endowment of that period. In individual choice, unlike market pricing, this condition is often violated empirically and it is not normative. More realistic and more popular models allow for nonlinear utility. Then marginal utility depends on the endowment, and the models of the preceding section become:

$$\textit{Non-Discounted Utility: } \sum_{t=0}^{T} U(x_t); \tag{4.8}$$

$$\textit{Constant Discounted Utility: } \sum_{t=0}^{T} \lambda^t U(x_t); \tag{4.9}$$

$$\textit{Time-Dependent Discounted Utility: } \sum_{t=0}^{T} U_t(x_t). \tag{4.10}$$

Continuity and monotonicity of $\succcurlyeq$ readily imply $\lambda > 0$ and strict increasingness and continuity of U and all U_ts. Eq. 4.9 is Samuelson's (1937) discounted utility, the most popular model for intertemporal choice. Each utility model reduces to the corresponding value model if utility is linear. In particular, in time-dependent discounted utility, if the $U_t(\alpha)$ are linear, they can be written as $\lambda_t \times \alpha$ and we can renormalize them such that $\lambda_0 = 1$, resulting in time-dependent discounted value results.

The mathematics underlying the preference axiomatizations of the utility models in Eqs. 4.8 - 4.10 is more advanced than for Theorems 4.1, 4.2, and 4.3. Whereas these theorems solved linear equalities, we now have to deal with nonlinear equalities, with nonlinear utilities intervening. Fortunately, this increased mathematical complexity does not show up in the preference conditions and, consequently, in the empirical tests of the models. The relevant PV preference conditions are obtained directly from those defined in Section 4.3 by adding dependence on the endowment levels e_0 and e_t. This way we readily obtain Theorems 4.4, 4.5, and 4.6 from Theorems 4.1, 4.2, and 4.3, respectively.

Theorem 4.4. *The following two statements are equivalent:*

1. *Non-discounted utility holds.*

2. $\pi = \pi(\varphi, e_0, e_t)$. [7]

The following uniqueness result holds for Statement 1: A real-valued time-independent constant μ can be added to U (U as in Eq. 4.8), and U can be multiplied by a positive constant ν.

An implication, displayed in the last two columns of Table 4.2, is that if π is any function of φ, e_0, and e_t, then it must be of the form displayed there.

Several authors have argued that any discounting, even if consistent over time, is irrational, and have thus recommended using non-discounted utility for intertemporal choice (Jevons 1871 pp. 72-73; Ramsey 1928; Rawls 1971). Condition 4.4.2 characterizes this proposal. Studies providing preference axiomatizations for non-discounted utility (sums of utilities) include Kopylov (2010), Krantz et al. (1971), Marinacci (1998), Pivato (2014), and Wakker (1986). We now turn to discounting.

Theorem 4.5. *The following two statements are equivalent:*

1. *Constant discounted utility holds.*

2. $\pi = \pi(\varphi, t, e_0, e_t) = \tau(\varphi, t+1, e'_1 = e_0, e'_{t+1} = e_t)$.[8]

[7] Here t refers to the period where φ is added and e_t specifies the endowment in that period. Note that, whereas π depends on the level of e_t, it does not depend explicitly on period t, which we denote by suppressing t from the arguments of $\pi(\ldots)$.

[8] The condition implies that the endowment levels $e_j, j \neq 0, t$ of the PV endowment e, and similarly the endowment levels $e'_j, j \neq 1, t+1$ of the tomorrow-value endowment e', are immaterial. Existence of π and the equality imply that τ also exists. The condition

The following uniqueness result holds for Statement 1: A real-valued time-independent constant μ can be added to U (U and λ as in Eq 4.9) *and U can be multiplied by a positive constant ν. λ is uniquely determined.*

The first preference axiomatization of constant discounted utility was in Koopmans (1960), with generalizations in Harvey (1995), Bleichrodt, Rohde, and Wakker (2008), and Kopylov (2010). Condition 4.5.2 requires that the same tradeoffs are made tomorrow as today. Such requirements have sometimes been taken as rationality requirements (see the introduction). The following theorem generalizes Theorem 4.3.

Theorem 4.6. *The following two statements are equivalent:*

1. *Time-dependent discounted utility holds.*

2. $\pi = \pi(\varphi, t, e_0, e_t)$.

The following uniqueness result holds for Statement 1: A time-dependent real constant μ_t can be added to every U_t (U_t as in Eq. 4.10), *and all U_t's can be multiplied by a joint positive constant ν.*

Statement 4.6.2 expresses that the tradeoffs between periods 0 and t are independent of what happens in the other periods, reflecting a kind of separability. Time-dependent discounted utility is a general additive representation, which has been axiomatized several times before.[9] It implies intertemporal

holds for all $0 \leq t \leq T-1$, and whenever $e'_1 = e_0$ and $e'_{t+1} = e_t$. We use the convenient argument matching notation popular in programming languages (R: see R Core Team 2013, Python: see Python Software Foundation 2013, Scala: see Ecole Polytechnique Federale de Lausanne 2013, among others) to express the latter two restrictions. Here the formal argument of a function (e'_1 or e'_{t+1}) is assigned a value (e_0 or e_t).

[9]See Debreu (1960), Gorman (1968), Krantz et al. (1971), Wakker (1989).

separability, which is arguably the most questionable assumption of most intertemporal choice models (Baucells and Sarin 2007; Dolan and Kahneman 2008). Another generalization of time-dependent discounted value (eq. 4.7) can be considered, which is intermediate between Eqs. 4.9 and 4.10.

$$\sum_{t=0}^{T} \lambda_t U(x_t). \tag{4.11}$$

We have not yet succeeded in finding an appealing present value condition for this representation.

4.5 Applications to contexts other than intertemporal choice

The mathematical results of the previous sections and the preference conditions used can be applied in contexts other than intertemporal choice. For instance, Theorem 4.3 is of special interest for decision under uncertainty, capturing nonarbitrage in finance. To see this point, we reinterpret the periods t as *states of nature*. Exactly one state obtains, but it is uncertain which one (Savage 1954). Now $x = (x_0, \ldots, x_T)$ refers to an uncertain prospect yielding outcomes x_t if state of nature t obtains. For simplicity, we focus on a single time for all outcomes here, so that discounting plays no role. The next section will consider both uncertainty and time. If we divide the discount weights λ_t by their sum $\sum_{t=0}^{T} \lambda_t$ (relaxing the requirement of $\lambda_0 = 1$), they sum to 1, and the representation becomes subjective expected value.

Subjective expected value was first axiomatized by de Finetti (1937) using a no-book argument, which is equivalent to the no-arbitrage condition of finance. In finance, the representation is as-if risk neutral, and the decision maker is the market which sets rational prices for state-contingent assets.

For state j, a state-contingent asset $x = (0, \ldots, 0, 1, 0, \ldots, 0)$ yields outcome 1 if j happens and nothing otherwise. In this interpretation, $lambda_j$ becomes the market price of this state-contingent asset, and PV is the offsetting quantity of state-0-contingent assets. Condition 4.3.2 provides the most concise formulation of the no-book and the no-arbitrage principle presently available in the literature.

For decision under uncertainty, certainty equivalents are more natural quantities than state-contingent prices. Reformulating our conditions in terms of certainty equivalents is a topic for future research. For decision under risk, Eq. 4.8 can be interpreted as von Neumann-Morgenstern expected utility for equal-probability lotteries, which essentially covers all lotteries with rational probabilities (writing every probability i/j as i probabilities $1/j$). Eq. 4.8 can also be interpreted as ambiguity under complete absence of information (Gravel, Marchant, and Sen 2012). Eq. 4.10 is Debreu's (1960) additively separable utility. Here again, our Statement 4.6.2 provides the most concise preference axiomatization presently available in the literature.

4.6 Time and uncertainty: Aggregating over two dimensions

This section applies our technique to aggregations over two dimensions. We consider the special case where one dimension refers to time and the other refers to uncertainty. In applications, usually both time and uncertainty play a role (Smith and McCardle 1999). We assume periods $0, \ldots, T$ and states of nature $0, \ldots, n$. Exactly one state is true but the decision maker is uncertain which one. We consider $(T + 1) \times (n + 1)$ tuples $(x_{0,0}, \ldots, x_{T,n})$ yielding outcome $x_{t,s}$ in period t if state of nature s is true. Such tuples are called *act sequences*. Thus, every period yields an *act* (map from states to $\mathbb{R}$), and every state of nature yields an outcome sequence. *Constant discounted expected value* is

$$\sum_{t=0}^{T}\sum_{s=0}^{n}\lambda^t p_s x_{t,s} \tag{4.12}$$

with $\lambda > 0$, $p_s > 0$ for all s, and $\sum p_s = 1$. Constant discounted expected value is the common evaluation system used in cost-effectiveness studies and by financial markets. In the latter case, the p_js and λ are the parameters. They are subjective from the market perspective. The evaluation formula is both arbitrage-free and time consistent (under the common time invariance).

We use state-contingent present values, defined as follows, and using payoffs in state 0 and period 0 for calibration: $\pi = \pi(\varphi, t, s, e_{0,0}, e_{0,1}, \ldots, e_{T,n})$ is such that

$$(e_{0,0} + \pi, e_{0,1}, \ldots, e_{t,s}, \ldots, e_{T,n}) \sim (e_{0,0}, e_{0,1}, \ldots, e_{t,s} + \varphi, \ldots, e_{T,n}). \tag{4.13}$$

The following reinforcement of monotonicity is common in decision under uncertainty. First, we identify a sure outcome sequence $(x_0, \ldots, x_T)$ with the act sequence that assigns x_t to each (t, s), and we induce preferences over outcome sequences ($T + 1$ tuples) this way. Second, we define *dominance* to hold if: (a) Preferences over outcome sequences satisfy monotonicity and, (b) replacing an outcome sequence contingent on a state of nature s by a weakly (strictly) preferred outcome sequence leads to a weakly (strictly) preferred $(T + 1) \times (n + 1)$ tuple. In the next theorem, we use the same notation for tomorrow's value τ as in Section 4.3, but now it is state-contingent. That is, we now use payoffs in state 0 and period 1 (tomorrow) for calibration: $\tau = \tau(\varphi, t, s, e_{0,0}, e_{0,1}, \ldots, e_{T,n})$ is such that

$$(e_{0,0}, \ldots, e_{0,n}, e_{1,0} + \tau, e_{1,1} \ldots, e_{t,s}, \ldots, e_{T,n}) \sim (e_{0,0}, e_{0,1}, \ldots, e_{t,s} + \varphi, \ldots, e_{T,n}). \tag{4.14}$$

Theorem 4.7. *Assume that $\succcurlyeq$ is a binary relation on $\mathbb{R}^{(T+1)\times(n+1)}$. It is represented by constant discounted expected value if and only if it is a continuous weak order satisfying dominance and:*

$$\pi(\varphi, t, s) = \tau(\varphi, t+1, s). \tag{4.15}$$

The parameters $\lambda, p_1, \ldots, p_n$ (as in Eq. 4.12) are uniquely determined.

For extending this result to nonlinear utility, expected utility for the aggregation over the states of nature (using an analog of Eq. 4.11) is of special interest. We leave this as a topic for future research. There is currently much interest in models with both risk and time, and their interactions. Baucells and Heukamp (2012) proposed a very general decision model. As in the preceding section, it is also desirable to obtain results in terms of present certainty equivalents rather than in terms of present contingent payments here. For example, Smith (1998) considered a combination of risk and time in a theoretical study, combining present values with certainty equivalents, Ahlbrecht and Weber (1997) did the same in an experimental study, and Pelsser and Stadje (2014) considered market pricings as in Theorem 4.7.

4.7 Proofs and clarification of the empirical status of PV conditions

We first present the proofs of Theorems 4.1 - 4.6. We first present the proofs of Theorems 4.1-4.6. We present them from most to least general because this approach is most clarifying and most efficient. The presentation of the proofs clarifies the relationship between our PV conditions and well-known preference conditions, showing that PV conditions indeed are preference conditions. In each proof, we start from our PV condition, which is always weaker than the conditions that are derived and that are commonly used in the literature. We thus show that our PV conditions give stronger results. Because each Statement 2. is immediately implied by substitution of the functional, we throughout assume Statement 2. and derive Statement 1. and the uniqueness results.

We first present the proofs of Theorems 4.1-4.6. We present them from most to least general because this approach is most clarifying and most efficient. The presentation of the proofs clarifies the relationship between our PV conditions and well-known preference conditions, showing that PV conditions indeed are preference conditions. In each proof, we start from our PV condition, which is always weaker than the conditions that are derived and that are commonly used in the literature. We thus show that our PV conditions give stronger results. Because each Statement 2. is immediately implied by substitution of the functional, we throughout assume Statement 2. and derive Statement 1. and the uniqueness results.

Proof of Theorem 4.6

The uniqueness results for Statement 4.6.1, which uses Eq. 4.10, follow from well-known uniqueness results in the literature (Krantz et al. 1971 Theorem 6.13; Wakker 1989 Observation III.6.6). We next derive Statement 4.6.1.

The equality $\pi(\varphi, t, e) = \pi(\varphi, t, e_0, e_t)$ in Condition 4.6.2 means that π is independent of $e_j, j \neq 0, t$. This holds if and only if

$$(e_0 + \pi, e_1, \ldots, e_{t-1}, e_t, e_{t+1}, \ldots e_T) \sim (e_0, e_1, \ldots, e_{t-1}, e_t + \varphi, e_{t+1}, \ldots e_T) \Rightarrow$$

$$(e_0 + \pi, e'_1, \ldots, e'_{t-1}, e_t, e'_{t+1}, \ldots e'_T) \sim (e_0, e'_1, \ldots, e'_{t-1}, e_t + \varphi, e'_{t+1}, \ldots e'_T) \quad (4.16)$$

This holds if and only if the implication holds with twice preference $\succcurlyeq$

instead of indifference $\sim$.[10] Eq. 4.16 with preference instead of indifference is known as *separability* of $\{0,t\}$ (Gorman 1968). By repeated application of Gorman (1968), separability of every set $\{0,t\}$ holds if and only if $\succcurlyeq$ is *separable*; i.e., every subset of $\{0,\ldots,T\}$ is *separable* (preferences are independent of the levels where outcomes outside this subset are kept fixed, as with separability of $\{0,t\}$). This holds if and only if an additively decomposable representation holds[11], which we call time-dependent discounted utility in the main text. □

Proof of Theorem 4.5

The uniqueness results for Statement 4.5.1, which uses Eq. 4.9, follow from those in Theorem 4.6, where, in terms of Eq. 4.10, λ is the proportion U_{t+1}/U_t for any t. It is useful to note that the sum of weights, $\sum \lambda^t$, is the same for each outcome sequence, implying that there is no special role for utility value 0. We next derive Statement 4.5.1.

The equality $\pi(\varphi, t, e_0, e_t) = \tau(\varphi, t+1, e'_1 = e_0, e'_{t+1} = e_t)$ in Condition 4.5.2 means that π is independent of $e_j, j \neq 0, t, T$ (as in Theorem 4.6, but now only for $t < T$), but also of whether it is measured in period 0 or period 1. This holds if and only if, writing α for $e_0 = e'_1$ and β for $e_t = e'_{t+1}$,

$$(\alpha+\pi, e_1, \ldots, e_{t-1}, \beta, e_{t+1}, \ldots e_T) \sim (\alpha, e_1, \ldots, e_{t-1}, \beta+\varphi, e_{t+1}, \ldots e_T) \Leftrightarrow$$
$$(e'_0, \alpha+\pi, \ldots, e'_{t-1}, e'_t, \beta, \ldots e'_T) \sim (e'_0, \alpha, \ldots, e'_{t-1}, e'_t, \beta+\varphi, \ldots e'_T) \quad (4.17)$$

It implies separability of $\{0,t\}$ for all $t < T$, as in Theorem 4.6, but now, instead of separability of $\{0,T\}$ we have separability of all $\{1,t+1\}$ for all

[10] If the upper indifference is changed into a strict preference, then we find $\pi' < \pi$ to give indifference. The lower indifference follows with π' instead of π. By monotonicity, replacing π' by π leads to the lower strict preference.

[11] See Debreu (1960), Gorman (1968), Krantz et al. (1971 Theorem 6.13), Wakker (1989 Theorem III.6.6).

$t < T$. The latter separability can, for instance, be seen by replacing all primes in Eq. 4.17 by double primes, which should not affect the second indifference because of the maintained equivalence with the first indifference. By repeated application of Gorman (1968), we still get separability of $\succcurlyeq$. Hence the above condition holds if and only if: time-dependent discounted utility holds with, further, U_0, U_t additively representing the same preference relation over $\mathbb{R}^2$ as U_1, U_{t+1} do. We can set $U_t(0) = 0$ for all t. Then, by standard uniqueness results (Wakker 1989 Observation III.6.6′), $U_1/U_0 = U_{t+1}/U_t = \lambda$ for a positive constant λ. This proves the equivalence in Theorem 4.5. Because Condition 4.5.2 implies constant discounted utility, it implies *stationarity*, used by Koopmans (1960) to axiomatize the model. The latter condition is defined as follows:

$$\begin{aligned}(x_0, x_1, \ldots, x_{T-1}, c_T) \succcurlyeq (y_0, y_1, \ldots, y_{T-1}, c_T) \Leftrightarrow \\ (c_0, x_0, \ldots, x_{T-1}) \succcurlyeq (c_0, y_0, \ldots, y_{T-1}).\end{aligned} \tag{4.18}$$

Our condition is weaker by considering tradeoffs between two periods, keeping the outcomes in all other periods fixed. □

Proof of Theorem 4.4

The uniqueness results for Statement 4.4.1, which uses Eq. 4.8, follow from those in Theorem 4.5. We next derive Statement 4.4.1.

The equality $\pi(\varphi, t, e) = \pi(\varphi, e_0, e_t)$ in Condition 4.4.2 means that π is not only independent of $e_j, j \neq 0, t$, as in Theorem 4.6, but also of t. This holds if and only if[12]

$$
\begin{gathered}
(e_0 + \pi, e_1, \ldots, e_{t-1}, e_t, e_{t+1}, \ldots e_T) \sim \\
(e_0, e_1, \ldots, e_{t-1}, e_t + \varphi, e_{t+1}, \ldots e_T) \\
\text{implies} \\
(e_0 + \pi, e'_1, \ldots, e'_{t'-1}, (e_t)_{t'}, e'_{t'+1}, \ldots e'_T) \sim \\
(e_0, e'_1, \ldots, e'_{t'-1}, (e_t)_{t'} + \varphi, e'_{t'+1}, \ldots e'_T) \qquad .
\end{gathered}
\tag{4.19}
$$

It readily follows that the above condition holds if and only if we have all the conditions of Theorem 4.6 and its representation, with the extra condition $U_t(e_t + \varphi) - U_t(e_t) = U_{t'}(e_t + \varphi) - U_{t'}(e_t)$, implying that we can take all functions U_t the same, independent of t. It implies symmetry, the condition commonly used in the literature to axiomatize non-discounted utility. *Symmetry* requires invariance of preference under every permutation of the outcomes. Symmetry immediately implies that π is independent of t, which is what Condition 4.4.2 adds to Condition 4.6.2. This shows once again that the PV conditions are weak compared to conditions commonly used in the literature. $\square$

PROOF OF THEOREMS 4.1, 4.2, AND 4.3

The uniqueness of the discount parameters in Theorems 4.2 and 4.3, based on Eqs. 4.5 and 4.7, follows from the uniqueness results of Theorems 4.6 and 4.5. We next derive the Statements 1.

The proof of Theorem 4.3 [4.2, 4.1] readily follows from Theorem 4.6, [4.5, 4.4], as follows. The theorems to be proved are the linear counterparts of the theorems from which they follow. The preference conditions are always the same except that dependence of the endowment levels e_0, e_t has been

[12]In the following equation we write $(e_t)_{t'}$ for $e'_{t'}$ to indicate that the t' level of e' is the same as e_t, the tth level of e.

dropped. In the notation of Theorem 4.6 this means that $U_t(e_t+\varphi)-U_t(e_t)$ is independent of e_t, which implies linearity of U_t. Similarly, the utility functions in Theorems 4.5 and 4.4 are linear. Then Theorems 4.3, 4.2, and 4.1 follow. □

For completeness, we show how the conditions of Theorems 4.1 - 4.3 can be restated directly in terms of preferences:

The equality $\pi=\pi(\varphi,t)$ in Condition 4.3.2 holds if and only if

$$\begin{gathered}(e_0+\pi, e_1, \ldots, e_{t-1}, e_t, e_{t+1}, \ldots e_T) \sim \\ (e_0, e_1, \ldots, e_{t-1}, e_t+\varphi, e_{t+1}, \ldots e_T) \\ \text{implies} \\ (e_0'+\pi, e_1', \ldots, e_{t-1}', e_t', e_{t+1}', \ldots e_T') \sim \\ (e_0', e_1', \ldots, e_{t-1}', e_t'+\varphi, e_{t+1}', \ldots e_T') \qquad . \end{gathered} \tag{4.20}$$

The equality $\pi=\pi(\varphi,t)=\tau(\varphi,t+1)$ in Condition 4.2.2 holds if and only if

$$\begin{gathered}(e_0+\pi, e_1, \ldots, e_{t-1}, e_t, e_{t+1}, \ldots e_T) \sim \\ (e_0, e_1, \ldots, e_{t-1}, e_t+\varphi, e_{t+1}, \ldots e_T) \\ \text{if and only if} \\ (e_0', e_1'+\pi, \ldots, e_{t-1}', e_t', e_{t+1}', \ldots e_T') \sim \\ (e_0', e_1', \ldots, e_{t-1}', e_t', e_{t+1}'+\varphi, \ldots e_T') \qquad . \end{gathered} \tag{4.21}$$

The equality $\pi = \pi(\varphi)$ in Condition 4.1.2 holds if and only if

$$\begin{gathered}(e_0 + \pi, e_1, \ldots, e_{t-1}, e_t, e_{t+1}, \ldots e_T) \sim \\ (e_0, e_1, \ldots, e_{t-1}, e_t + \varphi, e_{t+1}, \ldots e_T) \\ \text{implies} \\ (e'_0 + \pi, e'_1, \ldots, e'_{t'-1}, e_{t'}, e'_{t'+1}, \ldots e'_T) \sim \\ (e'_0, e'_1, \ldots, e'_{t'-1}, e_{t'} + \varphi, e'_{t'+1}, \ldots e'_T)\end{gathered} \quad . \tag{4.22}$$

We next compare the PV conditions in Theorems 4.1 - 4.3 with other conditions used in the literature to axiomatize the models in question. Condition 4.3.2 implies that the extra value of an extra outcome φ is independent of the level e_t to which it is added. This is implied by the well known *additivity* condition, requiring that a preference $x \succcurlyeq y$ is not affected by adding the same constant α to x_t and y_t. By adding the same constant to e_t and $e_t + \varphi$, we can change them into e'_t and $e'_t + \varphi$, implying our preference condition. Additivity is necessary and sufficient for time-dependent discounted utility (Wakker 2010 Theorem 1.6.1). It is more restrictive than our condition because we only consider tradeoffs between two periods, keeping the outcomes in all other periods fixed. Similar observations apply to the elementary Theorem 4.2 and the trivial Theorem 4.1.

Proof of Theorem 4.7

As before, necessity of the preference conditions is obvious, so that we assume the preference conditions and derive constant discounted expected value and the uniqueness results.

Statement 2. implies that π can be written as $\pi(\varphi, t, s)$ and is independent of the endowments. By Theorem 4.3, treating the paired indexes (t, s) as one index with index $(0, 0)$ here playing the role of index 0 in Theorem 4.3, we obtain a linear representation $\sum_{t=0}^{T} \sum_{s=0}^{n} \mu_{t,s} x_{t,s}$. We do not impose the restriction that $\mu_{0,0} = 1$ here and, hence, the weights are uniquely determined up to one common positive factor.

By dominance, for every fixed s we have the same preference relation over outcome sequences. Hence, each such reference relation is represented by a positive constant (depending on s) times $\sum_{t=0}^{T} \mu_{t,0} x_{t,s}$. This follows from the uniqueness result of Theorem 4.3, now applied with s kept fixed, and with the requirement $\lambda_0 = 1$ (here $\mu_{0,0}$) dropped. We can rewrite the representation as $\sum_{t=0}^{T} \sum_{s=0}^{n} \lambda_t p_s x_{t,s}$ with the p_s's summing to 1 and, hence, uniquely determined. We can renormalize further so that $\lambda_0 = 1$, after which all weights are uniquely determined. By Theorem 4.2, $\lambda_t = \lambda^t$ for $\lambda = \lambda_1$. Thus constant discounted expected value holds and uniqueness of the weights has also been established. $\square$

4.8 Discussion

The primary purpose of preference axiomatizations is to make decision models with theoretical constructs directly observable, by restating their existence (Statements 1. in our theorems) in terms of preference conditions (Statements 2. in our theorems). The simpler the preference conditions, the better they clarify the empirical meaning of the decision models. Similarity between the preference conditions and the functional helps to clarify the empirical meaning of the decision model. Hence this paper has introduced preference axiomatizations that are as simple as possible and that reflect the corresponding decision models as well and transparently as possible.

The conditions in our Statements 2. use fewer words and characters than any conditions previously proposed in the literature, which provides an objective criterion for our claim that they are the most concise conditions presently existing. Further, we think that our conditions are easy to understand and test because present values are familiar objects. PVs can be used as the goal functions to be optimized in intertemporal choice. At the same time, they are directly defined in terms of preferences and, hence, the subjective and behavioral character of preference axiomatizations is not lost by using PVs. Our efficient results are based on this dual nature of PVs.

We used tomorrow's values in the characterizations of constant discounting, but conditions entirely in terms of present values are also possible. For example $\pi(\varphi, t) = \sqrt{\pi(\varphi, t-1)\pi(\varphi, t+1)}$ characterizes constant discounted value. We were unable to find an easy way to extend this condition to nonlinear utility. An alternative condition is

$$\pi(\varphi, t) = \pi(\pi(\varphi, 1), t-1),$$

reflecting that the recursive structure at period t should be the same as at period 1 under constant discounting. The condition can be extended to nonlinear utility by specifying the relevant e levels:

$$\pi(\varphi, t, e_0, e_t) = \pi(\pi(\varphi, 1, e'_0 = e_{t-1}, e'_1 = e_t), t-1, e_0, e_{t-1}).$$

Because of the many e levels, the latter condition is not very transparent. Our formulations using tomorrow's value are more transparent. This case suggests that it is not always easy to find simple reformulations of preference conditions in terms of present values. We were also unable to find an easy condition in terms of present values for the representation $\sum_{t=0}^{T} \lambda_t U(x_t)$ (Eq. 4.11). This second case is of special interest in decision under uncertainty where each t designates a state of nature and the representation reflects the general expected utility representation.

The preference conditions directly corresponding with our present value conditions are weaker (leading to stronger theorems) than the ones commonly used in the literature. First, our present value conditions relate to indifferences rather than preferences. Conditions for indifferences are logically weaker, making their implications logically stronger.[13]

An empirical advantage of our preference conditions is that they can be directly tested using statistical techniques such as analyses of variance and regressions. For example, if we take PV as the dependent variable, Eq. 4.3 predicts that φ, t, e_0, and e_1 may be significant predictors, but the e_j's with $j \neq 0, t$ are not. We can test this prediction using standard regression analyses. These allow us to use the sophisticated probabilistic error theories underlying econometric regressions, which are easier to use than the more recently developed error theories for preferences (Wilcox 2008). There is extensive data on the present values of future options in the financial market, which can be used to test the various independence conditions proposed in this paper. For individual choice, we are not aware of tests of preference axioms using regression techniques. Such tests become possible through the theorems presented here.

In the main text, we confined our analysis to periods with upper bound T. Many papers have studied extensions of representations to infinitely many periods. Usually, in the first stage representation results are established for finitely many periods. Then in the next stage, the extension to infinitely many periods, continuity conditions are added to avoid diverging or undefined summations. Such two-stage techniques can readily be used to extend

[13]Our conditions only involve the simplest tradeoffs possible, involving the change of the present outcome and one future outcome. This further enhances their generality. For example, commonly used stationarity conditions are more restrictive than the conditions in our Statements 4.4.2 and 4.5.2. The derivations of the full force preference conditions used in the literature from our preference conditions are based on known techniques (including Gorman 1968), and in this sense are not very innovative.

our results to infinitely many periods, where we can simply copy the second stage of previous analyses. An advanced general reference is Pivato (2014). Further references for nondiscounted utility include Alcantud and Dubey (2014), Basu and Mitra (2007), and Marinacci (1998); for constant discounted utility, see Harvey (1995), Bleichrodt, Rohde, and Wakker (2008), and Kopylov (2010); for time-dependent discounted utility, see Hubner and Suck (1993), Streufert (1995), and Wakker and Zank (1999).

We assumed that present values always exist, which implies that utility (in period 0) is unbounded from both sides. These restrictions can be dropped if we modify the preference conditions to hold only if all present values involved exist. In proofs of theorems, we first obtain the preference conditions in full force for every outcome sequence only in a neighborhood of that outcome sequence. This neighborhood is small enough to ensure that all present values required there exist. Next we combine these local representations into one global representation using the techniques of Chateauneuf and Wakker (1993). Their technique works for our most general model, time-dependent discounted utility and, hence, covers all cases considered in this paper.

The follow-up paper Keskin (2015) provides extensions of our results to some popular hyperbolic discount models, while still maintaining intertemporal separability. Extensions to more general intertemporal models are a topic for future research, as are further extensions to other optimization contexts.

4.9 Conclusion

We have introduced a new kind of preference condition for intertemporal choice, which requirings (quantitative) present values to be independent of particular other variables. The quantitative index should be directly observable, so that the independence requirements are observable preference conditions that can be directly tested qualitatively and can be used in theoretical preference axiomatization. Unlike usual preference conditions, our conditions and their independence requirements can also be directly tested quantitatively. Our conditions are more concise and transparent than conditions proposed before in the literature, and they are weaker, leading to stronger theorems. The technique of expressing preference conditions as independence conditions for directly observable quantitative indexes can be extended to other decision contexts such as decision under uncertainty to give new concise conditions that can easily be tested empirically.

4.10 References

Ahlbrecht, Martin and Martin Weber 1997. An Empirical Study on Intertemporal Decision Making under Risk. Management Science. 43, 813 - 826.

Alcantud, José C.R. and Ram Sewak Dubey 014. Ordering Infinite Utility Streams: Efficiency, Continuity, and no Impatience. Mathematical Social Sciences. 72, 33 - 40.

Attema, Arthur E. 012. Developments in Time Preference and Their Implications for Medical Decision Making. Journal of the Operational Research Society. 63, 1388 - 1399.

Basu, Kaushik and Tapan Mitra 007 . Utilitarianism for Infinite Utility Streams: A New Welfare Criterion and Its Axiomatic Characterization. Journal of Economic Theory. 133, 350 - 373.

Baucells, Manel and Franz H. Heukamp 012 . Probability and Time Tradeoff. Management Science. 58, 831 - 842.

Baucells, Manel and Rakesh K. Sarin 007 . Satiation in Discounted Utility. Operations Research. 55, 170 - 181.

Bleichrodt, Han, Kirsten I.M. Rohde, and Peter P. Wakker 008 . Koopmans' Constant Discounting for Intertemporal Choice: A Simplification and a Generalization. Journal of Mathematical Psychology. 52, 341 - 347.

Broome, John R. 1991. Weighing Goods. Basil Blackwell, Oxford.

Campbell, John Y. and Robert J. Shiller 1987 . Cointegration and Tests of Present Value Models. Journal of Political Economy. 95, 1062 - 1088.

Chateauneuf, Alain and Peter P. Wakker 1993 . From Local to Global Additive Representation. Journal of Mathematical Economics. 22, 523 - 545.

de Finetti, Bruno 1937. La Prévision: Ses Lois Logiques, ses Sources Subjectives. Annales de l'Institut Henri Poincaré. 7, 1 - 68.

de Wit, Johan 1671. Waardije van Lyf-Renten naer Proportie van Los-

Renten. The Worth of Life Annuities Compared to Redemption Bonds.

Debreu, Gérard 1960. Topological Methods in Cardinal Utility Theory. In. Kenneth J. Arrow, Samuel Karlin, and Patrick Suppes 1959, eds) Mathematical Methods in the Social Sciences., 1626. Stanford University Press, Stanford, CA.

Dolan, Paul and Daniel Kahneman 008. Interpretations of Utility and Their Implications for the Valuation of Health. Economic Journal. 118, 215 - 234.

École Polytechnique Fédérale de Lausanne 000-2013, Scala Programming Language, URL:

`http:\www.scala-lang.org`

Epper, Thomas, Helga Fehr-Duda, and Adrian Bruhin 011 . Viewing the Future through a Warped Lens: Why Uncertainty Generates Hyperbolic Discounting. Journal of Risk and Uncertainty. 43, 163 - 203.

Fisher, Irving 1930. The Theory of Interest. MacMillan, New York.

Frederick, Shane, George F. Loewenstein, and Ted O'Donoghue 002. Time Discounting and Time Preference: A Critical Review. Journal of Economic Literature. 40, 351 - 401.

Gold, Marthe R., Joanna E. Siegel, Louise B. Russell, and Milton C. Weinstein 1996. Cost-Effectiveness in Health and Medicine. Oxford University Press, New York.

Gorman, William M. 1968. The Structure of Utility Functions. Review of Economic Studies. 35, 367 - 390.

Gravel, Nicholas, Thierry Marchant, and Arunava Sen 012. Uniform Expected Utility Criteria for Decision Making under Ignorance or Objective Ambiguity. Journal of Mathematical Psychology. 56, 297 - 315.

Harvey, Charles M. 1995. Proportional Discounting of Future Costs and Benefits. Mathematics of Operations Research. 20, 381 - 399.

Hubner, Ronald and Reinhard Suck 1993. Algebraic Representation of Additive Structure with an Infinite Number of Components. Journal of Mathematical Psychology. 37, 629 - 639.

Hull, John C. 013. Options, Futures, and Other Derivatives. Englewood Cliffs, Prentice-Hall, NJ (9th edn.).

Ingersoll, Jonathan E. and Stephen A. Ross 1992. Waiting to Invest: Investment and Uncertainty. Journal of Business. 65, 1 - 29.

Jevons, W. Stanley 18711. The Theory of Political Economy. London.

Ju, Nengjiu and Jianjun Miao 012. Ambiguity, Learning, and Asset Returns. Econometrica. 80, 559 - 591.

Keller, L. Robin and Craig W. Kirkwood 1999. The Founding of INFORMS: A Decision Analysis Perspective. Operations Research. 47, 16 - 28.

Keskin, Umut 015. Characterizing Non-Classical Models of Intertemporal Choice by Present Values. mimeo; URL:

`http:\www.bilgi.edu.tr/site_media/uploads/staff/umut-keskin/publications/presentvalue-nonclassicalmodels.pdf`

Koopmans, Tjalling C. 1960. Stationary Ordinal Utility and Impatience. Econometrica. 28, 287 - 309.

Kopylov, Igor 010. Simple Axioms for Countably Additive Subjective Probability. Journal of Mathematical Economics. 46, 867 - 876.

Krantz, David H., R. Duncan Luce, Patrick Suppes, and Amos Tversky 1971. Foundations of Measurement, Vol. I (Additive and Polynomial Representations). Academic Press, New York.

Laibson, David I. 1997. Golden Eggs and Hyperbolic Discounting. Quarterly Journal of Economics. 112, 443 - 477.

LeRoy, Stephen F. and Richard D. Porter 1981. The Present-Value Relation: Tests Based on Implied Variance Bounds. Econometrica. 49, 555 - 574.

Loewenstein, George F. and Drazen Prelec 1993 . Preferences for Sequences of Outcomes. Psychological Review. 100, 91 - 108.

Luce, R. Duncan 2000 . Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches Lawrence Erlbaum Publishers, London.

Luhmann, Christian C. 013. Discounting of Delayed Rewards is not Hyperbolic. Journal of Experimental Psychology: Learning, Memory, and Cognition. 39, 1274 - 1279.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini 006. Dynamic Variational Preference. Journal of Economic Theory. 128, n4 - 44.

Marinacci, Massimo 1998. An Axiomatic Approach to Complete Patience and Time Invariance. Journal of Economic Theory. 83, 105 - 144.

Parfit, Derek 1984. Reasons and Persons. Clarendon Press, Oxford.

Pelsser, Antoon and Mitja Stadje 014. Time-Consistent and Market-Consistent Evaluations. Mathematical Finance. 24, 25 - 65.

Pigou, Arthur C. 1920. The Economics of Welfare (edn. 1952: MacMillan, London.)

Pivato, Marcus 014 . Additive Representation of Separable Preferences over Infinite Products. Theory and Decision. 73, 31 - 83.

Python Software Foundation 1990-2013. Python Reference Manual, URL: `http:\www.python.org`

R Core Team 013. . R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL:

`Http:\www.R-project.org/.`

Ramsey, Frank P. 1928. A Mathematical Theory of Saving. Economic Journal. 38, 543 - 559.

Rawls, John 1971. A Theory of Justice. Harvard University Press, Cambridge MA.

Samuelson, Paul A. 1937. A Note on Measurement of Utility. Review of Economic Studies. 4 (issue 2, February 1937, 155 - 161.

Savage, Leonard J. 1954. The Foundations of Statistics. Wiley, New York. (Second edition 1972, Dover Publications, New York.)

Smith, James E. 1998 . Evaluating Income Streams: A Decision Analysis Approach. Management Science. 44, 1690 - 1708.

Smith, James E. and Kevin F. McCardle 1999 . Options in the Real World: Lessons Learned in Evaluating Oil and Gas Investments. Operations Research. 47, 1 - 15.

Soman, Dilip, George Ainslie, Shane Frederick, Xiuping Li, John Lynch, Page Moreau, Andrew Mitchell, Daniel Read, Alan Sawyer, Yaacov Trope, Klaus Wertenbroch, and Gal Zauberman 005. The Psychology of Intertemporal Discounting: Why are Distant Events Valued Differently from Proximal Ones?. Marketing Letters. 16, 347 - 360.

Streufert, Peter A. 1995. A General Theory of Separability for Preferences Defined on a Countably Infinite Product Space. Journal of Mathematical Economics. 24, 407 - 434.

Strotz, Robert H. 1956. Myopia and Inconsistency in Dynamic Utility Maximization. Review of Economic Studies. 23 (issue 3, June 1956, 165 - 180.

Takeuchi, Kan 010,. Non-Parametric Test of Time Consistency: Present Bias and Future Bias. Games and Economic Behavior. 71, 456 - 478.

Tsuchiya, Aki and Paul Dolan 005. The QALY Model and Individual Preferences for Health States and Health Profiles over Time: A Systematic Review of the Literature. Medical Decision Making. 25, 460 - 467.

Wakker, Peter P. 1986 . The Repetitions Approach to Characterize Cardinal Utility. Theory and Decision. 20, 33 - 40.

Wakker, Peter P. 1989. Additive Representations of Preferences, A New Foundation of Decision Analysis. Kluwer, Dordrecht.

Wakker, Peter P. 010. Prospect Theory for Risk and Ambiguity. Cambridge University Press, Cambridge, UK.

Wakker, Peter P. and Horst Zank 1999. State Dependent Expected Utility for Savage's State Space; Or: Bayesian Statistics without Prior Probabilities. Mathematics of Operations Research. 24, 8 - 34.

Wilcox, Nathaniel T. 008. Stochastic Models for Binary Discrete Choice under Risk: A Critical Primer and Econometric Comparison. James C. Cox and Glenn W. Harrison, (eds.) Risk Aversion in Experiments.; Research in Experimental Economics 12, 197 - 292, Emerald, Bingley.

Chapter 5

Characterizing Non-Classical Models of Intertemporal Choice by Present Values

Umut Keskin[a]

[a]: Istanbul Bilgi University, Istanbul, Turkey

Abstract

When a sequence of outcomes needs to be evaluated for making decisions in an intertemporal setting, the common practice in economic theory is to compute the discounted utility of the sequence using constant discounting. Despite its prevalence, this particular functional form has frequently been challenged on empirical grounds. As a result, many alternative models have been proposed to explain violations of this model. This paper studies two such models, namely variation aversion and decreasing impatience, that can accommodate common violations of constant discounted utility. Earlier work on these models presented their characterizations in terms of preference conditions that can be cumbersome and hence difficult to use in experimental studies for testing the theories. It is shown that the functional forms in these models can more easily be characterized by conditions based on present value. Being a familiar and intuitive concept, present value simplifies the existing characterizations and is empirically easy to observe.

Keywords: present value, intertemporal choice, discounted utility, variation aversion, hyperbolic discounting, decreasing impatience.

5.1 Introduction

Since its introduction by Samuelson (1937), the constant discounted utility (CDU) function has been the most prominent tool for intertemporal choice in economics. Its particular functional form has two important properties: it is additively separable and uses constant discounting. Although analytically useful, both properties have been shown to be violated empirically (Loewenstein & Prelec (1992), Loewenstein & Prelec (1993), Loewenstein & Thaler (1989)). This paper studies two appealing alternatives to CDU, variation aversion Gilboa (1989) and decreasing impatience Prelec (2004), that can accommodate the violations of these two properties and gives characterizations of these models in terms of present value.

Present value is defined as follows: Assume that an agent is endowed with a stream of payments spread over a finite time. In this endowment, suppose that we change the amount in one of the future periods, keeping other periods' amounts the same. This gives a new endowment stream that differs from the original one in only one period. Then we ask how much change in the current period wealth within the original endowment would make the agent indifferent to this new stream. The answer is called the present value (PV).

To illustrate, this general definition entails the PV formula commonly used in financial bookkeeping. In this case, the present value of K dollars received in time t is $\beta^t K$, where β is usually taken to be $1/(1+r)$ with market interest rate r. It was shown by Bleichrodt et al. (2013) that this particular formulation is justifiable only for a particular type of preference once one adapts PV as described above. One important shortcoming of using prescribed formulas for present value is that it ignores consumer preferences. For example, the explicit expression for PV above does not allow the valuation of future monetary amounts to depend on current income or income in other periods; whereas the PV used in this paper is the subjective valuation of K dollars of time t as of today. Hence, it may depend on any relevant variable. It is a subjective measure that does not assume any particular form of preference for the agent.

The PV characterizations in this paper are intended to help researchers who carry out experiments that test the validity of the aforementioned models. Most experiments on intertemporal choice models are inevitably carried out from today's point of view. For example, subjects are asked to submit their evaluations of some future monetary changes as of today. The responses they give are direct revelations of how they perceive future outcomes and this is also what lies at the heart of the construction of PV: the idea of bringing future outcomes back to today for evaluation, and allowing for the most general form of doing so. Therefore as an empirical tool, PV is a natural concept to be used in experiments since decision makers are already familiar with the concept. Also, as the propositions below will make more clear, the simple behavioral conditions I state on PV are more compact and more easily grasped than the behavioral conditions imposed on preference relations. Hence the PV approach simplifies the analysis and the testing of theories to a great extent.

The idea of present value characterization instead of preference conditions first occurred in Bleichrodt et al. (2013) in which, decision models that are more commonly used in economics and finance were characterized. In this sense, I study some deviations from these benchmark models. At a time when political correctness was not a requirement for academic papers, Carver (1918) called behavior that is not in line with widely accepted norms of economics, "eternally feminine". Needless to say, I do not sympathize with the author's approach of associating any behavior that does not comply with certain norms, with femininity. Yet, as an analogy, this study is the "eternally feminine" counterpart of our previous work, Bleichrodt et al. (2013). In this respect, the intuitive PV characterizations below shed more light on the rationale behind the behavior described by the relevant models. This contributes to the discussions on the normative aspects of CDU model versus its particular violations that are analyzed in this paper.

5.2 The Model

Let $S = \{0, \ldots, n\}$ be the set of time periods with 0 being the current time. At each time point i, an individual receives a monetary outcome $x_i \in \mathbb{R}$. That makes our object of study the streams of payments $\{x_0, \ldots, x_n\}$. These streams will sometimes be called *prospects*. Agents are assumed to have a continuous, complete and transitive preference relation $\succsim$ on the set of all such prospects, $\mathbb{R}^{n+1}$. $\succ$ and $\sim$ are the asymmetric and symmetric parts of $\succsim$.

For any $x \in \mathbb{R}^{n+1}$, and $\alpha \in \mathbb{R}$, $\alpha_i x$ denotes the prospect obtained by replacing the i$^{\text{th}}$ component of x by α. Similarly, $\alpha_i \gamma_j x$ replaces the i$^{\text{th}}$ and j$^{\text{th}}$ components of x by α and γ respectively for $i \neq j$. In addition to being a continuous complete preorder, $\succsim$ will be assumed to satisfy *strict monotonicity* throughout the paper.In other words, the following will hold for $\succsim$:

For any two prospects $x, y \in \mathbb{R}^{n+1}$, if $x_i \geq y_i$ $\forall i \in S$ and $x_j > y_j$ for at least one $j \in S$, then $x \succ y$.

Present value is defined as follows.

Definition 5.1. Suppose that an individual with preference relation $\succsim$ on

$\mathbb{R}^{n+1}$ is initially endowed with an income stream $x \in \mathbb{R}^{n+1}$. Further, suppose that x is changed by ϕ units only in period i. Then π is called the individual's *present value* of ϕ if it satisfies the following condition:

$$(x_0 + \pi, x_1, \ldots x_i, \ldots, x_n) \sim (x_0, \ldots, x_i + \phi, \ldots, x_n).$$

By continuity, existence of π is guaranteed. Due to strict monotonicity, it is unique. In its most general form, the PV may depend on x, i, and ϕ, and such dependence will be denoted by $\pi_i(x, \phi)$. In the case where any of these variables is irrelevant for π, it will be omitted. It follows that the PV of any amount received in period 0 is itself i.e., $\pi_0(x, \phi) = \phi$.

Given a preference relation $\succsim$ on an arbitrary set X, a function $V : X \to \mathbb{R}$ is called a *representing function* for $\succsim$, if for all $x, \ y \in X$, $x \succsim y$ if and only if $V(x) \geq V(y)$. The CDU model assumes that the representing function for $\succsim$ in the above setting takes the following form:

$$\sum_{i=0}^{n} \beta^i u(x_i)$$

where, $0 < \beta < 1$ discounts future outcomes and u is a real valued function on monetary payoffs.

5.3 Variation Aversion

Consider the following example from Gilboa (1989). Preferences are defined on the set of income streams spread over four periods. In each period, the payment can be high (H) or low (L). Assume that the agent dislikes variation in her periodical payments and we observe the preferences below:

$$(H, H, L, L) \sim (L, L, H, H) \succ (H, L, H, L) \sim (L, H, L, H) \tag{5.1}$$

We observe the above preferences for such an agent because (H, H, L, L) and (L, L, H, H) contain less variation than (H, L, H, L) and (L, H, L, H) between periods. Such distaste for variation may be due to adjustment costs or psychological reasons. Especially when this variation effect is more pronounced and opportunity cost is low (for instance, due to low interest rate or closely spread time points), (5.1) can be observed Gilboa (1989) for a more detailed discussion of such preferences). It can be seen from (5.1) that the constant discounted utility model cannot accommodate these preferences. In fact, a closer investigation shows that no additively separable representation can be used in this case. In other words, even a representation of the general form $\sum_{i=0}^{3} u_i(x_i)$, where $x_i \in \{H, T\}$ cannot explain (5.1). In this section, I will analyze a model developed by Gilboa (1989) that can explain (5.1).

As the underlying reasoning suggests, any attempt to model preferences in (5.1) should take into account the variation in utility terms between periods, $|u(x_i) - u(x_{i-1})|$. Gilboa (1989) introduced a preference condition, the *variation preserving sure thing principle*, which leads to the following representation:

$$\sum_{i=0}^{n} (\lambda_i u(x_i) + \tau_i |u(x_i) - u(x_{i-1})|). \tag{5.2}$$

In (5.2), discounted utility is adjusted by a weighted sum of utility variations in each period. Once these variations are incorporated, (5.2) can explain distaste for fluctuations in income. I take Gilboa's model as benchmark and give a characterization of (5.2) in terms of present value.

In the context of choice under uncertainty, Savage (1954) imposed a preference condition called sure thing principle (STP) to guarantee a representation that is additively separable among states. Gilboa's (1989) variation preserving sure thing principle is a modification of this condition so that non-separability among adjacent periods is allowed. This axiom is presented next.

Subsets of S of the form $\{i, i+1, \ldots, j\}$ for $i \leq j$ are called *intervals* and are denoted by $[i, j]$.

Definition 5.2. Let $A = [i, j] \subset S$ be an interval and let $x, x', y, y' \in \mathbb{R}^{n+1}$ be such that

$$\begin{aligned} x_k = y_k, &\qquad x'_k = y'_k \qquad \forall k \in A \\ x_k = x'_k, &\qquad y_k = y'_k \qquad \forall k \in A^{\mathsf{c}} \\ x_k = x'_k = y_k = y'_k &\quad \text{ for } \ k = i-1,\ j+1. \end{aligned}$$

Then $\succsim$ is said to satisfy *variation preserving sure thing principle* if $x \succsim y$ if and only if $x' \succsim y'$.

This condition can be explained as follows. Suppose that x and y assume common values on a subset $A \subset S$. Then the preference between them does not change when the common values of these two prospects on A are replaced with a different common set of values, as long as -and this is where it differs from STP- this replacement does not alter the variation in x and y in different degrees. Define $x_{-1} = x_{n+1} = 0$. Gilboa (1989) gave the following characterization of (5.2).

Theorem 5.1. (Gilboa 1989) The following two statements are equivalent:

(i) The representing function for $\succsim$ is of the following form:

$$\sum_{i=0}^{n} (\lambda_i u(x_i) + \tau_i |u(x_i) - u(x_{i-1})|)$$

for all $x \in \mathbb{R}$, where $u : \mathbb{R} \to \mathbb{R}$ is unique up to an increasing affine transformation and

$$\begin{aligned} |\lambda_i| &\geq |\tau_i| + |\tau_{i+1}| \ \ \forall i < n \\ |\lambda_n| &\geq |\tau_n| \text{ and } \tau_0 = 0. \end{aligned}$$

(ii) $\succsim$ satisfies variation preserving sure thing principle.

In the above theorem, u is interpreted as the utility function for monetary outcomes. Since it is unique up to an increasing affine transformation, we can rescale it so as to get $u(0) = 0$. Then we have $u(x_{-1}) = u(x_{n+1}) = 0$. The same argument is also valid for the main result below, Proposition 5.1.

5.3.1 Main Result

Instead of characterizing (5.2) by variation preserving sure thing principle, I now present a simpler and empirically more operational axiomatization in terms of present value behavior. Let $\Delta(\alpha, \beta)$ denote the difference $\alpha - \beta$.

Proposition 5.2. The following two statements are equivalent:

(i) The representing function for $\succsim$ is of the following form:

$$\sum_{i=0}^{n} (\lambda_i u(x_i) + \tau_i |u(x_i) - u(x_{i-1})|)$$

for all $x \in \mathbb{R}$, where $u : \mathbb{R} \to \mathbb{R}$ is unique up to an increasing affine transformation and

$$\begin{aligned} |\lambda_i| &\geq |\tau_i| + |\tau_{i+1}| \quad \forall i < n \\ |\lambda_n| &\geq |\tau_n| \text{ and } \tau_0 = 0. \end{aligned}$$

(ii) The present value so constructed in Definition 2.1 depends on $\phi, i, x_0, x_i, \Delta(x_1, x_0), \Delta(x_i, x_{i-1})$ and $\Delta(x_{i+1}, x_i)$:

$$\pi = \pi_i(\phi, x_0, x_i, \Delta(x_1, x_0), \Delta(x_i, x_{i-1}), \Delta(x_{i+1}, x_i)).$$

For the result above, I will further assume[1] that u is infinitely differebtiable.

[1] Debreu (1972) gives behavioral conditions on differentiability.

Proof. See Appendix A. □

In Bleichrodt et al. (2013), it was shown that $\succsim$ can be represented by an additively separable function $\sum u_i$ if and only if the agent's present value depends on the amount given in the future, his current income, his income at the period that the change happens and how far this period is (i.e., $\pi = \pi_i(\phi, x_0, x_i)$). Thus for the agent described above, in addition to these factors, variations that take place with respect to next period income and the income right before and after the future distortion also affect the perception of present value. And this is all that is needed to characterize such an agent. Compared to the variation preserving sure thing principle, this is a shorter, easier and more intuitive condition. Therefore, it clarifies the rationale behind such preferences. This section focused on the adjacent period separability. According to Proposition 5.2, as long as an agent takes the neighboring periods' income levels into consideration in her evaluation of future income changes, she deviates from the standard CDU. It would be difficult to claim that such consideration is irrational in any sense. Hence, with this clarification it becomes easier to argue that additive separability in intertemporal choice models is hard to defend from a normative point of view.

5.4 Decreasing Impatience

In this section, I will put my framework in a broader context by extending the results to a model developed by Prelec (2004). In addition to its additively separable form, one of the most distinctive features of CDU is that the degree of impatience for a given length of period stays the same no matter how far this period is from today. More formally, if $x \in \mathbb{R}^{n+1}$, $\alpha, \beta \in \mathbb{R}$ satisfy

$\alpha_i 0_{i+k} x \sim 0_i \beta_{i+k} x$ for some $i \in S$, then $\alpha_{i'} 0_{i'+k} x \sim 0_{i'} \beta_{i'+k} x$ for all $i' \in S$ as long as $i' + k \leq n$. For example, if you are indifferent between 10 dollars today and 11 dollars in one week; then according to CDU, you should be indifferent between 10 dollars in eight weeks and 11 dollars in nine weeks.

This property of CDU has been challenged on empirical grounds (Thaler (1981), Loewenstein & Prelec (1992)). These studies reveal that people do prefer sooner but worse outcomes to better but later outcomes when the comparisons are made for the near future. However, this preference reverses as the comparisons are delayed and people accept to wait more to obtain the better outcome. One way to explain such preferences has been to assert that people tend to decrease their rate of discount for farther away future outcomes. A particular functional form that incorporated this idea is the quasi hyperbolic discounting utility function which is also one of the most commonly used models as alternatives to CDU. In this model, any stream $x \in \mathbb{R}^{n+1}$ is evaluated by $u(x_0) + \beta \sum_{i \in S \setminus \{0\}} \delta^i u(x_i)$. It has first been used by Phelps & Pollak (1968) and was further developed by Laibson (1997). Numerous studies employed this functional form to explain economic phenomena such as procrastination (O'Donoghue & Rabin (1999)) and addiction (Gruber & Koszegi(2000)). It is a special case of a more general class of preferences axiomatized by Prelec (2004), the decreasing impatience (DI) preferences.[2]

Many studies on hyperbolic discounting and DI, including Prelec (2004), deal with simple prospects (i.e. two dated outcomes x and y received in periods t and s, denoted (x, t) and (y, s) respectively) but I study sequences

[2]Although commonly used in many studies nowadays, hyperbolic discounting or its generalizations are not the only models that can accommodate the violations mentioned above (see Rubinstein (2003)) for an alternative explanation and a critique of hyperbolic discounting).

of outcomes instead of such dated outcomes.

Prelec (2004) assumes that preferences are representable by the function $\sum_{i=0}^{n} \Phi(i)u(x_i)$, where $\Phi : \mathbb{R} \to \mathbb{R}$ is the discount function and $u : \mathbb{R} \to \mathbb{R}$ is the utility function for monetary payoffs in each period. For the rest of this section, this assumption will be maintained. The main result in Prelec (2004) is that $\ln(\Phi(i))$ is convex (in other words proportional changes in time lead to less and less discounting) if and only if the agent is decreasingly impatient. DI is defined as follows (generalized to the sequences of outcomes setting):

Definition 5.3. $\succsim$ exhibits DI if for all $i, j, k \in S$ with $i < j,\ i+k,\ j+k \in S$, $\alpha > \beta$ and $x \in \mathbb{R}^{n+1}$, $0_i\alpha_j x \sim \beta_i 0_j x$ implies $0_i\alpha_{j+k}x \succsim \beta_{i+k}0_j x$ (strict DI, if the latter preference is strict).

We have the following result from Prelec (2004) adapted to the $n-$period setting:

Theorem 5.3. (Prelec 2004) The following two statements are equivalent:

1. $\succsim$ satisfies DI.

2. $\ln(\Phi(i))$ is convex.

When $\succsim$ is representable by a function of the form $\sum_{i=0}^{n} \Phi(i)u(x_i)$, using the definition of PV one can see that the present value of any amount ϕ given at time i under an endowment x depends only on i, ϕ, x_0 and x_i:

$$\pi = \pi_i(\phi, x_0, x_i)$$

In accordance with the unified PV characterization system, I will provide an alternative axiomatization as follows: In accordance with the unified PV characterization system, I will provide an alternative axiomatization as follows:

Proposition 5.4. The following two statements are equivalent:

1. The present value so constructed in Definition 5.1 satisfies the following:

   For all $x \in \mathbb{R}^{n+1}$, $i, j, k \in S$ with $i < j$, $i + k$, $j + k \in S$ and $\phi, \phi' \in \mathbb{R}$ with $\phi > \phi'$, if $\pi_j(\phi, x_0, 0) = \pi_i(\phi', x_0, 0)$, then $\pi_{j+k}(\phi, x_0, 0) \geq \pi_{i+k}(\phi', x_0, 0)$.

2. $\ln(\Phi(i))$ is convex.

Proof. See Appendix B □

Therefore for those who are decreasingly impatient, the present value depends on the amount given in the future, their current income, their income at the period that change happens and also how far the change is. A DI agent behaves as follows: While today's worth of later but better outcome is relatively small compared to the worth of sooner but worse outcome, this changes as time elapses and he leans towards the better outcome since its worth today improves relative to the sooner but worse outcome.

5.5 Conclusion

I presented a generalized present value foundation for two non-classical models of intertemporal choice, variation aversion and decreasing impatience. Although my results are mainly analytical, my goal is to provide a natural road map for experimental researchers who wish to test the validity of intertemporal decision models. In this respect, I presented easily implementable and testable conditions in terms of present values for these models instead of less tractable preference conditions.

5.6 Appendices

Appendix 5.A: Proof of Proposition 5.2

Recall that present value, π, is defined through the indifference

$$(x_0 + \pi)_0 x_i x \sim x_0 (x_i + \phi)_i x. \tag{5.3}$$

Suppose that the representing function for $\succsim$ is as in (i) in Proposition 5.2. Then for $i < n$, (5.3) implies

$$\begin{aligned}\lambda_0 u(x_0+\pi) + \tau_1 |u(x_1) - u(x_0+\pi)| + \lambda_i u(x_i) +\\ \tau_i |u(x_i) - u(x_{i-1})| + \tau_{i+1}|u(x_{i+1}) - u(x_i)|\\ = \lambda_0 u(x_0) + \tau_1 |u(x_1) - u(x_0)| + \lambda_i u(x_i+\phi) +\\ \tau_i |u(x_i+\phi) - u(x_{i-1})| + \tau_{i+1}|u(x_{i+1}) - u(x_i+\phi)|.\end{aligned} \tag{5.4}$$

Showing that (ii) holds is a matter of solving (5.4) for π for different cases. I believe that this arithmetic is not crucial for our purposes. Therefore I will present the result for only one illustrative case:

$$u(x_0) \geq 0, \quad \phi > 0, \quad x_i \geq x_{i-1}, \quad x_{i+1} \geq x_i + \phi, \tag{5.5}$$

and

$$\frac{D_i}{D_0}(u(x_i+\phi) - u(x_i)) \leq u(x_1) - u(x_0) \tag{5.6}$$

where $D_i = \lambda_i + \tau_i - \tau_{i+1}$, and $D_0 = \lambda_0 - \tau_1$. We need (5.6) to determine the sign of $u(x_1) - u(x_0 + \pi)$ in (4). From (5.4), we get

$$D_0 u(x_0+\pi) + D_i u(x_i) = D_0 u(x_0) + D_i u(x_i+\phi)$$

Note that, $D_0 \neq 0$, for otherwise one would have $D_i = 0$ too and this would lead (5.3) to hold true for infinitely many values of π; which violates monotonicity. Hence, D_i/D_0 is well defined. Then with the assumptions made in (5.5) and (5.6), we have

$$\pi = u^{-1}(u(x_0) + \frac{D_i}{D_0}[u(x_i + \phi) - u(x_i)]) - x_0. \tag{5.7}$$

In (5.7), π depends on x_0, x_i and ϕ explicitly, and on i through D_i in the formula. Dependence on $\Delta(x_{i+1}, x_i)$ and $\Delta(x_i, x_{i-1})$ is due to the conditions assumed in (5.5). In (5.6), we see dependence of π on $u(x_1) - u(x_0)$. Using a Taylor expansion around x_0, we translate this into a dependence on $\Delta(x_1, x_0)$ (Recall that we assumed infinite differentiability). We have

$$u(x_1) - u(x_0) = \sum_{j=1}^{\infty} \frac{u^{(j)}(x_0)}{j!}[\Delta(x_1, x_0)]^j$$

where, $u^{(j)}(x_0)$ is the j^{th} derivative of u at x_0. Similarly,

$$u(x_i - \phi) - u(x_i) = \sum_{j=1}^{\infty} \frac{u^{(j)}(x_i)}{j!}\phi^j.$$

So in terms of explanatory variables for π, (5.6) gives us a dependence on $\Delta(x_1, x_0)$, x_0, x_i and ϕ. The case for $i = n$ is investigated in the same manner, only noting that by definition $x_{n+1} = 0$. Therefore,

$$\pi = \pi_i(\phi, x_0, x_i, \Delta(x_1, x_0), \Delta(x_i, x_{i-1}), \Delta(x_{i+1}, x_i)).$$

Conversely, assume that

$$\pi = \pi_i(\phi, x_0, x_i, \Delta(x_1, x_0), \Delta(x_i, x_{i-1}), \Delta(x_{i+1}, x_i)).$$

I will show that $\succsim$ satisfies variation preserving sure thing principle which will imply that it can be represented by (5.2). Let $A = [i, j] \subset S$ be an interval and $x, x', y, y' \in \mathbb{R}^{n+1}$ be such that

$$\begin{aligned} x_k = y_k, \qquad & x'_k = y'_k \qquad \forall k \in A \\ x_k = x'_k, \qquad & y_k = y'_k \qquad \forall k \in A^{\mathsf{c}} \\ x_k = x'_k = y_k = y'_k \quad & \text{for} \quad k = i-1, \ j+1. \end{aligned}$$

Assume that $j+1<n$ and suppose $x \succsim y$. First I will consider indifference, $x \sim y$. I *discount* x_ks to period 0 one by one. For this purpose, let us define the recursive sequence of p_ks as follows:

$$
\begin{aligned}
p_0 &= x_0 \\
p_1 &= p_0 + \pi_n(x_n, p_0, 0, \Delta(x_1, p_0), \Delta(0, x_{n-1}), \Delta(x_{n+1}, 0)) \\
p_2 &= p_1 + \pi_{n-1}(x_{n-1}, p_1, 0, \Delta(x_1, p_1), \Delta(0, x_{n-2}), \Delta(0, 0)) \\
&\vdots \\
p_k &= p_{k-1} + \\
&\pi_{n+1-k}(x_{n+1-k}, p_{k-1}, 0, \Delta(x_1, p_{k-1}), \Delta(0, x_{n-k}), \Delta(0, 0)). \\
&\vdots
\end{aligned}
$$

for $0 \leq k \leq n-(j+2)$. Also, let

$$
\begin{aligned}
\tilde{p}_1 &= p_{n-(j+2)} + \\
&\pi_{i-2}(x_{i-2}, p_{n-(j+2)}, 0, \Delta(x_1, p_{n-(j+2)}), \Delta(0, x_{i-3}), \Delta(x_{i-1}, 0)) \\
\tilde{p}_2 &= \tilde{p}_1 + \pi_{i-3}(x_{i-3}, \tilde{p}_1, 0, \Delta(x_1, \tilde{p}_1), \Delta(0, x_{i-4}), \Delta(0, 0)) \\
\tilde{p}_3 &= \tilde{p}_2 + \pi_{i-4}(x_{i-4}, \tilde{p}_2, 0, \Delta(x_1, \tilde{p}_2), \Delta(0, x_{i-5}), \Delta(0, 0)) \\
&\vdots \\
\tilde{p}_{i-3} &= \tilde{p}_{i-4} + \pi_2(x_2, \tilde{p}_{i-4}, 0, \Delta(x_1, \tilde{p}_{i-4}), \Delta(0, x_1), \Delta(0, 0)) \\
\tilde{p}_{i-2} &= \tilde{p}_{i-3} + \pi_1(x_1, \tilde{p}_{i-3}, 0, \Delta(0, \tilde{p}_{i-3})\Delta(\tilde{p}_{i-3}, 0)\Delta(0, 0)).
\end{aligned}
$$

We apply the same procedure to y. I will call the resulting values p' and $\tilde{p}'$; analogous to p and $\tilde{p}$ defined above. For notational simplicity, let $\tilde{p}_{i-2} = \Pi$ and $\tilde{p}'_{i-2} = \Pi'$. Using the definition of present value, we get

$$
\begin{aligned}
x &\sim (p_1, x_1, \ldots, x_{n-1}, 0) \\
&\sim (p_2, x_1, \ldots, x_{n-2}, 0, 0).
\end{aligned}
$$

Iteratively *discounting* each x_k for $k \in A^{\mathsf{c}} \setminus \{0, i-1, j+1, j+2\}$, we obtain

$$x \sim (\Pi, 0, \ldots, 0, x_{i-1}, \ldots, x_{j+2}, 0 \ldots, 0). \tag{5.8}$$

Similarly,

$$y \sim (\Pi', 0, \ldots, 0, y_{i-1}, \ldots, y_{j+2}, 0 \ldots, 0). \tag{5.9}$$

Rewriting Π and y_{j+2} as $\Pi = \Pi' + (\Pi - \Pi')$ and $y_{j+2} = x_{j+2} + (y_{j+2} - x_{j+2})$ respectively and using (8) and (9), we get

$$(\Pi' + (\Pi - \Pi'), 0, \ldots, 0, x_{i-1}, \ldots, x_{j+2}, 0 \ldots, 0) \sim$$
$$(\Pi', 0, \ldots, 0, y_{i-1}, \ldots, y_{j+1}, x_{j+2} + (y_{j+2} - x_{j+2}), 0 \ldots, 0).$$

Then using the fact that $y_k = x_k$ for $k = i-1, \ldots, j+1$ results in

$$(\Pi' + (\Pi - \Pi'), 0, \ldots, 0, x_{i-1}, \ldots, x_{j+2}, 0 \ldots, 0) \sim$$
$$(\Pi', 0, \ldots, 0, x_{i-1}, \ldots, x_{j+1}, x_{j+2} + (y_{j+2} - x_{j+2}), 0 \ldots, 0).$$

By the definition of present value, this means

$$\pi_{j+2}(y_{j+2} - x_{j+2}, \Pi', 0, x_{j+1}, x_{j+2}, 0) = \Pi - \Pi'.$$

Since $\pi_{j+2}(.)$ is independent of $x_i, \ldots, x_j$, we can replace these by $x'_i, \ldots, x'_j$:

$$(\Pi' + (\Pi - \Pi'), 0, \ldots, 0, x_{i-1}, x'_i, \ldots, x'_j, x_{j+1}, x_{j+2}, 0 \ldots, 0) \sim$$
$$(\Pi', 0, \ldots, 0, x_{i-1}, x'_i, \ldots, x'_j, x_{j+1}, x_{j+2} + (y_{j+2} - x_{j+2}), 0 \ldots, 0).$$

Noting again that $y'_k = x'_k$ for $k = i-1, \ldots, j+1$, we have

$$(\Pi' + (\Pi - \Pi'), 0, \ldots, 0, x_{i-1}, x'_i, \ldots, x'_j, x_{j+1}, x_{j+2}, 0 \ldots, 0) \sim$$
$$(\Pi', 0, \ldots, 0, y_{i-1}, y'_i, \ldots, y'_j, y_{j+1}, x_{j+2} + (y_{j+2} - x_{j+2}), 0 \ldots, 0).$$

Recall that Π and Π' are constructed by discounting each x_k and y_k to period 0. Now we apply the reverse operation to them and forward these components back to their original positions. This gives

$$(\Pi' + (\Pi - \Pi'), 0, \ldots, 0, x_{i-1}, x'_i, \ldots, x'_j, x_{j+1}, x_{j+2}, 0 \ldots, 0) \sim x'$$

and

$$(\Pi', 0, \ldots, 0, y_{i-1}, y'_i, \ldots, y'_j, y_{j+1}, x_{j+2} + (y_{j+2} - x_{j+2}), 0 \ldots, 0) \sim y'.$$

Hence,

$$x' \sim y'$$

as we wanted to show.

If $x \succ y$, then by solvability we can find $\varepsilon > 0$, such that $x \sim (y_0 + \varepsilon, g_1, \ldots, y_n)$. Afterwards, the same argument above is repeated to show that $x' \succ y'$.

If $j + 1 = n$, then we start rolling x_ks and y_ks back to present time from period $i - 2$ to obtain Π and Π'. Once Π and Π' are defined as such, the rest of the proof follows the same steps as above, and once again we obtain $x' \succsim y'$. $\square$

Appendix 5.B: Proof of Proposition 5.4

Suppose that statement i holds in the proposition. Recall that when the preferences are representable by a a function of the form $\sum_{i=0}^{n} \Phi(i)u(x_i)$, π depends on i, ϕ, x_0 and x_i in the context of Definiton 5.1. Now take any $x \in \mathbb{R}^{n+1}$, $i, j, k \in S$ with $i < j,\ i + k,\ j + k \in S$ and $\phi, \phi\prime \in \mathbb{R}$ with $\phi > \phi\prime$ that satisfies $\pi_j(\phi, x_0, 0) = \pi_i(\phi', x_0, 0)$. Then by the definition of PV, one has

$$\begin{aligned} (x_0 + \pi_j(\phi, x_0, 0))_0 0_i x &\sim \phi_j x \quad \text{and} \\ (x_0 + \pi_i(\phi', x_0, 0))_0 0_i x &\sim \phi'_i x. \end{aligned}$$

Since $\pi_j(.)$ does not depend on x_i and $\pi_i(.)$ does not depend on x_j, the indifference relations above can be rewritten as

$$\begin{aligned}(x_0 + \pi_j(\phi, x_0, 0))_0 0_j 0_j x &\sim 0_i \phi_i x \quad \text{and} \\ (x_0 + \pi_i(\phi', x_0, 0))_0 0_i 0_j x &\sim \phi'_i 0_j x.\end{aligned}$$

Then, since $\pi_j(\phi, x_0, 0) = \pi_i(\phi', x_0, 0)$ we have

$$0_i \phi_j x \sim \phi'_i 0_j x.$$

Statement (i) in Proposition 5.4 implies

$$\pi_{j+k}(\phi, x_0, 0) \geq \pi_i(\phi', x_0, 0).$$

Using the same argument above, we get

$$0_i \phi_{j+k} x \succsim \phi'_{i+k} 0_j x.$$

Therefore, DI holds which implies that $\ln(\Phi(i))$ is convex.

Conversely, suppose that $\ln(\Phi(i))$ is convex. Then DI holds by Prelec's (2004) theorem. Applying the first part of the proof from the reverse order, we can see that statement i holds. $\square$

5.7 References

Bleichrodt, A., Keskin, U., Rohde, K., Spinu, V., & Wakker, P. 2013. Discounted Utility and Present Value - A Close Relation, forthcoming in Operations Research

Debreu, G. 1972. Smooth Preferences. Econometrica: Journal of the Econo- metric Society. 603-615.

Gilboa, I. 1989. Expectation and Variation in Multi-period Decisions. Econo- metrica: Journal of the Econometric Society. 1153 -169.

Gruber, J., & Koszegi, B. 2000. Is Addiction Rational: Theory and Evidence (Tech. Rep.). National Bureau of Economic Research.

Laibson, D. 1997. Golden Eggs and Hyperbolic Discounting. The Quarterly Journal of Economics. 112 2, 443-478.

Loewenstein, G., & Prelec, D. 1992. Anomalies in Intertemporal Choice: Evidence and an Interpretation. The Quarterly Journal of Economics. 107 2, 573-597.

Loewenstein, G., & Prelec, D. 1993. Preferences for Sequences of Outcomes. Psychological Review. 100 1, 91.

Loewenstein, G., & Thaler, R. 1989. Anomalies: Intertemporal Choice. The Journal of Economic Perspectives. 3 , 181-193.

O'Donoghue, T., & Rabin, M. 1999. Doing It Now or Later. American Economic Review. 103-124.

Phelps, E., & Pollak, R. 1968. On Second-best National Saving and Game- equilibrium Growth. The Review of Economic Studies. 35 2, 185-199.

Prelec, D. 2004. Decreasing Impatience: A Criterion for Non-stationary Time Preference and Hyperbolic Discounting. The Scandinavian Journal of Economics. 106. 511-532.

Rubinstein, A. 2003. Economics and Psychology? The Case of Hyperbolic Discounting. International Economic Review. 44 4. 1207-1216.

Samuelson, P. 1937. A Note on Measurement of Utility. The Review of Economic Studies. 4 2. 155-161.

Chapter 6

Conclusion

This thesis consists of four papers on consumers' decision making under uncertainty and intertemporal choice. In the first paper, we set up a lab experiment to analyze how subjects update their beliefs and change their ambiguity attitudes upon receipt of new information regarding the uncertainty. We decomposed subjects' ambiguity attitudes into two parts: pessimism and likelihood insensitivity. We found that the effect of new information was relatively more pronounced in the likelihood insensitivity part with a considerable decrease compared to a less significant affect on pessimist behavior. Our results suggest that subjects deviate from expected utility for any level of information, but we did observe that they become closer to being expected utility maximizers as they received more information.

In the second paper, we placed the notion of independence in a decision theoretic context and stated preference conditions through this new construct to axiomatize decision rules for uncertainty. We showed that the symmetry of independence (along with standard conditions) is necessary and sufficient for the Bayesian model. This implies that many commonly used Non-Bayesian decision rules cannot accommodate symmetry of independence. We discussed the implications of these impossibility results. We also showed that the two stage model of Anscombe & Aumann would better fit the theory if the order of stages were reversed.

The third paper was about several models of intertemporal choice popular in economics and finance. Most of the empirical studies on intertemporal choice involve eliciting consumer's present values of future payments. Therefore we introduced preference conditions purely based on subjective present values to characterize the aforementioned models of intertemporal choice. These new characterizations are more clear and natural because they are built on something that subjects are more familiar with, instead of more complex structures.

In the last paper, we extended our results from the third paper to a broader context. The third paper had characterized more commonly used rational choice models, whereas the last paper used the principles we had developed there to analyze frequently observed empirical departures from those rational models.

Through our results from these four papers mentioned above, we provided new insights into individual behavior and its (ir)rationality within the context of uncertainty and dynamic settings.

Chapter 7

Summary

In this thesis we provided new results in two important branches of individual behavior in economics, namely choice under uncertainty and intertemporal choice. Specifically we focused on rationality in these two fields by means of four papers.

First we developed a method to decompose beliefs and ambiguity attitudes. Through this separation, we were able to provide a theoretical analysis of the effect of receiving an uncertainty resolving piece of information on both parts in isolation. Then to test our model empirically, we set up a lab experiment to see how these components get affected by the arrival of new information separately. In our experimental framework, we elicited subjects' ask prices of options that are Initial Public Offerings for three different information conditions regarding the past returns of the relevant option. We found that pessimism component was relatively unaffected, whereas likelihood insensitivity diminished as more information about the historical performance of the stocks became available. We observed that the estimated beliefs, when corrected for ambiguity attitudes, converged to true frequencies. Subjects moved in the direction of subjective expected utility as more information was provided, however substantial deviations remained even in the maximum information condition.

In our second paper, our focus was on an important aspect of decision under uncertainty: independence. We stated preference conditions capturing independence in a statistical sense, and examined independence in various models of decision making under uncertainty. Leaving aside the standard technical preference conditions, we showed that symmetry of independence is a necessary and sufficient condition for Bayesian decision rule. Although the symmetry of this basic notion appears to be natural, we showed that it is in fact quite restrictive in the sense that no other decision rule can accommodate it. Nonsymmetric independencies can be applied to non-Bayesian (ambiguity) models, where we derive their implications. In particular, these implications reveal a problem for the two stage Anscombe-Aumann framework, which is one of the most common frameworks in the ambiguity literature today. We show that this problem can be avoided if we simply reverse the order of stages in this framework.

In the third paper of the thesis, we studied intertemporal choice. Existing characterizations of intertemporal choice models involve axioms that are not directed towards their empirical testability due to their theoretical nature. Considering the fact that models have to be tested for their appropriateness and accuracy, we presented new preference conditions based on a very natural and intuitive concept, the present value, which serve this purpose. Lies in the heart of our new characterizations is the independence this present value from other relevant factors and variables. We showed how similar types of preference conditions, imposing independence conditions between directly observable quantities, can be developed for other multi-criteria optimization problems and can simplify behavioral axiomatizations there.

In our last paper, we provide an extension of our results from the third paper. Despite its prevalence, constant discounted utility has frequently been challenged on empirical grounds. As a result, many alternative models have been proposed to explain violations of this model. The fourth paper of the thesis studied two such models, namely variation aversion and decreasing impatience, that can accommodate common violations of constant discounted utility. It was shown that the functional forms in these models can easily be characterized by conditions based on present value.

Chapter 8

Dutch Summary

Dit proefschrift geeft nieuwe resultaten voor twee belangrijke gebieden van individueel keuzegedrag in de economie, namelijk: keuzes onder onzekerheid en intertemporele keuzes. In het bijzonder besteedt het aandacht aan rationaliteit binnen deze twee gebieden. Dit proefschrift bestaat uit vier artikelen.

Ten eerste wordt een methode ontwikkeld om geloof te ontbinden in twee componenten: subjectieve kans en attitude t.o.v. ambiguteit. Door deze ontbinding wordt het mogelijk een theoretische analyse te geven van de gesoleerde effecten van het ontvangen van nieuwe informatie op de beide componenten. Vervolgens, om de methode empirisch te toetsen, is een laboratorium experiment opgezet om de genoemde gesoleerde effecten te meten. Vraagprijzen van proefpersonen worden gemeten van opties op Initial Public Offerenings, voor drie verschillende niveaus van informatie over opbrengsten uit het verleden van de relevante opties. We vinden dat de pessimistische (attitude) component vrijwel niet benvloed wordt, maar dat de likelihood sensitivity (betreffende geloof) component duidelijk gereduceerd wordt naarmate meer informatie beschibaar komt. De geschatte subjectieve kansen convergeren, na correctie voor ambiguteits attitude, naar de ware frequenties.

Proefpersonen komen steeds dichter bij het klassieke subjectief verwachte nut naarmate meer informatie verschaft wordt, maar er blijven substantile afwijkingen zelfs bij het hoogste niveau van informatie.

Het tweede paper bekijkt een belangrijk aspect van beslissen bij onzekerheid: onafhankelijkheid. Preferentie-condities worden gegeven die onafhankelijkheid in de statistische zin bepalen, en onafhankelijkheid in diverse modellen van beslissen bij onzekerheid. De standaard technische condities hier buiten beschouwing latend, wordt getoond dat symmetrie van onafhankelijkheid een nodige en voldoende voorwaarde is voor Bayesiaanse beslissingsregels. Hoewel symmetrie van dit fundamentele concept natuurlijk lijkt, volgt in feite dat het beperkend is in de zin dat geen andere beslisregel er aan kan voldoen. Niet-symmetrische onafhankelijkheden kunnen worden toegepast in niet-Bayesiaanse (ambiguteits) modellen, waarvoor enige gevolgen worden afgeleid. In het bijzonder brengen deze gevolgen een probleem aan het licht voor het twee-staps Anscombe-Aumann model, n van de populairste modellen voor ambiguteit. Het blijkt dat dit probleem vermeden kan worden als we de volgorde van de twee stappen omkeren.

Het derde paper van dit proefschrift onderzoekt intertemporele keuzes. Bestaande karakterisaties van intertemporele keuze-modellen gebruiken axiomas die niet sterk gericht zijn op hun empirische toetsbaarheid. Omdat modellen getoetst moeten worden op hun gepastheid en nauwkeurigheid, levert dit paper nieuwe preferentie-condities die gebaseerd zijn op een zeer natuurlijk en intutief concept: de contante waarde. Centraal in deze nieuwe karakterisaties is de onafhankelijkheid van deze contante waarde van andere factoren en variabelen. Getoond wordt hoe vergelijkbare typen van preferentie condities, onafhankelijkheden opleggend aan direct observeerbare quantitatieve grootheden, ontwikkeld kunnen worden voor andere multi-criteria optimalisatie problemen en daar kunnen dienen om preferentie-axiomatiseringen te vereenvoudigen.

Het laatste paper geeft een uitbreiding van de resultaten van het derde paper. Zijn wijd verbreide toepassingen niettegenstaande, is het model van constante discontering vaak bekritiseerd vanwege zijn vele empirische schendingen. Dientengevolge zijn veel alternatieve modellen voorgesteld om de schendingen te verklaren. Het vierde paper van die proefschrift bestudeert twee van zulke alternatieve modellen, namelijk variatie-afkeer en dalende ongeduldigheid, welke de schendingen van constante discontering kunnen verklaren. Deze vormen kunnen gemakkelijk gekarakteriseerd worden door condities gebaseerd op contante waarden.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

635 P. MULLER, *Labor Market Policies and Job Search*
636 N. KETEL, *Empirical Studies in Labor and Education Economics*
637 T.E. YENILMEZ, *Three Essays in International Trade and Development*
638 L.P. DE BRUIJN, *Essays on Forecasting and Latent Values*
639 S. VRIEND, *Profiling, Auditing and Public Policy: Applications in Labor and Health Economics*
640 M.L. ERGUN, *Fat Tails in Financial Markets*
641 T. HOMAR, *Intervention in Systemic Banking Crises*
642 R. LIT, *Time Varying Parameter Models for Discrete Valued Time Series*
643 R.H. KLEIJN, *Essays on Bayesian Model Averaging using Economic Time Series*
644 S. MUNS, *Essays on Systemic Risk*
645 B.M. SADABA, *Essays on the Empirics of International Financial Markets*
646 H. KOC, *Essays on Preventive Care and Health Behaviors*
647 V.V.M. MISHEVA, *The Long Run Effects of a Bad Start*
648 W. LI, *Essays on Empirical Monetary Policy*
649 J.P. HUANG, *Topics on Social and Economic Networks*
650 K.A. RYSZKA, *Resource Extraction and the Green Paradox: Accounting for Political Economy Issues and Climate Policies in a Heterogeneous World*
651 J.R. ZWEERINK, *Retirement Decisions, Job Loss and Mortality*
652 M. K. KAGAN, *Issues in Climate Change Economics: Uncertainty, Renewable Energy Innovation and Fossil Fuel Scarcity*
653 T.V. WANG, *The Rich Domain of Decision Making Explored: The Non-Triviality of the Choosing Process*
654 D.A.R. BONAM, *The Curse of Sovereign Debt and Implications for Fiscal Policy*
655 Z. SHARIF, *Essays on Strategic Communication*
656 B. RAVESTEIJN, *Measuring the Impact of Public Policies on Socioeconomic Disparities in Health*
657 M. KOUDSTAAL, *Common Wisdom versus Facts; How Entrepreneurs Differ in Their Behavioral Traits from Others*
658 N. PETER, *Essays in Empirical Microeconomics*
659 Z. WANG, *People on the Move: Barriers of Culture, Networks, and Language*
660 Z. HUANG, *Decision Making under Uncertainty-An Investigation from Economic and Psychological Perspective*

661 J. CIZEL, *Essays in Credit Risk, Banking, and Financial Regulation*

662 I. MIKOLAJUN, *Empirical Essays in International Economics*

663 J. BAKENS, *Economic Impacts of Immigrants and Ethnic Diversity on Cities*

664 I. BARRA, *Bayesian Analysis of Latent Variable Models in Finance*

665 S. OZTURK, *Price Discovery and Liquidity in the High Frequency World*

666 J. JI, *Three Essays in Empirical Finance*

667 H. SCHMITTDIEL, *Paid to Quit, Cheat, and Confess*

668 A. DIMITROPOULOS, *Low Emission Vehicles: Consumer Demand and Fiscal Policy*

669 G.H. VAN HEUVELEN, *Export Prices, Trade Dynamics and Economic Development*

670 A. RUSECKAITE, *New Flexible Models and Design Construction Algorithms for Mixtures and Binary Dependent Variables*

671 Y. LIU, *Time-varying Correlation and Common Structures in Volatility*

672 S. HE, *Cooperation, Coordination and Competition: Theory and Experiment*

673 C.G.F. VAN DER KWAAK, *The Macroeconomics of Banking*

674 D.H.J. CHEN, *Essays on Collective Funded Pension Schemes*

675 F.J.T. SNIEKERS, *On the Functioning of Markets with Frictions*

676 F. GOMEZ MARTINEZ, *Essays in Experimental Industrial Organization: How Information and Communication affect Market Outcomes*

677 J.A. ATTEY, *Causes and Macroeconomic Consequences of Time Variations in Wage Indexation*

678 T. BOOT, *Macroeconomic Forecasting under Regime Switching, Structural Breaks and High-dimensional Data*

679 I. TIKOUDIS, *Urban Second-best Road Pricing: Spatial General Equilibrium Perspectives*

680 F.A. FELS□, *Empirical Studies of Consumer and Government Purchase Decisions*

681 Y. GAO, *Stability and Adaptivity: Preferences over Time and under Risk*

682 M.J. ZAMOJSKI, *Panta Rhei, Measurement and Discovery of Change in Financial Markets*

683 P.R. DENDERSKI, *Essays on Information and Heterogeneity in Macroeconomics*

684 U. TURMUNKH, *Ambiguity in Social Dilemmas*