universität wien

DISSERTATION

Titel der Dissertation

"Non-Expected Utility vs. Expected Utility Theory in Consumption/Savings Decisions over the Life Cycle"

Verfasser

Diplom-Volkswirt Torsten Jürgen Lisson

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, August 2011

Studienkennzahl lt. Studienblatt:	A 094 140
Dissertationsgebiet lt. Studienblatt:	Volkswirtschaftslehre (Economics)
Betreuer:	Univ.-Prof. Dipl.-Ing. Dr. Gerhard Sorger

Acknowledgments

First of all I want to thank my supervisors Gerhard Sorger and Robert Kunst for their guidance, patience, critique and helpful suggestions while the development and especially in the final stage of this dissertation.

Furthermore, I want to thank my fellow students at the university of Vienna for interesting and helpful discussions on my research. Special thanks to Emanuel Gasteiger for being a good friend and colleague throughout our studies, and hopefully also in the time to come.

I am also grateful to David Laibson, Andrea Repetto and especially Jeremy Tobacman for providing me their computer code of their paper "Estimating Discount Functions with Consumption Choices over the Lifecycle", which was a very helpful guideline for the computations in the dissertation.

I received helpful comments from Larry Blume, Ana-Begona Ania-Martinez and the participants of the Graduate Seminars at the University of Vienna.

Last but not least, I want to thank my family for their great support.

Contents

1 Introduction

This thesis faces the question whether Non-Expected Utility models are better in explaining individual behavior in dynamic settings than Expected Utility Theory. In particular, we concentrate on two Non-Expected Utility models, namely Rank-Dependent Utility Theory proposed by Quiggin (1982) and Cumulative Prospect Theory put forward by Kahneman and Tversky (1992). For the evaluation of the performance of the theories we use a structural buffer stock consumption model for households, based on Laibson, Repetto and Tobacman (2007). Our theoretical framework nests all three theories, i.e. Expected Utility Theory, Rank-Dependent Utility Theory and Cumulative Prospect Theory. The distinction of these theories is pinned down by two parameters, which we estimate in a Method of Simulated Moments approach.

To the best of our knowledge, this is the first comparison of Expected Utility Theory and Rank-Dependent Utility and/or Cumulative Prospect Theory in a dynamic environment. In general, the development of Non-Expected Utility theories was motivated by experimental evidence indicating that people violate principles of Expected Utility Theory. These experiments, like from e.g. Allais (1953), Ellsberg (1961) and Kahneman and Tversky (1979, 1992), considered one time decisions, i.e. not various decisions made over a longer time period. In contrast to those we simulate consumption and savings decisions of households over the whole life cycle and compare these with real life data, which we take from the SOEP[1] survey and the SAVE[2] study. Our main objective is to find the theory that minimizes the distance between the simulated and the empirical profiles. The simulation procedure we apply exhibits a stochastic income process and the risk of survival. These elements define the risky components in the model, which are evaluated by the households over their life cycle, in order to make their decisions. The way in which risk is evaluated represents the main difference between Expected Utility Theory and Rank-Dependent Utility or Cumulative Prospect Theory. However, using Non-Expected Utility models in a dynamic setting is prone to some critique, mainly outlined in Machina (1989). The critique deals with some undesirable

[1]Socio-Economic Panel (SOEP), data for years 1984-2008, SOEP, 2009.

[2]SAVE (Sparen und Altersvorsorge in Deutschland), data for the years 2000 to 2008, SAVE 2009.

implications on economic behavior. This is, at least to our understanding, why Rank-Dependent Utility and Cumulative Prospect Theory are not used in dynamic models in the literature so far. In order to overcome these critique points we find that for the derivation of the dynamic decisions in our model one should use Folding Back[3] and especially should **not** apply the Reduction of Compound Lotteries Axiom. One consequence of this solution procedure is that it requires an evaluation of a large decision tree to obtain a simulation of the consumption and savings decisions of the households. This numerical simulation process is based on the work of Carroll (1992, 1997), Deaton (1991), Zeldes (1989), Gourinchas and Parker (2002), Hubbard, Skinner and Zeldes (1994,1995) and Laibson, Repetto and Tobacman (2007).

The Method of Simulated Moments approach, first introduced by McFadden (1989), Pakes and Pollard (1989), Duffie and Singleton (1993), which we apply consists of two stages. The first stage, sometimes called the calibration section, pins down parameters e.g. for the stochastic income process, the variation of the effective household size, interest rates and so forth. In the second stage the parameters of interest, which distinguish the three risk evaluation theories, are estimated. This is done by minimizing the distance between the simulated moments, which we obtain from the solution of our model, and the empirical moments, which we estimate from the available data. We can not show that the estimated parameters are asymptotically normally distributed. Due to this aspect we perform a bootstrap method to check the significance of our estimates. The final result of the estimation procedure is that Rank-Dependent Utility Theory is the risk evaluation theory which fits the data best.

The thesis is organized as follows: First, the model and the extensions to and deviations from the approach of Laibson, Repetto and Tobacman (2007) are described. The subsequent section displays some empirical evidence being responsible for the development of Non-Expected Utility Theories and introduces the corresponding value and probability weighting functions for Rank-Dependent Utility and Cumulative Prospect Theory. Section 4 gives a simplified example of the dynamic setting occurring in the model and considers two ways how such a dynamic problem can be solved. Section 5 discusses the implications of the proposed procedure of Section 4 in combination with Rank-

[3]Synonymous expressions in the literature are Backwards Induction or Rolling Back.

Dependent Utility and Cumulative Prospect Theory on the critique given in the literature. The following section concludes the theoretical part. Section 7 gives a rough description of the SOEP and SAVE datasets we use. Section 8 gives a brief overview of the estimation procedure. Sections 9 and 10 deal with the calibration of the first stage parameters needed for the simulated moments. Then, the empirical moments are defined and estimated from data. Section 12 gives detailed information about the simulation process. Section 13 defines the Method of Simulated Moments estimator. Then, the results from the estimation are presented and the significance of these is discussed. The next two sections interpret and check the robustness of the results. Section 17 discusses some extensions and future research questions. The last section concludes.

2 Framework

The model, which we use, is based on the simulation literature pioneered by Carroll (1992, 1997), Deaton (1991), Hubbard, Skinner and Zeldes (1994,1995), Gourinchas and Parker (2002), and Laibson, Repetto and Tobacman (2003). In most terms, the model is very similar to Laibson, Repetto and Tobacman (2007). We deviate in the way that we use a more general form of the instantaneous utility or value function, respectively. This enables us to compare various risk evaluation systems in the context of consumption and savings decisions over the life cycle within the estimation process. These risk evaluation systems are in particular Expected Utility Theory, Rank-Dependent Utility and Cumulative Prospect Theory. Furthermore, the model uses exponential discounting instead of hyperbolic discounting[4] as in Laibson, Repetto and Tobacman (2007).

The general model consists of 7 parts: demographics, income, liquid assets, illiquid assets, dynamic and static budget constraint, utility/value function, and equilibrium.

[4] Actually, Laibson, Repetto and Tobacman (2007) use a quasi-hyperbolic discount function in their model.

2.1 Demographics

In a household, which is considered as one decision unit, always lives a head and a spouse. The number of children and dependent adults in the household varies over time. Economic life of a household starts with the age of 20 and ends at 90.[5] The household is retired if the household is older than T, which will be estimated from data later on. There is an exogenous probability of surviving period t, labeled s_t. The effective household size in period t, n_t, is equal to the number of adults and 0.4 times the number of children in the household. The factor 0.4 stems from Blundell, Browning, Meghir (1994) to achieve "equivalent" household size. The idea is to control for the different needs of adults and children.[6]

For the simulation and calibration in the upcoming sections households will be categorized by their education. This is due to the aspect that there are differences in retirement age and especially in income.

2.2 Income

The income Y_t contains all after-tax transfers and wages, i.e. labor income, inheritances, private defined-benefit pensions and all government payments of period t. It is assumed that labor is supplied inelastically, according to this Y_t is exogenous. The income equation is separated over time: i) the income while the household is in the workforce, i.e. younger than T, and ii) the income when the household is retired. The working age income is modeled as

$$ln(Y_t) = y_t = HH_t + f^W(t) + u_t + v_t^W,$$

where $f^W(t)$ is a polynomial in age, v_t^W is an iid shock and u_t is an AR(1) process, which later on will be substituted by a Markov process to reduce computation time for the solution of the model.[7] The AR(1) process is included to cover shocks, which are influencing income over more periods, whereas v_t^W

[5]The reference age for the household is the age of the head of the household.

[6]Blundell, Browning, Meghir (1994) note that one could also use more sophisticated equivalence scales that allow for economies of scale and age differences of children.

[7]The particular form of the income equations for working life and retirement will be estimated later on in the calibration section. There will be also a detailed discussion about the substitution of the AR(1) process by a Markov Process.

are transitory shocks affecting only one period. Additionally we implement variables concerning household composition, summarized in HH_t.

The income process during retirement is modeled in a simpler way and without persistent shocks, i.e.

$$ln(Y_t) = y_t = c + v_t^R,$$

where c is a constant and v_t^R represents transitory shocks in retirement. As mentioned before these income processes will be estimated later on.

2.3 Liquid Assets

The liquid assets at the beginning of period t are denoted by X_t. Hence, the disposable money of a household in period t is $Y_t + X_t$, i.e. income plus liquid assets. It is allowed that X_t can be negative, i.e. the households have the possibility to borrow on credit cards or overdraft credits. A credit line is assumed to restrict the credits and is defined by $X_t \geq -\lambda \cdot \bar{Y}_t$, where $\bar{Y}_t$ is the age-specific income at age t and λ is a fraction calibrated later on.

2.4 Illiquid Assets

The illiquid assets at the beginning of period t are denoted by Z_t. This variable is supposed to be positive throughout the life cycle, i.e. $Z_t \geq 0, \forall t$. There are two types of returns: i.) capital gains from interest and ii.) a consumption flow. The consumption flow in period t is defined as $\gamma \cdot Z_t = 0.05 \cdot Z_t$, and can be seen e.g. as the consumption of living in a house you own.[8] Transaction costs for selling the illiquid assets are assumed to be so high, that Z is only increasing over time, until it is bequeathed to the next generation. Laibson, Repetto and Tobacman (2007) argue, that the construction of Z does not match any illiquid asset in particular, but has some properties of home equity.[9] They justify these

[8]The assumption of $\gamma = 0.05$ is taken from Laibson, Repetto and Tobacman (2007). We check the influence of other values for γ in the robustness section.

[9]The intuition Laibson, Repetto and Tobacman (2007) give is: *"Consider a consumer who owns a house of fixed real value H and derives annual consumption flows from the house of $\gamma \cdot H$. Suppose the consumer has a mortgage of size M and home equity of H-M. The real cost of the mortgage is $\eta \cdot M$, where $\eta = i \cdot (1 - \tau_T) - \pi$ is the nominal mortgage interest rate adjusted for inflation and the tax deductibility of interest payments. If we assume $\eta \approx \gamma$, the net benefit to the homeowner is $\gamma \cdot H - \eta \cdot M \approx \gamma \cdot (H - M) = \gamma \cdot Z$."*

assumptions by four observations, whereas in the environment of our setting only two of these apply, as the other two are linked to quasi-hyperbolic discounting. Their first argument is, that selling houses, cars or pension plans entails at least small transaction costs and delays. The second, and for the approach here the most important, is that one extremely reduces the choice set of each household, i.e. computation time of the Method of Simulated Moments is enormously decreased by this construction of the illiquid asset.

2.5 Dynamic and Static Budget Constraints

The net investments in liquid and illiquid assets in period t are I_t^X and I_t^Z, respectively. The dynamic budget constraints are defined by

$$X_{t+1} = R^X \cdot (X_t + I_t^X)$$

$$Z_{t+1} = R^Z \cdot (Z_t + I_t^Z),$$

where R^X and R^Z are the interest rates for the liquid and illiquid asset, respectively. As mentioned in the description of the liquid asset, X can also be negative - this case was interpreted as borrowing on credit cards or overdraft credits. Therefore, the interest rate for the liquid asset has to be defined for positive and negative holdings, i.e.

$$R^X = \begin{cases} R^X_{CC} & \text{if } X_t + I_t^X < 0 \\ R^X & \text{if } X_t + I_t^X \geq 0. \end{cases}$$

The static budget constraint for each period t is given by

$$C_t = Y_t - I_t^X - I_t^Z,$$

where C_t is the consumption in period t. Consumption is calculated as the residual of the right hand side of the above equation throughout the whole simulation procedure.

2.6 Utility/Value Function

At this point we deviate from the approach of Laibson, Repetto and Tobacman (2007). We extend the instantaneous payoff function by a reference point r_{ref}, which will be later on estimated in the Method of Simulated Moments procedure. This utility function can be used for Cumulative Prospect Theory, as this theory contains such a reference point to describe a particular status quo, like e.g. past consumption. Hence, a representative household experiences at period t the instantaneous value/utility

$$u(C_t, Z_t, n_t) = \begin{cases} n_t \cdot \frac{\left(\frac{C_t+\gamma Z_t - r_{ref}}{n_t}\right)^{1-\rho} - 1}{1-\rho} & \text{for} \quad C_t + \gamma Z_t - r_{ref} \geq 0 \\ \\ -\lambda_L \cdot n_t \frac{\left(-\left(\frac{C_t+\gamma Z_t - r_{ref}}{n_t}\right)\right)^{1-\tau}}{1-\tau} & \text{for} \quad C_t + \gamma Z_t - r_{ref} < 0. \end{cases}$$

The function is twofold, as with a reference point the evaluation of the consumption (inclusive consumption flows from the illiquid asset) can be perceived as a loss. The construction of this utility function is based on the idea of Kahneman and Tversky (1992), which we will discuss in more detail in the next section.

For now, let ρ and τ be parameters determining the curvature of the functions, normally $\rho, \tau \in [0; 1]$, and λ_L is a weighting parameter if the consumption in relation to the reference point is perceived as negative.[10] The reference point was interpreted by Kahneman and Tversky (1979) as the current wealth level, when an individual was faced with an uncertain or risky problem. Here it will be interpreted as a reference consumption level - indicating whether current consumption is considered as a loss or a gain. This concept is called Habit Formation or Habit Persistence in the literature, see e.g. Duesenberry (1949), Pollak (1970) or Constantinides (1990). In these papers the reference point is defined by the consumption of the previous period or periods. The idea is that people/families tend to maintain their standard of living and perceive consumption below past consumption as a loss. In contrast to this kind of determination of the reference point, we will not use any past levels of consumption. The main reason why we proceed this way is that the concept of

[10]The weight λ_L refers to a concept labeled Loss Aversion by Kahneman and Tversky (1979). We will describe this aspect in more detail in the next section.

Habit Formation would enormously increase computation time in the simulation process. We will describe later on how we set up the reference points over the life cycle.

Finally, in order to determine the utility of a household over the whole time horizon, let us define the continuation payoff. However, the continuation payoff can only be calculated after the whole dynamic system, which is described in the next subsection, is solved. That is, after all investment decisions for X and Z in each period over the life cycle are made. Together with the expectations of the income these decisions determine consumption. Given that the continuation value in t, denoted V_t, is defined by

$$V_t = \sum_{i=1}^{N-t} \delta^i \left(\prod_{j=1}^{i-1} w_{h,l}(s_{t+j}) \right) [w_{h,l}(s_{t+i}) \cdot u(C_{t+i}, Z_{t+i}, n_{t+i}) \\ + (1 - w_{h,l}(s_{t+i})) \cdot b(X_{t+i}, Z_{t+i})],$$

where δ is the discount factor and s_t the survival rate. The function $w_{h,l}$ (h=high and l=low) is determined by a probability weighting function π and whether u plus the corresponding continuation value is bigger or smaller than b, i.e.

$$\begin{aligned} w_h(s_t) &= \pi(s_t) \quad \text{if} \quad u(C_t, Z_t, n_t) + V_t \geq b(X_t, Z_t) \\ w_l(s_t) &= 1 - \pi(1 - s_t) \quad \text{if} \quad u(C_t, Z_t, n_t) + V_t < b(X_t, Z_t). \end{aligned}$$

The reason why the function w is constructed that way and how the information about $u(C_t, Z_t, n_t) + V_t \stackrel{\leq}{>} b(X_t, Z_t)$ is received will be clarified later on. At the moment, it is just important to mention that this kind of representation enables the model to capture Expected Utility Theory[11], Rank-Dependent Utility and Cumulative Prospect Theory, at least the version used in this thesis. Recall, for the latter two theories the model has first to be solved, before the continuation payoff can be calculated. Actually, the overall value, i.e. over the whole life cycle, is not used for further examinations in this thesis, but would be necessary if this model is used for any kind of policy evaluation.

The function b represents a bequest motive of the household and is defined

[11] Note, if π is the identity function, then $w_h = w_l = s_t$. This leads to the continuation payoff for Expected Utility Theory.

by

$$b(X, Z) = (R - 1) \cdot max(0, X + 2/3 \cdot Z) \cdot b_B.$$

As in Laibson, Repetto and Tobacman (2007), liquidated bequeathed wealth is consumed by heirs as an annuity, therefore the factor of interest $(R - 1)$. The illiquid asset Z is multiplied by 2/3, because liquidating such assets is often associated with large transaction costs. The parameter b_B is a weighting unit to reach an appropriate equivalence to the instantaneous utility/value function. Laibson, Repetto and Tobacman (2007) set $b_B = \frac{u_1(\hat{y},0,\hat{n})}{1-\delta}$, where u_1 is the partial derivative of the instantaneous utility function u, $\hat{y}$ is the average labor income over the life cycle and $\hat{n}$ is the average effective household size.

2.7 Equilibrium

For the solution of the model, it is assumed that each period t is associated with a "Self t", the decision maker in period t. These Selves, for $t = (20, 21, 22, ...90)$,[12] are the players in an intra-personal game. This approach follows Strotz (1956), Pollak (1968), Phelps and Pollak (1968), and is also used by Laibson et al. (2007). At first, each Self t, i.e. player, realizes the state it is in. A state is defined by a vector of state variables $\Lambda_t = \{t, \bar{Y}_t + X_t, Z_t, m_t, v_t\}$, where t is the current period, $\bar{Y}_t$ the age-specific household income, X_t and Z_t the liquid and illiquid asset holdings at the beginning of t, u_t is the Markov State of the income in period t and v_t is the transitory shock in t.[13] After the realization of its state, Self t chooses a strategy by determining I_t^X and I_t^Z, taking the strategies of all the other Selves as given. Accordingly, the equilibrium is defined as: all strategies are optimal given the strategies of the other players. The optimal strategies for each period and its corresponding decisions are found via numerical Backward Induction - also called Folding Back or Rolling Back in the literature.[14] The objective function of Self t is given by

$$u(C_t, Z_t, n_t) + \delta E_t V_{t,t+1}(\Lambda_{t+1}),$$

[12] A finite horizon is assumed, as this implies a unique equilibrium - for details see Laibson (1997).

[13] Recall, the Markov process represents an AR(1) process - this will be discussed in more detail later.

[14] Note, the model described in this section can not be solved analytically, even under Expected Utility Theory.

where $V_{t,t+1}(\Lambda_{t+1})$ is the $t+1$ continuation payoff of Self t. The operator E_t represents the value of expectations according to the risk evaluation system used in the model, i.e. Expected Utility Theory, Rank-Dependent Utility or Cumulative Prospect Theory. This equation is maximized by choosing I_t^X and I_t^Z, which determine C_t.[15] The sequence of continuation payoffs is defined as

$$V_{t,t+1}(\Lambda_{t+1}) = w_{h,l}(s_{t+1})\left[u(C_{t+1}, Z_{t+1}, n_{t+1}) + \delta E_t V_{t+1,t+2}(\Lambda_{t+2})\right] + (1 - w_{h,l}(s_{t+1}))E_t b(\Lambda_{t+2}).$$

Here, the weighting function $w_{h,l}$ is as described above - due to the usage of Rank-Dependent Utility and Cumulative Prospect Theory. Recall, if the value of the term in the squared brackets is bigger than the value from the bequest function b, then w_h is used and vice versa. As a final remark, note, that the solution of the system starts at $t = 90$, which is the maximum age a household can reach, and runs backwards.

3 Non-Expected Utility in Decisions under Risk

This section deals with empirical evidence showing violations of Expected Utility Theory and introduces two Non-Expected Utility Theories, which are able to capture the empirical findings. These two alternative theories are Rank-Dependent Utility proposed by Quiggin (1982) and Cumulative Prospect Theory by Kahneman and Tversky (1992). The upcoming experiments and choice theories only consider decisions under risk, i.e. the probabilities of the possible events are known, in contrast to decisions under uncertainty, when the probabilities are not known.[16] Recall, in the model description in the previous section the utility/value function and the continuation payoffs were defined in such a way to permit the usage of Expected Utility Theory, Rank-Dependent Utility or Cumulative Prospect Theory.

[15]Consumption was defined by the static budget constraint: $C_t = Y_t - I_t^X - I_t^Z$, where I_t^X and I_t^Z are chosen by Self t and Y_t is given by the age-specific income $\bar{Y}$, the Markov state u_t and transitory shock v_t^W, which occur in period t.

[16]An example for a decision under risk is if the outcome is determined by rolling a die, whereas betting on horse races is an example for decision under uncertainty.

3.1 Why Non-Expected Utility Models?

Apart from the standard economic theory for choice under risk, namely Expected Utility Theory, there are many others, generally called Non-Expected Utility Models.[17] We will concentrate on two of those, namely Rank-Dependent Utility, also known as Anticipated Utility Theory, and Cumulative Prospect Theory. The reason why there exist so many alternative theories for choice under risk[18], is that there are empirical observations showing that decision makers systematically violate basic principles of Expected Utility Theory or indicating in some parts a different evaluation method.[19] In order to give a general impression, the upcoming part will exhibit some of the experiments executed by Kahneman and Tversky (1979).

First of all, let us define lotteries or prospects (these two terms are used synonymously) before we turn to the empirical evidence. A lottery $(x_1; p_1|...|x_n; p_n)$ is a gamble that yields with probability p_1 outcome x_1 and so forth. The probabilities of all the outcomes sum up to 1, i.e. $\sum_{i=1}^{n} p_i = 1$. The evaluation of such a lottery according to Expected Utility Theory is defined by

$$U(x_1; p_1|...|x_n; p_n) = \sum_{i=1}^{n} p_i \cdot u(x_i),$$

where $u(x_i)$ is a function, standardly used in economics, associating an utility value to an outcome x_i and p_i is the weight of this utility value, which it contributes to the value of the whole lottery.

One of the most famous experiments, which violates Expected Utility Theory, stems from Allais (1953) and is known as the Allais Paradox. A variation of Allais's experiment was executed by Kahneman and Tversky (1979) and they labeled this phenomenon the Certainty Effect. This effect refers to problem sets, where at least one lottery is involved, which provides an outcome with probability 1, i.e. with certainty. Consider the following example: People were asked to choose lottery A or B in Problem 1, and lottery C or D in Problem 2.

[17] For a list of some of these alternative models see Table C.1 in the Appendix and for a summary see Starmer (2000).

[18] There are also many alternative theories for choice under uncertainty, but these are not under consideration in this thesis.

[19] See e.g. Allais (1953), Ellsberg (1961), Kahneman and Tversky (1979, 1992) - for an overview see Yaqub, Haz and Hussain (2009).

Problem 1:[20] N=72

A: (2500;0.33 | 2400;0.66 | 0;0.01) [18]

B: (2400;1) [82*]

The number of participants was N=72, and the percentages of the chosen option are given in the squared brackets, i.e. in Problem 1 82 percent of the respondents preferred lottery B, namely receiving 2400 with certainty. The asterisk indicates that the results are significant at the 0.01 level.

Problem 2: N=72

C: (2500;0.33 | 0;0.67) [83*]

D: (2400;0.34 | 0;0.66) [17]

The choices from Problem 1 and 2, namely preferring B over A and C over D, are not consistent with Expected Utility Theory, at least for the majority of the respondents. This can be shown by using the evaluation principle of Expected Utility Theory, as defined above - with $u(0) = 0$. The preferences of Problem 1 yield the following inequation

$$u(2400) > 0.33 \cdot u(2500) + 0.66 \cdot u(2400)$$

or rewritten, if one summarizes equal outcomes

$$0.34 \cdot u(2400) > 0.33 \cdot u(2500).$$

But, looking at the preferences given by Problem 2, one gets

$$0.34 \cdot u(2400) < 0.33 \cdot u(2500).$$

[20]Kahneman and Tversky (1979) p.264: *"The outcomes refer to Israeli currency. To appreciate the significance of the amounts involved, note that the median net monthly income for a family is about 3000 Israeli pounds. The respondents were asked to imagine that they were actually faced with the choice described in the problem, and to indicate the decision they would have made in such a case. The responses were anonymous, and the instructions specified that there was no 'correct' answer to such problems, and that the aim of the study was to find out how people choose among risky prospects."*

These preferences obviously contradict each other, correspondingly the respondents did not obey Expected Utility Theory. The results of the above example exhibits a pattern, which has become known as the Common Consequence Effect. This effect disagrees with Expected Utility Theory in a similar way, only from another perspective. It is possible to show this by reformulating the lotteries A,B,C and D, without changing the actual problem.

Table 1: Common Consequence Effect

probability	0.01	0.33	0.66
A	0	2500	2400
B	2400	2400	2400
C	0	2500	0
D	2400	2400	0

In Table 1 one can see that in column 2 and 3 the lottery pairs A/B and C/D face the same probability-outcome composition, namely 0 and 2400 for the probability 0.01, and 2500 and 2400 for the probability 0.33. Thus, the only difference between Problem 1 and 2 is in column 3, where A and B share a same consequence (2400 for the probability 0.66), as well as C and D (0 with a probability of 0.66). According to Expected Utility Theory the common consequence in Problem 1 should not affect the preference between the lotteries A and B. The same should hold for the lotteries C and D. This implies that an individual confronted with Problem 1 and 2 should ignore the third column and concentrate on the first two columns to make her/his decision. As stated, the problems are identical, if the third column is omitted. Hence, according to Expected Utility Theory a decision maker should either prefer A over B and C over D, or B over A and D over C, but the above experiment indicates otherwise.

Another example, which violates Expected Utility Theory and also demonstrates the Certainty Effect, is the following problem pair:

Problem 3: N=95

A: (4000;0.80 | 0;0.20) [20]

B: (3000;1) [80*]

Problem 4: N=95

C: (4000;0.20 | 0;0.80) [65*]

D: (3000;0.25 | 0;0.75) [35]

Here, we apply the evaluation principle of Expected Utility Theory once again to see the violation. For Problem 3 one gets

$$0.8 \cdot u(4000) < u(3000) \Leftrightarrow \frac{u(3000)}{u(4000)} > \frac{4}{5}.$$

But, Problem 4 implies the opposite, namely

$$0.2 \cdot u(4000) > 0.25 \cdot u(3000) \Leftrightarrow \frac{u(3000)}{u(4000)} < \frac{4}{5}.$$

An important feature of this problem pair is that lottery C=(4000;0.8 | 0;0.2) can be written in terms of lottery A. This yields the lottery (A;0.25 | 0;0.75), i.e. lottery C is now a chance of 0.25 winning lottery A and 0.75 getting nothing. The same is possible for D, which leads to D=(B;0.25 | 0;0.75). The point to see is that the reformulations of C and D share the same ratio, i.e. a chance of 0.25 winning A or B, respectively. Hence, it should be again a choice between A and B as in Problem 3. This logic is apparently not obeyed by the test subjects. This effect is called the Common Ratio Effect in the literature and is a direct violation of the Independence Axiom, which is central in Expected Utility Theory. This axiom states that if B is preferred to A, then any probability mixture (B;p | 0;1-p) must be preferred to the corresponding mixture (A;p | 0;1-p)[21]. These kind of experimental results were also found by others, like Tversky (1975), Hagen (1979), MacCrimmon and Larsson (1979), and Chew and Waller (1986).

Another aspect, which Kahneman and Tversky (1979) found, is labeled the Reflection Effect. This effect deals with the question of how do people act if the outcomes of lotteries are negative. According to this finding, there is no direct violation of Expected Utility Theory, but it exhibits a new aspect which Kahneman and Tversky incorporated in their new theory. For the examination

[21] The Independence Axiom will be discussed in more detail in Section 5.2.

of this question they simply reversed the outcomes of the Problems 3 and 4. Their results are displayed in Table 2.

Table 2: Positive and negative prospects

	positive outcomes			negative outcomes		
Problem 3	(4000;0.8)[a]	$\prec$	(3000;1)	(-4000;0.8)	$\succ$	(-3000;1)
N=95	[20]		[80*]	[92*]		[8]
Problem 4	(4000;0.2)	$\succ$	(3000;0.25)	(-4000;0.2)	$\prec$	(-3000;0.25)
N=95	[65*]		[35]	[42]		[58]

[a] The lotteries in the table are reduced - the outcomes equal to 0 and the corresponding probabilities are omitted, i.e. lottery A $(4000; 0.8|0; 0.2)$ is now in short (4000;0.8).

The percentage values (in squared brackets) indicate that the pattern of choice from positive to negative outcomes has changed - is reflected. These findings imply a risk averse attitude in the positive domain, i.e. faced with positive outcomes, and a risk seeking attitude in the negative domain. Hence, the utility/value function of a corresponding respondent ought to be convex for negative values and concave for positive ones. This kind of behavior was also detected by e.g. Markowitz (1959), Williams (1966), and Fishburn and Kochenberger (1979).

As stated above, there are many other experiments, which exhibit violations of Expected Utility Theory. However, this section is intended to give just a small impression about how such experiments were conducted to identify deviations from Expected Utility Theory. In the next section, we will show some further empirical examples, which are responsible for the particular development of Prospect Theory, Rank-Dependent Utility and Cumulative Prospect Theory.

3.2 (Cumulative) Prospect Theory and Rank-Dependent Utility

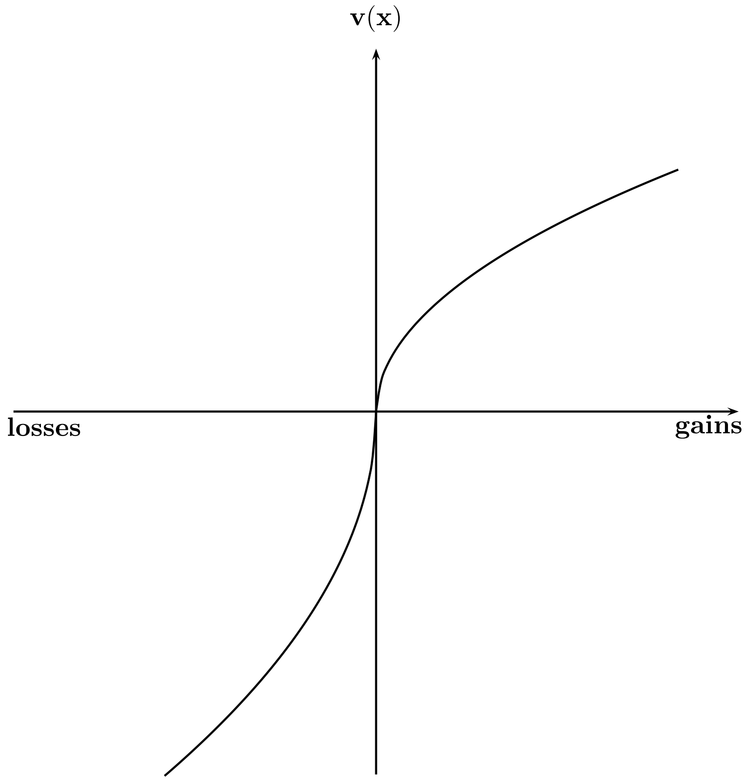
Based on their experiments Kahneman and Tversky (1979) derived an alternative to Expected Utility Theory, called Prospect Theory. This new theory is able to capture most of their empirical findings. The innovations are a new kind of value function and a weighting function for the probabilities.

The evaluation of a lottery $(x_1; p_1|...|x_n; p_n)$ according to Prospect Theory, in order to calculate the value V, is defined as

$$V = \pi(p_1) \cdot v(x_1) + ... + \pi(p_n) \cdot v(x_n),$$

where v is a value function and π is a probability weighting function. The value function has properties like the one displayed in Figure 1.

Figure 1: Hypothetical Value Function



The value function is defined for the positive as well as the negative domain, in order to cope with positive and negative outcomes in a lottery. These are interpreted as gains or losses relative to a predetermined reference point, like e.g. previous (i.e. before a decision maker is faced with a decision problem) asset holdings, wealth or consumption. For the introduction of a reference point Kahneman and Tversky (1979) give an intuitive justification[22] as well as

[22]They reason that *"an object at a given temperature may be experienced as hot or cold to the touch depending on the temperature to which one has adapted. The same principle applies to non-sensory attributes such as health, prestige, and wealth. The same level of wealth, for example, may imply abject poverty for one person and great riches for another - depending on their current assets."* - Kahneman and Tversky (1979), p. 277.

empirical evidence. The following two problems were given to two different groups.

Problem 5: In addition to whatever you own, you have been given 1000. You are now asked to choose between (N=70):

A: (1000;0.50 | 0;0.50) [16]

B: (500;1) [84*]

Problem 6: In addition to whatever you own, you have been given 2000. You are now asked to choose between (N=68):

C: (-1000;0.50 | 0;0.50) [69*]

D: (-500;1) [31]

In the above example one can observe again the Reflection Effect. Note, if regarded in final states, i.e. implementing the "pre-payments" into the lotteries, then the problems are identical

$$A = (2000; 0.50|1000; 0.50) = C \text{ and } B = (1500; 1) = D.$$

Due to this experiment Kahneman and Tversky (1979) conclude that the carriers of value or utility are changes in wealth, rather than final asset holdings that include current wealth. There are also some other empirical examples in the literature, indicating that individuals tend to evaluate from a reference point, e.g. Kahneman, Knetsch and Thaler (1990).

Furthermore, the function in Figure 1 is concave in the positive and convex in the negative domain. This feature stems from experiments like the one shown in Table 2 in the previous section - indicating that individuals are risk averse if faced with positive outcomes and risk seeking if faced with negative ones.

Another property of the function in Figure 1, is that it is steeper in the negative than in the positive domain. This aspect, called Loss Aversion, is often phrased in the literature as "losses loom larger than gains". The observation stems from the fact that most individuals dislike symmetric bets, i.e. lotteries of the form $(-x; 0.5|x; 0.5)$.[23] Kahneman and Tversky (1979) also claim that the aversiveness of symmetric fair bets increases with the size of stake. Suppose $x > y \geq 0$, then according to their proposition $(-y; 0.5|y; 0.5) \succ (-x; 0.5|x; 0.5)$. This implies - using Prospect Theory -

$$\pi(0.5) \cdot v(y) + \pi(0.5) \cdot v(-y) > \pi(0.5) \cdot v(x) + \pi(0.5) \cdot v(-x)$$

or

$$v(-y) - v(-x) > v(x) - v(y).$$

If y is set equal to 0, then one gets $-v(-x) > v(x)$. Now, let y converge to x , so that $v(x) - v(y)$ and $v(-y) - v(-x)$ approximately represent the slope in the corresponding intervals[24], i.e. $v'(-x) > v'(x)$ - assuming that the derivative v' of v exists.

According to the finding that the carriers of value or utility are gains and losses, and the argument of Loss Aversion, Kahneman and Tversky (1992) defined the value function as

$$v(x) = \begin{cases} x^{\rho}, & x \geq 0 \\ -\lambda_L(-x)^{\tau}, & x < 0, \end{cases}$$

where λ_L is a coefficient of Loss Aversion. The parameters ρ and τ can be interpreted as coefficients of risk aversion and risk seeking, respectively, and represent, if $\rho, \tau \in [0, 1]$, a psychological principle known as Diminishing Sensitivity. This principle declares that people are more sensitive to changes close to their reference point than to changes more far away. For example, the difference between €100 and €200 is perceived as greater than the difference between €1 100 and €1 200, although the bigger gains are in general preferred.

One final remark should be made to conclude the introduction of the value

[23] Dislike means that e.g. if people are asked to choose between $(-x; 0.5|x; 0.5)$ and not participating in any lottery, the majority decides not to play, although both options have the same expected value. In contrast to the aspects discussed before Expected Utility Theory does not violate this one.

[24] The derivation is as follows: $\frac{v(-x)-v(-y)}{-x+y} > \frac{v(x)-v(y)}{x-y}$ the denominator is on both sides of the inequation equal in magnitude, i.e. one can conclude that $v'(-x) > v'(x)$ if y converges to x.

function of Prospect Theory. The convex and concave curvature of the value function is not necessarily required to impose risk seeking and risk averse behavior, as these attitudes can also be captured by the probability weighting function. But, Kahneman and Tversky (1979) justify their approach by the fact that other independent studies from various fields, like Barnes and Reinmuth (1976), Grayson (1960), Grether and Plott (1979), Halter and Dean (1971), and Swalm (1966), came up with similar results, i.e. most utility functions were concave for gains and convex for losses, and also exhibited steeper slopes for the negative domain.

Before we turn to the probability weighting function, mentioned as the second innovation of Prospect Theory above, we will introduce two new theories for choice under risk, namely Rank-Dependent Utility and Cumulative Prospect Theory. There are two reasons for this: First, in their 1979 paper Kahneman and Tversky solely sketched some properties of a hypothetical weighting function and did not define the function over the whole probability interval, especially near the endpoints 0 and 1. Second, the features Kahneman and Tversky (1979) outlined for a weighting function, mainly that $\pi(p) \neq p$ and $p + q = 1$ but $\pi(p) + \pi(q) < 1$, gave rise to the critique of violations of dominance.[25] Due to this critique a new way of handling the probability weighting function was brought forward. The first of these new theories was proposed by Quiggin (1982)[26] and is called Rank-Dependent Utility (or Anticipated Utility Theory as in the original paper).

According to this theory, the evaluation of a lottery $(x_1; p_1|...|x_n; p_n)$ is executed in the following way. First, all the outcomes of the lottery are ranked in an ascending (or descending)[27] manner, so that $x_1 < x_2 < ... < x_n$[28]. Then, the value V of the lottery is calculated by[29]

$$V = \sum_{i=1}^{n} h_i(p) \cdot u(x_i),$$

[25] Violations of stochastic dominance for simple lotteries - see Appendix A for an example.

[26] Quiggin (1982) for decisions under risk and Schmeidler (1989) for decisions under uncertainty.

[27] It does not matter in which order the outcomes are ranked, the principle stays the same, it is only reversed.

[28] If there are two identical outcomes, then the probabilities of these are combined to one outcome/probability pair.

[29] In the equation $p = (p_1, p_2, ..., p_n)$.

where u is a utility function as used in Expected Utility Theory. Here, the decision weights $h_i(p)$, satisfying $\sum_{i=1}^{n} h_i(p) = 1$, depend on all $p's$ and are defined as

$$h_i(p) = \pi\left(\sum_{j=i}^{n} p_j\right) - \pi\left(\sum_{j=i+1}^{n} p_j\right),$$

where π, with $\pi(1) = 1$ and $\pi(0) = 0$, is a probability weighting function as in Prospect Theory, which will be defined in more detail below. The above equation exhibits the difference between *"the chance of winning an outcome at least as good as x_i"* and *"the chance of winning an outcome strictly better than x_i"*. To illustrate the evaluation procedure consider the following game.[30] You roll a die once and observe the result $x = 1, ..., 6$. If x is even, you receive \$ x and if it is odd you pay \$ x. The outcomes are equiprobable, i.e. each is obtained with probability 1/6. Ranking the outcomes in an ascending fashion yields the lottery $(-5; 1/6|-3; 1/6|-1; 1/6|2; 1/6|4; 1/6|6; 1/6)$. The value V of this lottery is given by

$$\begin{aligned}
V &= [\pi(1) - \pi(5/6)] \cdot u(-5) \\
&\quad + [\pi(5/6) - \pi(4/6)] \cdot u(-3) \\
&\quad + [\pi(4/6) - \pi(3/6)] \cdot u(-1) \\
&\quad + [\pi(3/6) - \pi(2/6)] \cdot u(2) \\
&\quad + [\pi(2/6) - \pi(1/6)] \cdot u(4) \\
&\quad + [\pi(1/6) - \pi(0)] \cdot u(6).
\end{aligned}$$

For example, the last line of the above equation describes the weighted cumulative chance of receiving an outcome equal to or better than \$ 6. That is, a weighted probability of 1/6 receiving an outcome at least as good as \$ 6, minus the weighted chance of winning an outcome strictly better than \$ 6, which is 0 in the given lottery.

The second theory we want to introduce, is Cumulative Prospect Theory put forward by Kahneman and Tversky (1992), which is in principle an ex-

[30]This example is from Kahneman and Tversky (1992) - only the evaluation of the lottery is a little different, as they do not apply Rank-Dependent Utility.

tension of the original Prospect Theory by the cumulative functional used by Quiggin (1982). The main difference between Rank-Dependent Utility and Cumulative Prospect Theory is that the latter uses gains and losses as the carrier of values, instead of final terms. Hence, the value function is defined as in Prospect Theory with a predetermined reference point, which is responsible for outcomes being considered as positive or negative. The evaluation procedure is, at the beginning similar to Rank-Dependent Utility. That means all the outcomes are first ranked in an ascending (or descending) order. In the next step, the outcomes are divided in negative and positive ones, corresponding to the reference point, and evaluated separately. This leads to the fact that the calculation of the overall value V of a lottery $(x_1; p_1|...|x_n; p_n)$ is twofold[31], in the sense of

$$V = V^+ + V^-.$$

The outcomes $(x_1, ..., x_m)$ are perceived as losses and $(x_{m+1}, ..., x_n)$ as gains according to the reference point.

$$V^- = \sum_{i=1}^{m} h_i^-(p)v(x_i), \quad V^+ = \sum_{i=m+1}^{n} h_i^+(p)v(x_i)$$

Hence, V^+ summarizes the value of all positive outcomes and V^- the value of all negative outcomes. The cumulative weightings of the positive outcomes are calculated according to the intuition *"the chance of winning an outcome at least as good as x_i"* minus *"the chance of winning an outcome strictly better than x_i"*, whereas for the negative outcomes the prevailing intuition is *"the risk of receiving an outcome at least as bad as x_i"* minus *"the risk of receiving an outcome strictly worse than x_i"*. Furthermore, Kahneman and Tversky (1992) use different probability weighting functions for the positive and negative parts of the lottery, namely π^+ and π^-. This leads to the following definition of the cumulative weighting functions

$$h_i^+(p) = \pi^+\left(\sum_{j=i}^{n} p_j\right) - \pi^+\left(\sum_{j=i+1}^{n} p_j\right), \forall i > m$$

[31]Twofold only if the lottery contains positive and negative outcomes according to the reference point.

and

$$h_i^-(p) = \pi^-\left(\sum_{j=1}^{i} p_j\right) - \pi^-\left(\sum_{j=1}^{i-1} p_j\right), \forall i \leq m.$$

Although, this representation implies that the sum of all cumulative weights can be smaller or greater than 1,[32] it does not exhibit violations of stochastic dominance as Prospect Theory.

Kahneman and Tversky (1992) set up an experiment to receive detailed information about the value and the weighting functions. The general experiment setting was that the participants were confronted with a two-outcome lottery (e.g. (x;0.75 | 0;0.25) - a 75 percent chance of winning \$ x and a 25 percent chance of zero) and seven sure outcomes, logarithmically spaced between the two outcomes of the lottery, i.e. \$ x and 0. The subjects were asked to reveal their preferences between the lottery and each of the sure outcomes. In order to obtain more detailed information about the certainty equivalent[33] a second round was executed. In this round, the same question was posted, with the difference that the sure outcomes were now linearly spaced between a value 25 percent higher than the lowest amount accepted in the first round and a value 25 percent lower than the highest amount rejected. Then, the certainty equivalent was obtained by the middle of the lowest value accepted and the highest value rejected of the second round. The interpretation of the results was based on the principle that people who revealed a certainty equivalent, which was close to the expected value of the lottery were risk neutral. Risk neutrality means that the value function and the probability weighting function are the identity functions. On the other hand, if the certainty equivalent was smaller (bigger) than the expected value an individual was assumed to be risk averse (risk seeking). According to this approach, the fourfold pattern of risk attitudes was obtained - see Table 3.

The fourfold pattern describes that people seem to be risk averse if it comes to losses with low probabilities or gains with high probabilities, and to risk seeking behavior if it comes to losses with high probabilities or gains with low probabilities. The lotteries and their certainty equivalents are taken from the

[32]The sum can only be unequal 1, if there are positive and negative outcomes in the lottery.

[33]The certainty equivalent is the sure payoff, when the subject is indifferent between the risky lottery and this sure payoff.

Table 3: The Fourfold Pattern of Risk Attitudes[a]

	Gain	Loss
Low probability	C(100;0.05)=14 Risk Seeking	C(-100;0.05)=-8 Risk Aversion
High probability	C(100;0.95)=78 Risk Aversion	C(-100;0.95)=-84 Risk Seeking

[a] The source of this table is Tversky and Wakker (1995).

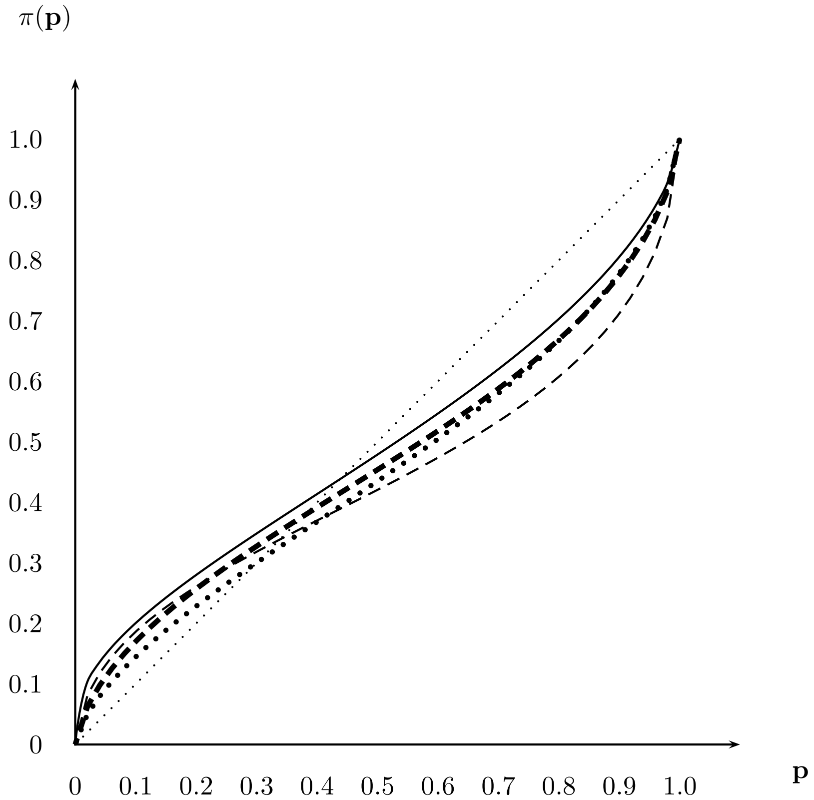
experiment results of Kahneman and Tversky (1992). For illustration, consider the example (100;0.05)=14 for a gain with low probability. The certainty equivalent of 14 stems from the experiment and is the median of the certainty equivalents of all participants. The expected value of the lottery (100;0.05 | 0;0.95) is $0.05 \cdot 100 + 0 \cdot 0.95 = 5$. Apparently, the certainty equivalent is bigger than the expected value. Thus, the subjects of the experiment exhibited a risk seeking attitude for the combination of gain and low probability. The fourfold pattern of risk was also detected in various other studies, like e.g. Fishburn and Kochenberger (1979), Kahneman and Tversky (1979), Hershey and Schoemaker (1980), Payne, Laughhunn and Crum (1981), Cohen, Jaffray and Said (1987), and Wehrung (1989).

Figure 2 shows some probability weighting functions featuring the fourfold pattern. One property of the functions, which has not been mentioned yet is that a change of moderate probabilities, like e.g. from 0.35 to 0.40, has less impact than a change at the endpoints, like e.g. from 0.95 to 1.00. This aspect is simply visible by the steep turns of the functions at the endpoints. In general, the weighting functions have an inverted S-shape and cross the 45-degree line at probabilities of around 0.3 to 0.4. This implies that low probabilities are overweighted and high probabilities are underweighted.

The functions in Figure 2 stem from different studies/estimates:

- $\pi(p) = \frac{p^{\varphi}}{(p^{\varphi}+(1-p)^{\varphi})^{(1/\varphi)}}$ from Kahneman and Tversky (1992) with estimates of $\varphi = 0.61$ for gains and $\varphi = 0.69$ for losses (dashed curves: normal=gains and bold=losses)
- $\pi(p) = \frac{\delta_p p^{\varphi}}{(\delta_p p^{\varphi}+(1-p)^{\varphi})}$ Fox and Tversky (1995) with estimates for φ from 0.69 to 0.72 and for δ_p from 0.76 to 0.77 (dotted curve)

Figure 2: Probability Weighting Functions



- $\pi(p) = exp(-\beta(-ln(p))^{\alpha})$ Prelec (1998), with a suggestion of $\alpha = 0.65$ and $\beta = 1$ (continuous curve)

The functions have the property that, for particular parameters, they reduce to the function $\pi(p) = p$. This property will be used in the estimation procedure later on. In order to show this, consider the weighting function from Prelec (1998), with $\beta = 1$ and $\alpha = 1$,[34] i.e.

$$\pi(p) = exp(-\beta(-ln(p))^{\alpha}) = exp(-(-ln(p))) = p.$$

The above equation shows that, with this parameter setting, the weighting function reduces to the identity function, i.e. as used in Expected Utility Theory. The only drawback of Prelec's proposal is that the function is not defined for $p = 0$. Thus, one has to determine $\pi(0) = 0$.

[34]Figure B.1 in the Appendix shows Prelec's probability weighting function for different α - the parameter α is restricted to the interval [0;1].

In the remaining part we will deviate in some respects from the definition of Kahneman and Tversky (1992). In the sense that the evaluation procedure under Cumulative Prospect Theory will be the same as under Rank-Dependent Utility. This implies that gains and losses are treated in the same way concerning the probability weighting function, i.e. there will be only one function for both domains. Thus, the value of a lottery $(x_1; p_1|...|x_n; p_n)$ under "quasi" Cumulative Prospect Theory will be determined by

$$V = \sum_{i=1}^{n} h_i(p) \cdot v(x_i),$$

where

$$h_i(p) = \pi\left(\sum_{j=i}^{n} p_j\right) - \pi\left(\sum_{j=i+1}^{n} p_j\right),$$

with e.g.

$$\pi(p) = exp(-(-ln(p))^{\alpha}) \quad \text{and} \quad v(\tilde{x}) = \begin{cases} (\tilde{x})^{\rho}, & \tilde{x} = x - r_{ref} \geq 0 \\ -\lambda_L(-\tilde{x})^{\tau}, & \tilde{x} = x - r_{ref} < 0. \end{cases}$$

This kind of representation will be used within the Method of Simulated Moments approach, as it permits to compare Expected Utility, Rank-Dependent Utility and Cumulative Prospect Theory in one estimation run. The idea is to estimate the crucial parameters in the setting, which are responsible for the differences between these theories. This implies the following cases for the estimation.

- If α and r_{ref} are significantly different from 1 and 0, respectively, then Cumulative Prospect Theory[35] fits the data best.

- If α is significantly different from 1 and r_{ref} is not significantly different from 0, then Rank-Dependent Utility is the preferred theory by the data.

- If α and r_{ref} are not significantly different from 1 and 0, respectively, then the data suggests that Expected Utility Theory is the prevailing theory.

[35] As the deviations from Cumulative Prospect Theory of Kahneman and Tversky (1992) do not violate any basic principles of this theory, we believe that the conclusions of the estimation can be drawn for the original version as well.

4 Dynamic Choice in the Model

In the model described in Section 2 a financial independent household exists from the age of 20 to the age of 90. This implies there are 71 periods in which the household has to decide how much investments it should put in liquid and illiquid assets each year, thereby determining the consumption each period. But, before a household can make its investment decisions, the risk of income, as there are transitory shocks and persistent shocks each year, has to resolve. Furthermore, there is also the risk of surviving the current period, which also influences the savings decisions.

In the model of Section 2, the dynamic choice problem over the life cycle is treated as an intra-personal game, i.e. a household at each period t is a player, denoted "Self t", in this game. Each Self t is in a particular state, which is determined by the decisions of the players and the resolution of risk in the previous periods. At the beginning of period t, the persistent and the transitory shocks are realized, thereby the income of the current year, i.e. age-specific income plus persistent shock plus transitory shock, is determined. Thus, Self t's state is defined by the asset holdings, the income and age-specific household size. After the perception of its situation the player has to make its investment decisions and is finally faced with the risk of survival of the current period. For optimal behavior in the context of the whole life cycle each player has also to take into account the decisions of all the subsequent players.

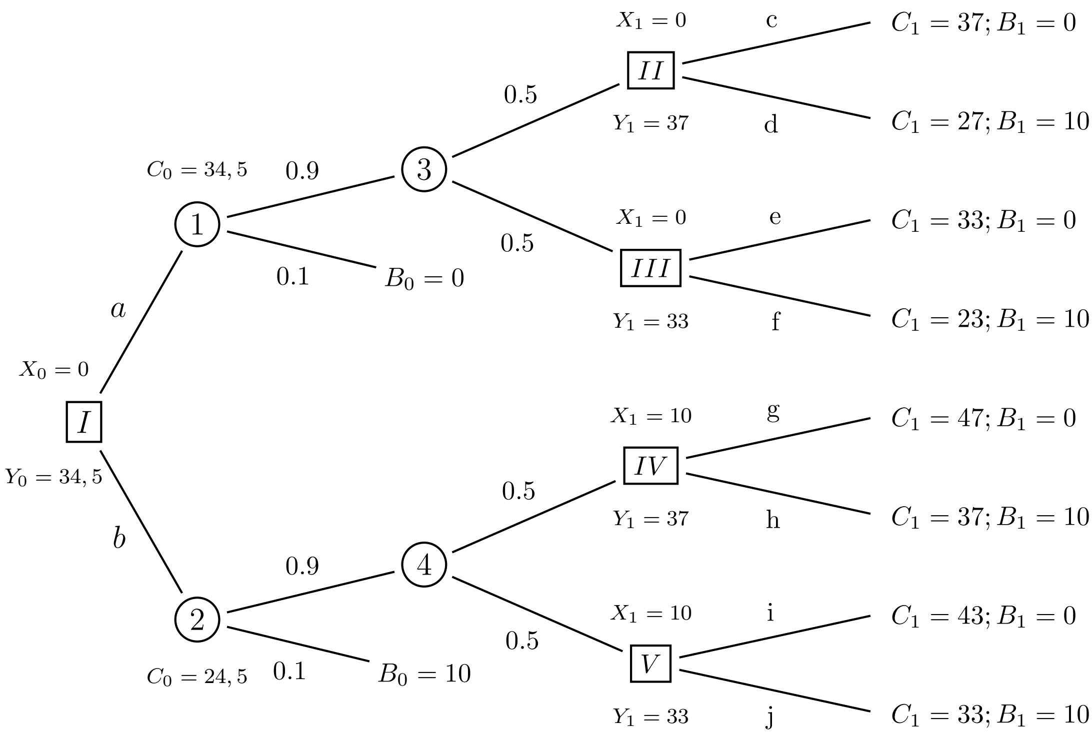
In order to illustrate the dynamic choice problem and how it is solved, consider the following example, which is very reduced in comparison to the problem arising in the model of Section 2, but still exhibits the same pattern. In the example, which is displayed in form of a decision tree in Figure 3, a household can live two periods, with a probability of 0.9 surviving the first year. The household consists of one person and the age-specific incomes are equal to €34 500 in the first and €35 000 in the second period. At the beginning of the second period, a transitory shock arises either reducing the income by €2 000 or increasing income by €2 000, each with probability 0.5. In contrast to the model, persistent shocks are omitted.

For the sake of simplicity the household has in each period only two options to choose, either saving an amount of €10 000 in liquid assets X (moving downwards in the tree) or saving nothing and consume all income (moving upwards in the tree). The rectangles in Figure 3 denote a node where a decision is made by the household, i.e. it chooses up or down. The possible decisions are represented by the small letters a to j. The circles describe nodes where nature is moving, i.e. at these points risk is incorporated and the corresponding probabilities of each event are given in the middle of the branches. The risk of survival is displayed at the nodes 1 and 2, with a 0.9 chance of surviving and a hazard of 0.1 to die. The transitory shocks occur at the nodes 3 and 4, each with probability 0.5. This example corresponds to an intra-personal game with two players. Self 2 - the player in the second period - has to make its decision in one of the four nodes II, III, IV or V. It depends on the decision of Self 1 - the player in period 1 - and how nature was moving at the chance nodes 1-4, in which of these four nodes Self 2 will end up and has to make its decision. Self 1 decides at node I and has to take into account the risky components at the nodes 1 to 4 and the decisions Self 2 would make when it is in the various possible states.

Finally, assume that the last thing happening in period 1 is the resolution of the risk of survival and the first thing in period 2, if the household lives in the second year, is the move of nature resolving the transitory shock.

For the evaluation of the tree let $u(C) = C^{0.6}$ be the utility function for consumption, and $b(B) = B^{0.5}$ be the bequest function. The probability function is $\pi(p) = exp(-(-ln(p))^{0.5})$, with $\pi(0) = 0$. For this example discounting of the second period outcomes is omitted. Furthermore, assume that a hypothetical reference point is equal to 0, which implies that the evaluation procedure of Rank-Dependent Utility and Cumulative Prospect Theory is identical, as there are only gains and no losses. As already stated above such a game can be solved via Folding Back, i.e. the evaluation starts at the right side of the decision tree. Thus, start with the choice problems of Self 2 at the right upper rectangle in the tree, namely decision node II. At this node the situation of Self 2 is defined by having an income of $Y_1 =$ €37 000 (€35 000 from the age-specific income plus €2 000 from a positive transitory shock) and liquid assets of $X_1 =$ €0 - this information is given above and below each decision node. Player 2 has at this point two options, either choosing the upper branch c or the lower branch d. The value V of such a branch is simply given by the sum of the utility from consumption and bequest. Hence, the value of branch c at

Figure 3: Two Period Example



node II, V^c_{II}, and of branch d, V^d_{II}, are given by

$$V^c_{II} = u(37000) + b(0) = 37000^{0.6} + 0 = 550.71$$

$$V^d_{II} = u(27000) + b(10000) = 27000^{0.6} + 10000^{0.5} = 555.85.$$

These values state that, if Self 2 is at decision node II, it will decide to choose branch d, as $V^d_{II} > V^c_{II}$, i.e. at this point the player would spend €27 000 of its disposable money for consumption and save €10 000 for bequest. The same is done for the rest of the possible decision nodes of Self 2. In summary the possible decisions of Self 2 are

- II: $V^c_{II} = 550.71 < 555.85 = V^d_{II} \Rightarrow$ player 2 chooses d
- III: $V^e_{III} = 514.17 > 514.03 = V^f_{III} \Rightarrow$ player 2 chooses e
- IV: $V^g_{IV} = 635.71 < 650.71 = V^h_{II} \Rightarrow$ player 2 chooses h

- V: $V_V^i = 602.67 < 614.17 = V_V^j \Rightarrow$ player 2 chooses j.

Thus, after the decisions of Self 2 at each possible node are known, the decision of Self 1 has to be considered. The problem at this point is that player 1 has to take into account the risk stemming from the transitory shock and the hazard of death, i.e. it has to deal with a lottery. In the approach applied here, the risky components are evaluated separately, i.e. first the risky lottery occurring from the income shock and then the risky lottery arising from the risk of survival are considered. This approach is called the Certainty Equivalent Mechanism in the literature, which is implicitly used in the Folding Back procedure applied here. Later on we will describe another way, usually applied in economics, of dealing with more than one risky component appearing in a subsequent fashion.

Now, consider chance node 3, on the time line this is the very beginning of period 2, when nature is moving and realizes a positive or a negative income shock, each with a probability of 0.5. It is known from the examination of the decision behavior of Self 2, that a positive income shock leading to node II generates the utility of $V_{II}^d = 555.85$, as Self 2 would choose branch d at this node. A negative shock, which leads to the decision node III, provides the utility $V_{III}^e = 514.17$. Thus, the lottery at chance node 3 is given by $(555.85; 0.5|514.17; 0.5)$, i.e. a chance of 0.5 ending up with an utility of either 555.85 or 514.17. This lottery is now evaluated according to Rank-Dependent Utility. First of all, as described in the previous section, the outcomes are ranked in an ascending fashion - that is $514.17 < 555.85$. The value of this lottery, denoted by v_3, is defined by[36]

$$\begin{aligned}
v_3 &= [(\pi(1) - \pi(0.5)] \cdot 514.17 \\
&\quad + [(\pi(0.5) - \pi(0)] \cdot 555.85 \\
&= [exp(-(-ln(1)^{0.5}) - exp(-(-ln(0.5)^{0.5})] \cdot 514.17 \\
&\quad + [exp(-(-ln(0.5)^{0.5}) - 0] \cdot 555.85 \\
&= 532.3.
\end{aligned}$$

The same procedure is applied to the lottery occurring at chance node 4. This

[36]Recall, the weighting function from Prelec (1998), namely $\pi(p) = exp(-\beta(-ln(p))^{\alpha})$, is not defined for $p = 0$, therefore one has to assume that $\pi(0) = 0$.

yields $v_4 = 630.06$. Now, let us proceed to the nodes 1 and 2, which represent the risk of survival. In order to arrive at chance node 1 Self 1 has to choose branch a. This implies that all the income of period 1 is used for consumption, i.e. $C_0 =$ €34 500. The lottery arising at node 1 is surviving period 1 with probability 0.9 and yielding the value of the lottery occurring at chance node 3, $v_3 = 532.3$, or not surviving and ending up with a bequest of $B_0 =$ €0, as nothing was saved by Self 1, with probability 0.1. This lottery $(0; 0.1|532.3; 0.9)$ is again evaluated by Rank-Dependent Utility and results in the value $v_1 = 384.76$. Note, that the value v_3 is already expressed in terms of utility, but $B_0 =$ €0 has still to be transformed by the bequest function, before the outcomes are ranked. This aspect is important, because it is necessary to express all outcomes in the same terms before the ranking is executed. Hence, the lottery at node 2 is correctly defined by $(100; 0.1|630.06; 0.9)$ and has the value $v_2 = 483.14$.

Now, there is only the decision node I left and Self 1 has to choose the branch (either a or b), which provides the highest utility to it over the whole time horizon. Branch a leads to the value v_1 and to the consumption of $C_0 =$ €34 500, i.e. the value of this branch is

$$V_a = u(C_0) + v_1 = 34500^{0.6} + 384.76 = 912.83.$$

On the other side, branch b provides the value v_2 and the consumption of $C_0 =$ €24 500, i.e.

$$V_b = u(C_0) + v_2 = 24500^{0.6} + 483.14 = 913.17.$$

As the value of branch b is bigger than the value of branch a, i.e. $V_a < V_b$, Self 1 will choose b.

This course of action, namely Folding Back, can also be executed with Expected Utility Theory or Cumulative Prospect Theory. Though, for the latter, it is necessary to determine a reference point and define the utility function according to this point, like shown in Section 3. It would also be possible to use a reference point for the bequest, but we will stick here to the interpretation that the reference point is related to a particular level of consumption.

As mentioned above, there is also another way to solve dynamic choice problems, like the one in the example above. This approach is usually applied in economics and uses the Reduction of Compound Lotteries Axiom, which is an axiom of Expected Utility Theory.

> *Reduction of Compound Lotteries Axiom*: Let $X = (x_1; p|x_2; (1-p))$ and $Y = (y_1; q|y_2; (1-q))$ be simple one-stage lotteries, and $A = (X; r|Y; (1-r))$ a compound lottery, leading with probability r to lottery X and with probability $1-r$ to Y. Then, the Reduction of Compound Lotteries axiom states that
>
> $$A \sim (x_1; p \cdot r|x_2; (1-p) \cdot r|y_1; q \cdot (1-r)|y_2; (1-q) \cdot (1-r)).$$

Hence, the axiom permits that probabilities of different, but connected, stages can be multiplied. The main implication is that a decision maker is indifferent according to the timing of the resolution of risk. This also includes that an individual is indifferent if nature is resolving the existing risk in one move or more. The application of the axiom results in lotteries with only one stage, i.e. nature is moving only once.

The intuitive assumption behind this axiom is that an individual has no other economic activities or decisions (including consumption/savings decisions) to undertake in the meantime (i.e. between the resolution of risk of different stages), so that it has no incentive to prefer single-stage lotteries over compound lotteries because of impatience and/or planning benefits - Machina (1989, pp. 1625-26).

There are also others, like Kreps and Porteus (1975), arguing that there is no reason to make this reduction assumption, if a sufficiently large time span is between different stages of a dynamic problem. This is the case in the model described in Section 2, as between two decisions passes a whole year.

The application of the Reduction of Compound Lotteries Axiom implies, that the decisions made along a decision tree are summarized in a strategy, which contains a choice for all possibly occurring decision nodes. For example in Figure 3 a decision maker can choose a at point I, if it survives the first period and a positive shock occurs, then it could take branch c at node II. In the case of survival and a negative shock it could choose e. These choices a,c, and e are called a strategy - in the following denoted by <ace>. A strategy contains a decision for each possible outcome of the risky elements in the dynamic problem. In the example of Figure 3 a decision maker can choose between 8 such strategies and will choose the one generating the highest utility according to its preferences. The strategies and the corresponding lotteries are displayed in Figure 4.

At the chance nodes 1-8 the lotteries belonging to the various strategies are exhibited, e.g. for strategy <ace> the three possible events are:

1. surviving the first period (probability of 0.9) and realization of a positive shock (probability of 0.5), i.e. a 0.9·0.5=0.45 chance[37] to end up with a consumption of C_0 = €34 500 in the first period, C_1 = €37 000 in the second period and a bequest of B_1 = €0 after death in the second period,

2. surviving the first period and realization of a negative shock, i.e. a chance of 0.45 for C_0 = €34 500, C_1 = €33 000 and B_1 = €0,

3. not surviving the first period with a probability of 0.1 yielding a consumption of C_0 = €34 500 in the first period and a bequest of B_0 = €0 after death in the first period.

Now, these strategies are evaluated by Rank-Dependent Utility. Note, that at first the outcomes of each branch have to be transformed by the corresponding utility/bequest function and then the sum of these utilities is used for the ranking, i.e.

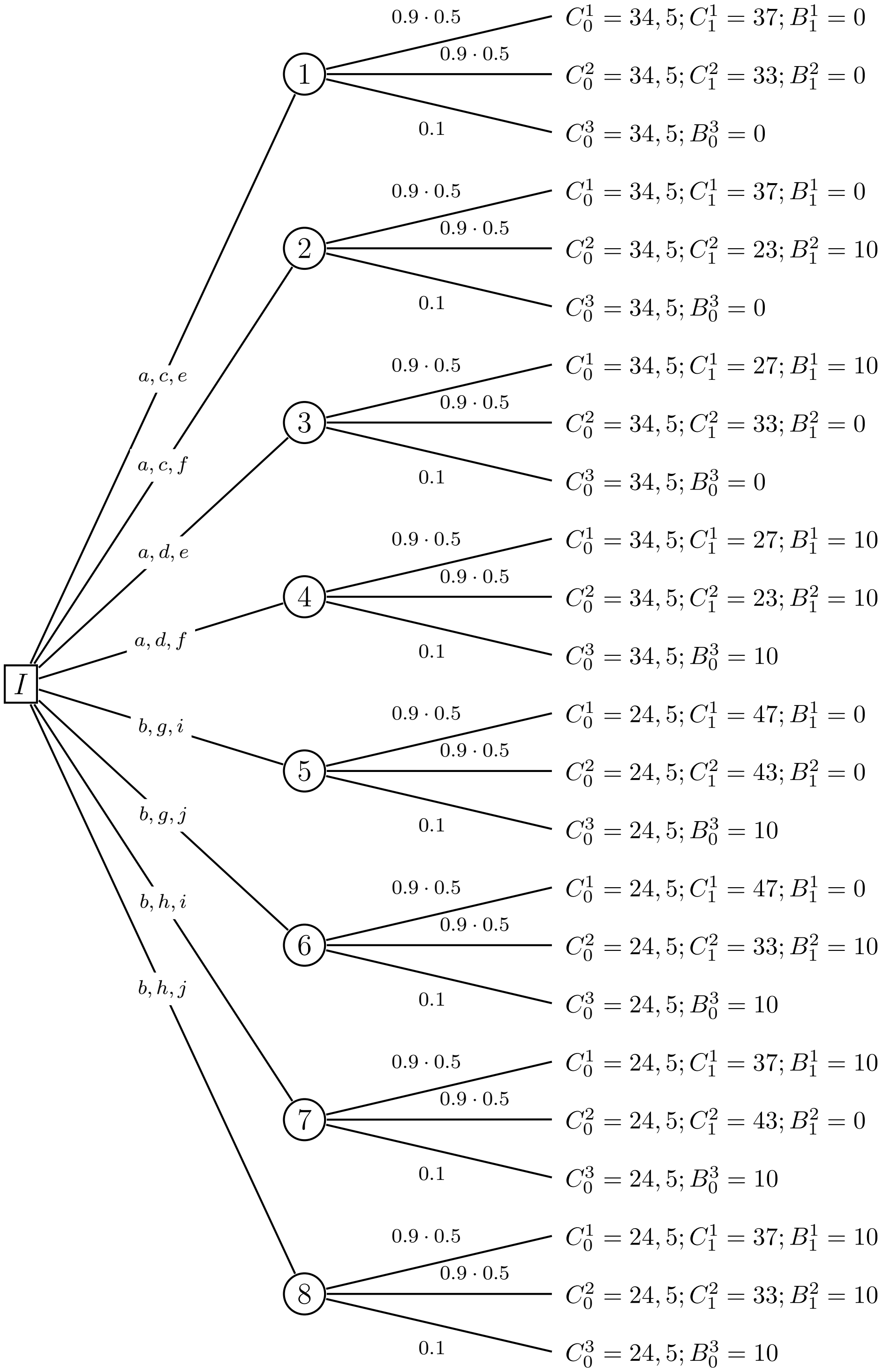
$$u(C_0^1) + u(C_1^1) + b(B_0^1) \underset{>}{\overset{\leq}{=}} u(C_0^2) + u(C_1^2) + b(B_0^2) \underset{>}{\overset{\leq}{=}} u(C_0^3) + b(B_0^3)$$

has to be clarified before starting the evaluation. The value of strategy <ace>, denoted by V_{ace}, is given by

$$\begin{aligned} V_{ace} &= [1 - \pi(0.9)] \cdot u(C_0^3) \\ &\quad + [\pi(0.9) - \pi(0.45)] \cdot (u(C_0^2) + u(C_1^2) + b(B_1^2)) \\ &\quad + [\pi(0.45) - \pi(0)] \cdot (u(C_0^1) + u(C_1^1) + b(B_1^1)) \\ &= 914.67. \end{aligned}$$

[37]The calculation of the event "surviving first period and positive income shock" is possible because of the Reduction of Compound Lotteries Axiom.

Figure 4: Decision Tree in Normal Form



The values of the remaining strategies are

$$
\begin{aligned}
V_{acf} &= 914.63 \\
V_{ade} &= 916.78 \\
V_{adf} &= 916.73 \\
V_{bgi} &= 906.89 \\
V_{bgj} &= 910.50 \\
V_{bhi} &= 913.03 \\
V_{bhj} &= 916.64.
\end{aligned}
$$

Comparing the results, a decision maker would choose strategy <ade>, if the dynamic problem is evaluated according to the procedure suggested by the Reduction of Compound Lotteries Axiom. However, the choice of this strategy differs from the choice obtained by the Backward Induction routine. As the results are not identical, the question arises, which procedure is the one to apply if preferences according to Rank Dependent Utility or Cumulative Prospect Theory prevail.

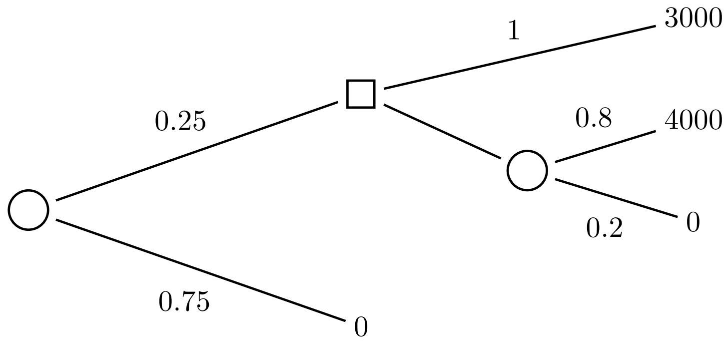
First, in order to answer this question, there is empirical evidence indicating that individuals do not obey the Reduction of Compound Lotteries Axiom, see e.g. Ronen (1971), Snowball and Brown (1979), Kahneman and Tversky (1979, 1981), Holler (1983) or Starmer and Sugden (1991). In an example from Kahneman and Tversky (1979) the participants were faced with the following problem.

> **Problem 7:** Consider the following two-stage game. In the first stage, there is a probability of 0.75 to end the game without winning anything, and a probability of 0.25 to move on to the second stage. If you reach the second stage you have a choice between:
>
> C′: (4000;0.80 | 0;0.20) [22]
>
> D′: (3000;1) [78*]
>
> Your choice must be made before the game starts, i.e. before the outcome of the first stage is known.

Again, the values in the squared brackets give the percentages of each option chosen by 141 participants. The graphical representation of Problem 7 in extensive form, i.e. without the usage of the Reduction of Compound Lotteries Axiom, is displayed in Figure 5. The decision tree in normal form, i.e. the sequential lottery was first reduced to a one-stage lottery by applying the Reduction of Compound Lotteries Axiom, is shown in Figure 6. Note, that the decision occurring in Figure 6 is the same as in Problem 4, where the majority preferred the lottery $(4000; 0.2|0; 0.8)$. Although, the Problems 4 and 7 are regarded as equivalent by the Reduction of Compound Lotteries Axiom, the choices in the two problem sets differ. Hence, it can be concluded that people do not obey this axiom, at least most of the people in the experiment of Kahneman and Tversky (1979). [38]

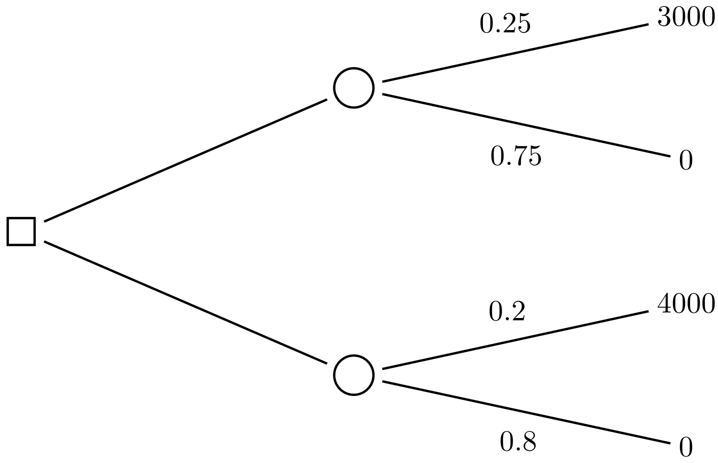
Furthermore, a central aspect is that strategies found with the Reduction of Compound Lotteries Axiom, if Rank-Dependent Utility or Cumulative Prospect Theory are used for evaluation, can be dynamically inconsistent. In this thesis Dynamic Consistency is considered as the feature that the choice at a any decision node, when arriving at this node, is equal to the planned choice made at the beginning of the dynamic problem - see also e.g. Machina (1989). Another definition of Dynamic Consistency, like in Epstein (1992) or Sarin and Wakker (1998), is equality in choice in both representations of a dynamic problem, i.e. extensive (Figure 3) and normal form (Figure 4). This

Figure 5: Problem 7 in sequential form



[38]Kahneman and Tversky state that this is a possible conclusion, but there could also be another unknown effect being responsible for the violation of the principles of Expected Utility Theory.

Figure 6: Problem 7 with Reduction Axiom



definition actually implies that the Reduction of Compound Lotteries Axiom holds and therefore is rather inappropriate for the approach put forward in this thesis.

The possible Dynamic Inconsistency arising with Rank-Dependent Utility preferences[39] and the application of the Reduction of Compound Lotteries Axiom can be observed in the problem given in Figure 7, which stems from Jaffray and Nielsen (2006). For the example the probability weighting function is again given by $\pi(p) = exp(-(-ln(p))^{0.5})$ and the values at the leaves of the decision tree are already expressed in terms of utility.

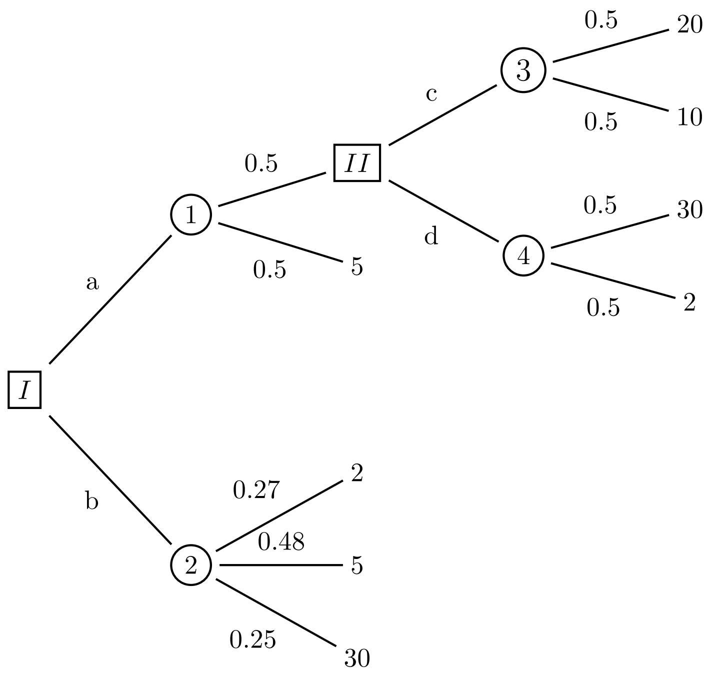
There are - from the root (rectangle I) of the decision tree - three strategies <ac>, <ad> and <b>, which evaluated by Rank-Dependent Utility with the Reduction of Compound Lotteries Axiom lead to

$$V_{ac} = (1 - \pi(0.5)) \cdot 5 + (\pi(0.5) - \pi(0.25)) \cdot 10 + \pi(0.25) \cdot 20 = 10.26$$
$$V_{ad} = (1 - \pi(0.75)) \cdot 2 + (\pi(0.75) - \pi(0.25)) \cdot 5 + \pi(0.25) \cdot 30 = 11.46$$
$$V_{b} = (1 - \pi(0.73)) \cdot 2 + (\pi(0.73) - \pi(0.25)) \cdot 5 + \pi(0.25) \cdot 30 = 11.41.$$

Hence, a decision maker, using this evaluation approach, would choose strategy <ad>. However, the choice of this strategy implies that if the decision maker happens to find itself, after the move of nature, at the decision node II,

[39] Again, the same holds for Cumulative Prospect Theory - it just has to be assumed, that the utility values were obtained after the transformation with the reference point.

Figure 7: Example from Jaffray and Nielsen (2006)



it has no longer an incentive to obey the initial strategy of choosing d. This can be seen by evaluating the lotteries at chance nodes 3 and 4, namely

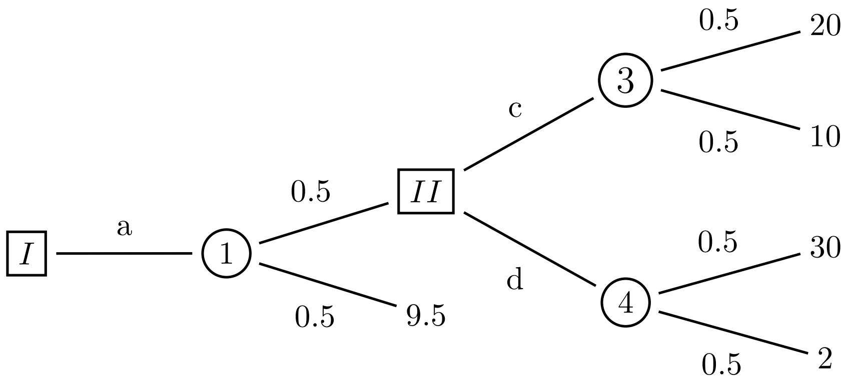
$$V_3 = (1 - \pi(0.5)) * 10 + \pi(0.5) * 20 = 14.35$$
$$V_4 = (1 - \pi(0.5)) * 2 + \pi(0.5) * 30 = 14.18.$$

Lottery 3 yields a bigger value as lottery 4, thus a decision maker would select branch c. This obviously differs from the strategy chosen at the very beginning of the decision tree. This kind of behavior pattern is dynamically inconsistent and can occur if Rank-Dependent Utility is used in combination with the Reduction of Compound Lotteries Axiom in dynamic problems with more than one stage requiring a decision.

A further problem arising if these preferences prevail and the Reduction of Compound Lotteries is used, is that past or irrelevant consequences do matter in the evaluation procedure. The following example illustrating this aspect stems also from Jaffray and Nielsen (2006) - see Figure 8.

From the examination of Figure 7 it is known that a decision maker with

Figure 8: Violation of Consequentialism



Rank-Dependent Utility preferences applying the Reduction of Compound Lotteries axiom would choose strategy <ad>, as $10.26 = V_{ac} < V_{ad} = 11.46$ - for now the lower subtree is omitted. Figure 8 exhibits a similar dynamic problem with the only difference that the outcome 5 is replaced by the outcome 9.5. The evaluation of the strategies in this problem, with V'_{ac} and V'_{ad} denoting the values of the strategies in the example of Figure 8, yields

$$V'_{ac} = (1 - \pi(0.5)) \cdot 9.5 + (\pi(0.5) - \pi(0.25)) \cdot 10 + \pi(0.25) \cdot 20 = 12.80$$
$$V'_{ad} = (1 - \pi(0.75)) \cdot 2 + (\pi(0.75) - \pi(0.25)) \cdot 9.5 + \pi(0.25) \cdot 30 = 12.70.$$

In this example a decision maker would choose strategy <ac>. This implies that in the evaluation procedure with the Reduction of Compound Lotteries Axiom the outcome change from 5 to 9.5 in the first stage changes the decision in the second stage, namely from d to c. This aspect violates the principle that irrelevant consequences should not influence the decision process. In the literature this principle is known as Consequentialism - see e.g. Hammond (1988) - and is automatically satisfied if Folding Back is applied to solve a decision tree. The intuitive question behind Consequentialism for the examples above is why should the outcomes 5 or 9.5 have any influence on the decisions in the second stage, although the first stage has already resolved and these two outcomes did not occur.

In summary, we will **not** use the Reduction of Compound Lotteries Axiom, due to these aspects. Hence, the answer to the question posed above, which procedure should be applied in a dynamic setting like in the example displayed in Figure 3, if preferences according to Rank-Dependent Utility or Cumulative Prospect Theory prevail, is that the Backward Induction routine should be used.

There is also a paper by Segal (1990), who suggests the usage of Rank-Dependent Utility preferences without the Reduction of Compound Lotteries Axiom, at least for two-stage lotteries. In contrast to the example used in this section Segal (1990) refers to two-stage lotteries as lotteries having other lotteries as outcomes.

5 Critique on Backwards Induction

The procedure of solving dynamic problems under Non-Expected Utility Theory preferences with Backward Induction, without the usage of the Reduction of Compound Lotteries Axiom, has some implications on the critique given in the literature, especially concerning Rank-Dependent Utility and Cumulative Prospect Theory.[40]

5.1 Stochastic Dominance

Before we consider the critique against Folding Back as outlined in Machina (1989), let us consider an argument put forward by Jaffray and Nielsen (2006). In their paper they also use Rank-Dependent Utility for dynamic decision making, but in contrast to the approach proposed in this thesis they use the Reduction of Compound Lotteries Axiom. Their procedure for solving dynamic problems is a form of Rolling Back with a rejection mechanism for dominated strategies. The argument by Jaffray and Nielsen (2006) is that Folding Back and Rank-Dependent Utility preferences alone lead to first-order-stochastically dominated strategies. Recall, each strategy in a dynamic setting, like in Figure 3, implies a one-stage lottery, see Figure 4, if the Reduction of Compound Lotteries Axiom is applied. Therefore, First-Order-Stochastic Dominance for strategies can be defined in terms of the corresponding one-stage lottery.

[40] At least for the version of Cumulative Prospect Theory used in this thesis.

(Strict) First-Order-Stochastic Dominance: The lottery L (strictly) stochastically dominates lottery L′ in first order if:

$$\sum_{i:u(c_i)\leq\kappa} p_i \leq \sum_{i:u(c_i)\leq\kappa} p'_i, \text{ for all } \kappa \in \Re \text{ and}$$
$$\left(\sum_{i:u(c_i)\leq\kappa} p_i < \sum_{i:u(c_i)\leq\kappa} p'_i, \text{ for at least one } \kappa \in \Re\,.\right)$$

In the example of Figure 7 strategy <b> would be chosen by a decision maker with Rank-Dependent Utility preferences applying the Backward Induction procedure. Following Jaffray and Nielsen (2006), this strategy is first-order-stochastically dominated by the strategy <ad>. This can be seen by comparing the cumulative distributions of the two lotteries, namely $(2; 0.27|5; 0.48|30; 0.25)$ for strategy <b> and $(2; 0.25|5; 0.5|30; 0.25)$ for strategy <ad>. Here, for κ only the possible outcomes 2, 5 and 30 are relevant.

$$\sum_{i:u(c_i)\leq 30} p_i^{ad} = 1 \leq 1 = \sum_{i:u(c_i)\leq 30} p_i^{b}$$
$$\sum_{i:u(c_i)\leq 5} p_i^{ad} = 0.75 \leq 0.75 = \sum_{i:u(c_i)\leq 5} p_i^{b}$$
$$\sum_{i:u(c_i)\leq 2} p_i^{ad} = 0.25 < 0.27 = \sum_{i:u(c_i)\leq 2} p_i^{b}$$

According to the definition the lottery arising from strategy <b> is first-order-stochastically dominated by the lottery provided by strategy <ad>. Thus, Jaffray and Nielsen (2006) conclude that the Backward Induction procedure selects a dominated strategy. But, this approach to determine First-Order-Stochastic Dominance uses the Reduction of Compound Lotteries Axiom to reduce the two-stage lottery of strategy <ad> to a one-stage lottery in order to calculate the cumulative distributions. This is necessary, as First-Order-Stochastic Dominance is a measure between one-stage lotteries. Due to this aspect it is questionable if First-Order-Stochastic Dominance is an appropriate measure to compare a one-stage lottery with a two-stage lottery.[41]

[41] Later on we will introduce a concept from Segal (1990) dealing with stochastic dominance for two-stage lotteries, that is implicitly based on First-Order-Stochastic Dominance, but does not require the usage of the Reduction of Compound Lotteries Axiom.

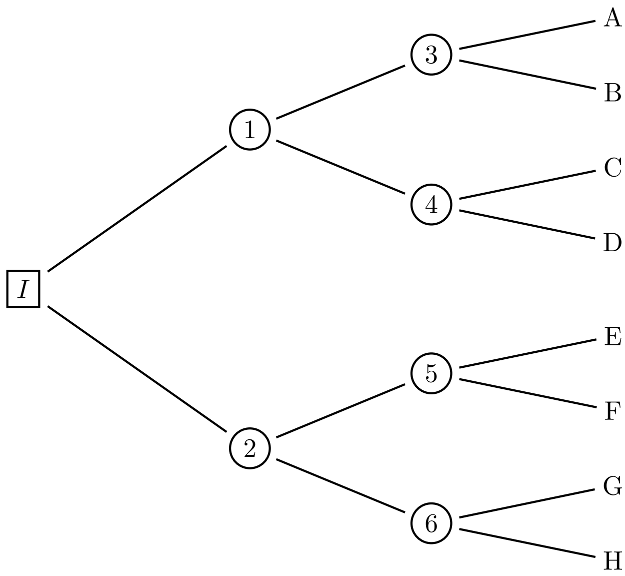
Independent of this question is the fact that a strategy not chosen from the procedure of Rolling Back the decision tree could be dynamically inconsistent, like shown in the previous section for the example of Figure 7.[42] This implies that such a strategy, even if it is dominant in the sense of the definition of First-Order-Stochastic Dominance, will not be obeyed by at least one of the players deciding at intermediate or final stages. Thus, in our opinion the approach to assess strategies should be, first to determine all dynamically consistent strategies and then to choose among these the one providing the highest overall value, i.e. the value over all stages. Especially in the dynamic setting occurring in our model, when the players of each stage are different Selves of the same decision maker. We do not see any incentive, why one should deviate from this principle, even if there are strategies first-order-stochastically dominating the one chosen by the Backward Induction routine. The reason is that every Self acts optimally given the decisions of the others Selves. This way the equilibrium was defined in Section 2 and results in the fact that no player has an incentive to deviate from the strategy obtained by Backward Induction. Hence, given the optimal behavior, i.e. selecting the option at each decision node providing the highest overall value, and Dynamic Consistency inevitably leads to the strategy found via Backwards Induction.

Another problem, which is similar to the problem put forward by Jaffray and Nielsen (2006), but differs in the origin of the underlying decision tree, concerns two-stage lotteries with solely chance nodes. See Figure 9 for a decision between two two-stage lotteries, where $A, ..., H \in \Re$. The problem is that the evaluation process of such two-stage lotteries under Non-Expected Utility Theory preferences using the Reduction of Compound Lotteries Axiom can also yield first-order-stochastically dominated lotteries. In contrast to the Jaffray-Nielsen example above it is not possible to argue that dynamic consistency is not satisfied, as there is only one decision to make.

For such two-stage lotteries, having other lotteries as outcomes, Segal (1990) proposes alternative ways to compare these with regard to stochastic dominance, but without the usage of the Reduction of Compound Lotteries Axiom. This is mainly due to the empirical evidence suggesting that individuals do not always obey this axiom. Furthermore, this new definition of stochastic

[42] At least as long as there are no commitments between the players acting at the different stages.

Figure 9: Two-stage lotteries with only chance nodes



dominance for two-stage lotteries leads to a more general setting distinguishing between one and two-stage lotteries. Segal (1990) also shows that one of these new concepts of stochastic dominance, which he calls Compound Dominance, is compatible with Rank-Dependent Utility.

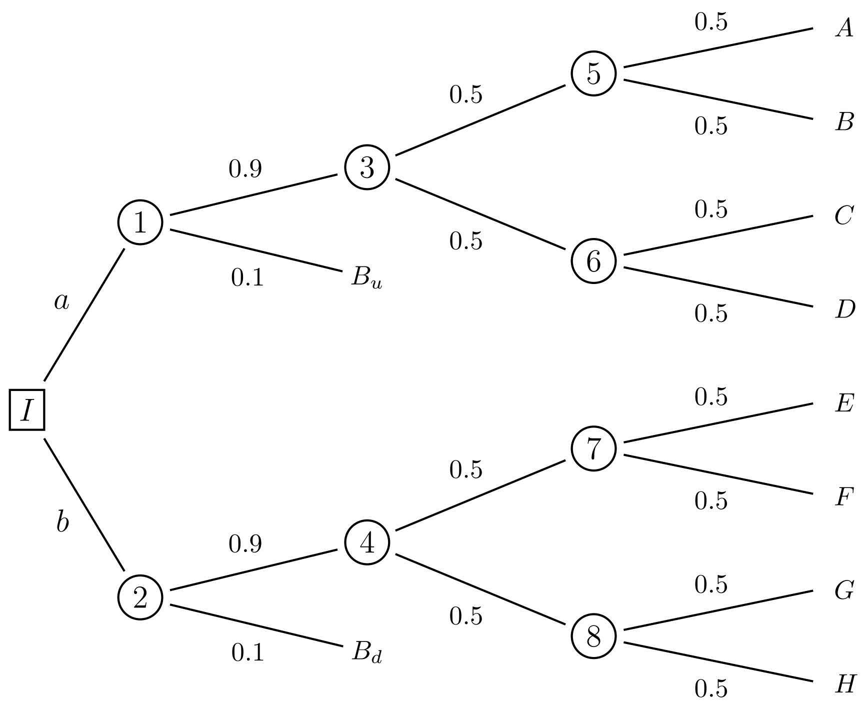
However, in the dynamic setting occurring in our model there are three-stage lotteries, i.e. three chance nodes between two consecutive decision nodes. For an example, see Figure 10, where a decision maker has to choose between two three-stage lotteries at choice node I, i.e. either choosing branch a or b. The three stages are determined in the model by

1. the chance of survival - chance nodes 1 and 2,
2. the risk of being in the high (upper branch) or in the low (lower branch) income class [43] - chance nodes 3 and 4, and
3. the risk of a high or a low transitory income shock - chance nodes 5 to 8.

Note, that the probabilities - in the upper and lower tree in Figure 10 - are identical at each branch in the three-stage lotteries. This circumstance stems

[43]This is the part of income represented by the Markov Process, which approximates an AR(1) process - this will be described in more detail in the calibration section

Figure 10: Three-stage Lottery Example



from the fact, that in each period the risk of survival and income are the same at each decision node, the only difference are the investments in the liquid and illiquid assets.

In order to determine stochastic dominance for lotteries, like those in Figure 10, we adopt a concept of Segal (1990) and extend it to such three-stage lotteries. The concept of Two-Stage-Stochastic Dominance by Segal (1990) is based on an idea of Kamae, Krengel and O'Brien (1977), and can be defined as[44]:

> *(Strict) Two-Stage-Stochastic Dominance*: The two-stage lottery $A = (X_1; q_1| \ldots | X_m; q_m)$ stochastically dominates the two-stage lottery $B = (Y_1; q_1| \ldots |Y_m; q_m)$, where X_i and Y_i, for $i = 1, ..., m$, are simple (one-stage) lotteries, by (strict) two-stage stochastic dominance, if and only if X_i first-order-stochastically dominates Y_i for all $i = 1, ..., m$ (and X_i strictly first-order-stochastically dominates Y_i for at least one $i = 1, ..., m$).

[44]This is not exactly the definition given by Segal (1990), but captures the concept.

The three-stage lotteries in Figure 10 for the choices a and b are denoted by L_a and L_b, respectively, and can be written as

$$L_a = \left(\left[(A; 0.5|B; 0.5); 0.5|(C; 0.5|D; 0.5); 0.5\right]; 0.9|B_u; 0.1\right)$$

and

$$L_b = \left(\left[(E; 0.5|F; 0.5); 0.5|(G; 0.5|H; 0.5); 0.5\right]; 0.9|B_d; 0.1\right).$$

In the squared brackets are the two-stage lotteries, if the household survives, and on the right hand side is the bequest of the current period, if the household does not survive - with $A, B, C, D, E, F, G, H, B_u, B_d \in \Re$, i.e. these are real numbers and no further lotteries. Substituting the simple lotteries $(A; 0.5|B; 0.5)$ by L_{AB} etc., we get

$$L_a = \left(\left[L_{AB}; 0.5|L_{CD}; 0.5\right]; 0.9|B_u; 0.1\right)$$

and

$$L_b = \left(\left[L_{EF}; 0.5|L_{GH}; 0.5\right]; 0.9|B_d; 0.1\right).$$

Assume that the certainty equivalent[45] of L_{AB}, $CE(L_{AB})$, is bigger than that of L_{CD}, and also $CE(L_{EF}) > CE(L_{GH})$. This assumption relies on the fact that the first lottery in the squared brackets is the one after nature has chosen that the decision maker is in the high income class, i.e. the household has more disposable income before the transitory shocks are resolved. Therefore, this lottery yields a higher value than the second lottery, as everything else is equal. Now, we propose a definition of stochastic dominance for the three-stage lotteries arising in the model.[46]

Definition: The three-stage lottery L_a stochastically dominates the three-stage lottery L_b if and only if:

C-1: L_{AB} first-order-stochastically dominates L_{EF},

[45] Received from the evaluation by Expected Utility Theory, Rank-Dependent Utility or Cumulative Prospect Theory.

[46] Actually, the three-stage lotteries consist of more outcome-probability pairs in the second and third stage. However, the principle used here for the simple setting is also applicable for these "bigger" three-stage lotteries.

C-2: L_{CD} first-order-stochastically dominates L_{GH},

C-3: B_u first-order-stochastically dominates B_d,

and L_a strictly stochastically dominates L_b, if at least one of these conditions satisfies strict First-Order-Stochastic Dominance.

Conditions C-1 and C-2 consider the lotteries in squared brackets. Actually, this is a comparison by Two-Stage-Stochastic Dominance, as defined above. The third condition is solely a comparison of real values from the bequest - with the possible relations $B_u>B_d$, $B_u=B_d$, or $B_u<B_d$. The proof that Rank-Dependent Utility[47] is compatible with this concept is shown in Appendix E. Note, this definition solely corresponds to the kind of three-stage lotteries occurring in the model described in Section 2.

5.2 Strategical Equivalence

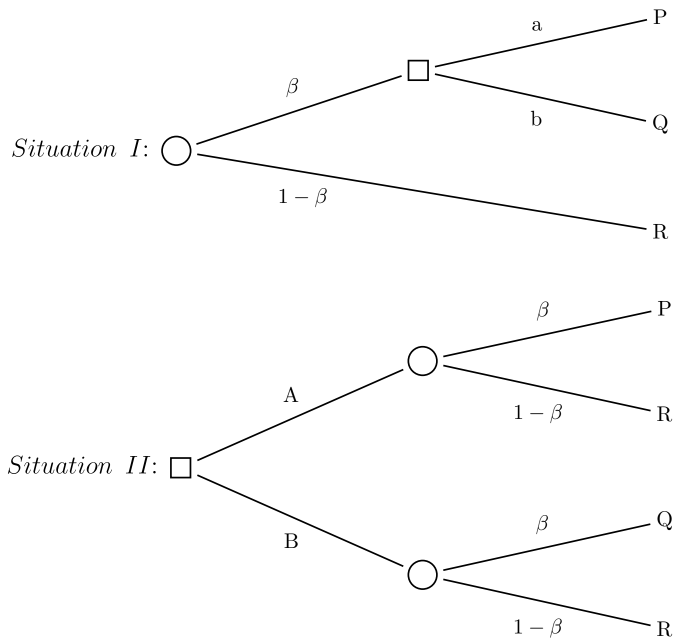
Further points of criticism for Folding Back with Non-Expected Utility preferences - in general - are summarized and given by Machina (1989). The point is that such a procedure implies some "undesirable" properties of behavior - at least "undesirable" from the perspective of the approach proposed in Machina (1989), which throughout uses the Reduction of Compound Lotteries Axiom.

The first of these properties is that a decision maker using Backward Induction with Non-Expected Utility preferences is not indifferent between strategically equivalent decision trees, which are trees with the same opportunity set. An opportunity set is the set of outcomes of all possible strategies in a dynamic problem - an example is given below. This aspect was also presented by Keeney and Winkler (1985), LaValle and Wapman (1986) and Hammond (1988). Consider the decision trees in Figure 11, where Situation I represents the dynamic problem in extensive form and Situation 2 in normal form.[48] Note, that the trees differ in accordance to the place of the decision node. Both representations share the same opportunity set, namely

[47]The same holds for the version of Cumulative Prospect Theory we use, as the probability weighting function has the same properties and the value function, even though twofold, is monotonically increasing as under Rank-Dependent Utility.

[48]This example is taken from LaValle and Wapman (1986).

Figure 11: Strategically Equivalent Decision Trees



- Situation I: strategy a: $(\beta; P|1-\beta; R)$, strategy b: $(\beta; Q|1-\beta; R)$
- Situation II: strategy A: $(\beta; P|1-\beta; R)$, strategy B: $(\beta; Q|1-\beta; R)$.

Due to this aspect, a decision maker should be indifferent between the two trees in Figure 11. LaValle and Wapman (1986) use the Chew-MacCrimmon utility function[49] and some specific lotteries for P, Q and R, to show that in the absence of the Independence Axiom (see definition below) Rolling Back a decision tree is not an appropriate approach. This conclusion refers to the aspect that the two representations of Figure 11 can lead to different strategies under the standard assumptions. However, this argument holds for Rank-Dependent Utility and Cumulative Prospect Theory only if the Reduction of Compound Lotteries Axiom is used, as the usage of this axiom results in the non-adherence of the Independence Axiom, which is defined as

> *Independence Axiom*: Let $X = (x_1; p|x_2; (1-p))$ and $Y = (y_1; q|y_2; (1-q))$ be simple one-stage lotteries. The lottery X is preferred (indifferent) to lottery Y, if and only if $(X; r|Z; (1-r))$ is preferred (indifferent)

[49] This is another Non-Expected Utility model proposed by Chew and MacCrimmon (1979).

to $(Y; r|Z; (1-r))$ for all lotteries Z and probabilities r, i.e.

$$X \succ (\sim) Y \Leftrightarrow (X; r|Z; (1-r)) \succ (\sim)(Y; r|Z; (1-r)).$$

Appendix F contains the proof that the Independence Axiom is satisfied by Rank-Dependent Utility[50] preferences, at least for the constellation used in the definition above, i.e. for lotteries having two lotteries as outcomes. The central aspect of the proof is the usage of the Folding Back procedure instead of the Reduction of Compound Lotteries Axiom.

A remaining problem is that a decision maker with Rank-Dependent Utility or Cumulative Prospect Theory preferences is not indifferent between the extensive and normal form, although these share the same opportunity set. The result of this indifference are different lottery values for the extensive and normal form. However, this has only implications if different forms of representation are used in subtrees at the same level. In the dynamic setting of our model all subtrees are represented in extensive form. This simply stems from the construction of the risky components over time, i.e. hazard of survival, persistent and transitory shocks.

There is also another point of view in the literature put forward by Hazan (1987), who directly comments on the article of Lavalle and Wapman (1986). He states that the usage of extensive form is mostly perfectly proper in the absence of the Independence Axiom (in his argumentation the Independence Axiom is not fulfilled, as the Reduction of Compound Lotteries Axiom is applied), but the transformation from extensive form to normal form is impermissible. The argument of Hazan (1987) is based on the fact that Non-Expected Utility preferences, like the Chew-MacCrimmon functional used by Lavalle and Wapman (1986), lead to dynamically inconsistent strategies, if the Reduction of Compound Lotteries Axiom and the transformation from extensive to normal from is used.

[50] Again, the proof also applies for the version of Cumulative Prospect Theory used in this thesis, because the probability weighting function and the value function feature in principle the same properties as those of Rank-Dependent Utility.

5.3 Aversion to Information

Another objection listed by Machina (1989) is that Non-Expected Utility maximizers exhibit Aversion to Information. Whenever there prevail preferences of the form X $\succ$ Y, i.e. lottery X is preferred to Y, and (Y if E; Z if $\neg$E) $\succ$ (X if E; Z if $\neg$E), i.e. the lottery, which provides lottery Y if event E occurs and lottery Z if event $\neg$ E (that is "not" E) occurs, is preferred. A decision maker with these kind of preferences would choose (Y if E; Z if $\neg$E) without any further information about the uncertain event E. However, if the decision maker could learn whether E will occur before choosing, then she/he would realize that she/he would decide to get X if E will occur. This leads to the conclusion that the decision maker would choose (X if E; Z if $\neg$E), which is the less preferred (compound) lottery in accordance to the preferences assumed above. Hence, the decision maker is averse to information, as the information would make him worse off. This argument was also found by Keasey (1984), Loomes and Sugden (1984), Wakker (1988) and Hilton (1989). But as in the subsection above, this argument holds because the assumed preferences violate the Independence Axiom. As noted above the Independence Axiom is not violated by Rank-Dependent Utility and Cumulative Prospect Theory, if the Reduction of Compound Lotteries Axiom is not used. This implies that this kind of preferences do not prevail under these theories and therefore the Aversion of Information critique does not apply.

5.4 Folding Back as a Formal Optimization Tool

The last critique point by Machina (1989, pp.1655-56), and according to him a more fundamental one, is that:

> *"...as a formal optimization tool, folding back is appropriate only when the objective function is separable across the various subdecisions of a problem, and this is simply not true for an individual with general non-separable (that is, general non-expected utility) preferences who is facing a dynamic choice situation."*

In order to see why Non-Expected Utility preferences are considered as generally non-separable in Machina (1989), regard the following two examples. The first example is a problem set showing again the Allais Paradox, as the one described in Section 2.

Problem 8a:

a_1: ($1 000 000; 1)

a_2: ($ 5 000 000 ; 0.1 | $ 1 000 000 ; 0.89 | $ 0 ; 0.01)

Problem 8b:

a_3: ($ 5 000 000 ; 0.1 | $ 0 ; 0.9)

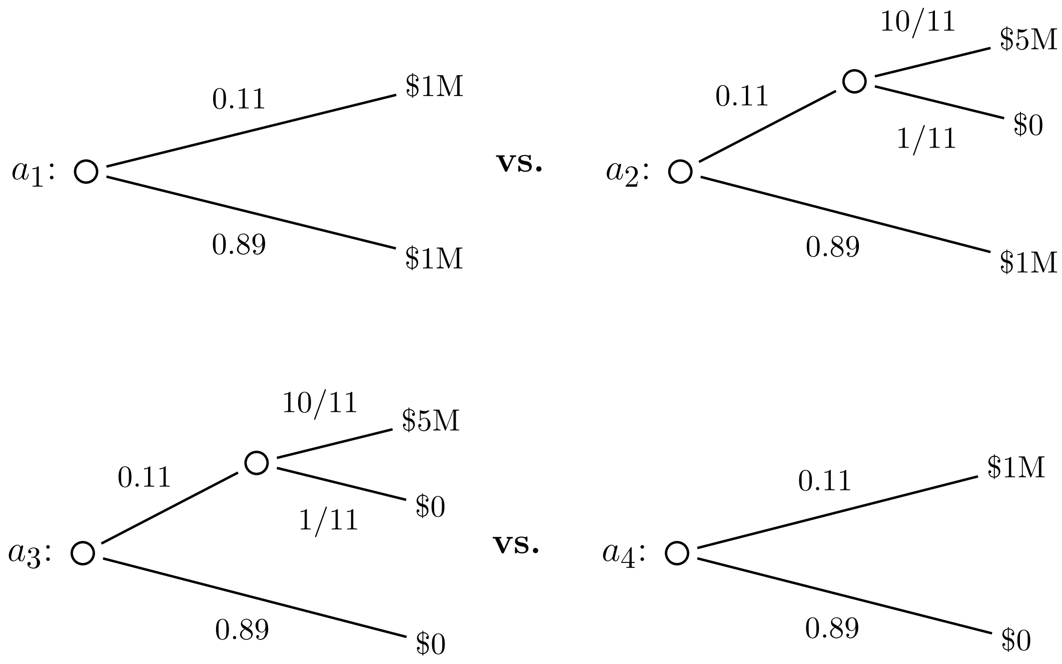
a_4: ($ 1 000 000 ; 0.11 | $ 0 ; 0.89)

This example violates Expected Utility Theory, as the majority of the subjects in these kind of choice problems prefer a_1 over a_2, and a_3 over a_4. Recall, this phenomenon was called the Common Consequence Effect in Subsection 3.1. For illustration the Reduction of Compound Lotteries Axiom is used to show why these choices violate Expected Utility Theory - consider Figure 12, where $1M is $ 1 000 000.[51]

Note, the only difference between the upper and the lower trees is that the lower branch (from the roots of the trees) is receiving $1M with a chance of 0.89 for a_1 and a_2, and $0 also with a chance of 0.89 for a_3 and a_4. The upper tree pair indicates, as a_1 was preferred to a_2, that an individual with such preferences would be willing to replace the upper sublottery ($5M; 10/11 | $ 0; 1/11) by a sure gain of $1M. However, the lower tree pair with a_3 preferred to a_4 shows, that such a change would not be accepted. This is a violation of a principle called Replacement Separability over Sublotteries and is the first aspect why Non-Expected Utility preferences are considered as non-separable in Machina (1989). This conclusion is drawn because Non-Expected Utility models are capable to capture the choices of individuals in such Allais-type problems. In order to see this, consider the evaluation of Problems 8a and 8b with Rank-Dependent Utility with the utility function $u(x) = ln(x+1)$ and the probability

[51] This figure is from Machina (1989).

Figure 12: Example for Replacement Separability



weighting function $\pi(p) = exp(-(-ln(p))^{0.5})$.

$$V_{a_1} = ln(1M + 1) = 13.82$$
$$V_{a_2} = (\pi(0.99) - \pi(0.1)) * ln(1M + 1) + \pi(0.1) * ln(5M + 1) = 12.85$$

and

$$V_{a_3} = \pi(0.1) * ln(5M + 1) = 3.38$$
$$V_{a_4} = \pi(0.11) * ln(1M + 1) = 3.13$$

Hence, Rank-Dependent Utility hits the preferences of the Allais Paradox example, as $V_{a_1} > V_{a_2}$ and $V_{a_3} > V_{a_4}$. But, not all combinations of utility and probability weighting functions imply this kind of preferences, e.g. the combination $u(x) = x$ and $\pi(p)$ does not.

Now, let us have a look at the lotteries displayed in Figure 12, which were generated by application of the Reduction of Compound Lotteries Axiom.[52] For the evaluation we use Folding Back with Rank-Dependent Utility. For the

[52]Here, the axiom is used to extend a one-stage lottery to a two-stage lottery.

lotteries a_1 and a_4 nothing changes, but for a_2 and a_3 the value V′ of sublottery ($5M;10/11 | $0;1/11) has to be calculated first, i.e.

$$V' = \pi(10/11) * ln(5M + 1) = 11.33.$$

Accordingly, the choice problems a_2 and a_3 are now determined by the lotteries[53] (11.33;0.11 | ln(1M+1);0.89) and (11.33;0.11 | ln(0+1);0.89), respectively. The values of the lotteries are[54]

$$\begin{aligned}
V_{a_1} &= 13.82 \\
V^F_{a_2} &= (1 - \pi(0.89)) * 11.33 + \pi(0.89) * ln(1M + 1) = 13.10 \\
&\text{and} \\
V^F_{a_3} &= \pi(0.11) * 11.33 = 2.56 \\
V_{a_4} &= 3.13.
\end{aligned}$$

In contrast to above, the choice pattern has changed - a_1 is still preferred to a_2, but a_4 is now chosen over a_3. This implies that Replacement Separability over Sublotteries is no longer violated in this example. Furthermore, as shown above the Independence Axiom (without the usage of the Reduction of Compound Lotteries Axiom) is satisfied in general by Rank-Dependent Utility. Therefore, Replacement Separability over Sublotteries is not violated, as the Independence Axiom implies separability. Thus, the first aspect in Machina (1989) can be rejected for Rank-Dependent Utility and Cumulative Prospect Theory, at least for the type used in this thesis.

The second example, which is used in Machina (1989) to show the non-separability of Non-Expected Utility, is the same as the one presented in Subsection 3.1 by Problems 3 and 4.

Problem 3: N=95

A: (4000;0.80 | 0;0.20) [20]

[53]Note, the value V′ is already in terms of utility, therefore the outcomes have to be transformed before the evaluation proceeds.

[54]Superscript F stands for Folding Back.

B: (3000;1) [80*]

Problem 4: N=95

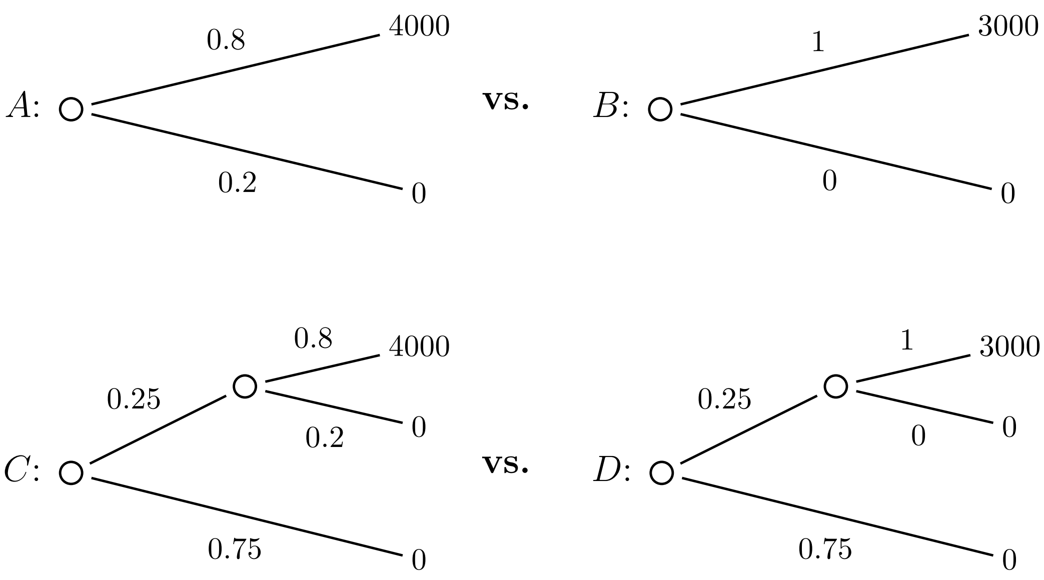
C: (4000;0.20 | 0;0.80) [65*]

D: (3000;0.25 | 0;0.75) [35]

The majority selected B in Problem 3, and most people chose lottery C in Problem 4. This phenomenon was called the Common Ratio Effect, as the second lottery pair is equal to the first pair multiplied by 0.25. Figure 13 shows that with the Reduction of Compound Lotteries Axiom these two problem sets share an identical choice situation.

In Figure 13, the upper sublotteries of C and D are the same as the lotteries

Figure 13: Example for Mixture Separability



given by A and B.[55] The rest of the trees of C and D is identical, thus it should be actual a choice between the two lotteries A and B. However, the results from

[55]This figure is also from Machina (1989).

Kahneman and Tversky (1979) and others indicate that people do not consider these choice patterns as generally equivalent. This is a violation of a principle called Mixture Separability over Sublotteries, which is the second aspect why Machina Non-Expected Utility models are regarded in Machina (1989) as generally non-separable, because they can capture preferences found in these kind of experiments.

At this point, let us again show that Rank-Dependent Utility is compatible with the choices given in the example above. Let $u(x) = x$ and $\pi(p) = exp(-(-ln(p))^{0.5})$, then the values of the lotteries are

$$V_A = \pi(0.8) \cdot 4000 = 2494.1$$
$$V_B = \pi(1) \cdot 3000 = 3000$$
and
$$V_C = \pi(0.2) \cdot 4000 = 1124.9$$
$$V_D = \pi(0.25) \cdot 3000 = 924.2.$$

Thus, according to Rank-Dependent Utility lottery B is chosen over A, and lottery C is preferred to D - these are the same preferences as found in the experiment.

Now, consider the Folding Back evaluation of the lower lottery pair in Figure 13. As stated above, the sublotteries are identical to A and B, therefore the values of these are already calculated by V_A and V_B. This yields the lotteries (2949.1;0.25 | 0;0.75) and (3000;0.25 | 0;0.75) for C and D, respectively. As the ranking is equal in both lotteries, i.e. $V_A > 0$ and $V_B > 0$, and the probabilities are identical it is obvious (as long as u(x) is continuously increasing) that lottery D will be chosen over C. This choice pattern is actually the same, as required by the Independence Axiom, which is generally satisfied by Rank-Dependent Utility. Hence, the critique that Mixture Separability over Sublotteries is violated and the conclusion that Rank-Dependent Utility, as well as Cumulative Prospect Theory, are non-separable is not justified, at least from the perspective of the procedure proposed in this thesis.

6 Conclusion Theoretical Part

The previous sections suggest that the model proposed in Section 2 should be solved via Folding Back without the application of the Reduction of Compound Lotteries Axiom, if preferences according to Rank-Dependent Utility or Cumulative Prospect Theory prevail. The most driving aspect of this suggestion is that this procedure ensures that the optimal or chosen strategy in a dynamic setting is dynamically consistent, i.e. this strategy is obeyed over the various stages of the dynamic problem. This is especially of interest in our consumption/savings model, as each year another Self of the household has to decide how to proceed. This implies that in each period, a household has sufficient time to reconsider the plan or strategy it had made earlier in life. The model was due to this aspect described as an intra-personal game with as many players as years in the life cycle setting. The Folding Back procedure guarantees that each player behaves optimal according to the decisions of the other players, which is exactly the definition of the equilibrium of the model.

Furthermore, a decision maker applying Backwards Induction without the Reduction of Compound Lotteries Axiom is a consequentialist, i.e. only relevant consequences influence his or her evaluation process. This implies for the model that each Self t solely considers the decisions of the players acting after period t. The decisions of the previous players only influence the situation, determined by the amount of the liquid and illiquid assets, Self t is in, which is the starting point of its evaluation process.

The central conclusion is that the Reduction of Compound Lotteries Axiom is not used to solve the dynamics of the model. This is due to the aspect that this axiom is the basis of the critique given in the literature that Non-Expected Utility preferences are generally non-separable, lead to dynamically inconsistent strategies and Aversion to Information. These points of criticism, as discussed in Section 5, can be overcome if Folding Back is used instead, at least for Rank-Dependent Utility and Cumulative Prospect Theory. This is mainly due to the fact that under the procedure, proposed in this thesis, the Independence Axiom is satisfied for these two theories. The aspect that Rank-Dependent Utility and Cumulative Prospect Theory preferences imply that strategically equivalent decision trees (normal form and extensive form) are not considered as indifferent can not be solved by the Backwards Induction routine. But, this does not influence the solution process of the model, as the dynamic structure is throughout in extensive form.

Most convincing, apart from the theoretical aspects, for the non-adherence of the Reduction of Compound Lotteries, is the empirical evidence that individuals do not obey this axiom, at least the majority of the participants in the various experiments. This is a crucial factor (at least in our opinion) as the introduction of the Non-Expected Utility Theories is based and justified by empirical evidence, indicating that human behavior according to choices under risk violate basic principles of Expected Utility Theory.

However, at this point it has to be mentioned that Expected Utility Theory is a normative theory and the alternative theories are descriptive. We analyze if consumption/savings decisions over the life cycle can be better explained by Rank-Dependent Utility or Cumulative Prospect Theory than by Expected Utility Theory. This implies that on the one hand theories are used in the model that are based on empirical findings and on the other hand a normative theory suggesting how people should behave based on logical grounds. It is exactly this comparison, which lies at the heart of the upcoming part.

As a final remark for the theoretical part, note that a special application property of the Reduction of Compound Lotteries Axiom is that most dynamic models can be solved analytically and thereby enabling the researchers to examine different scenarios in their models. However, the model proposed in Section 2 can only be solved numerically, because of the ranking procedure in Rank-Dependent Utility and Cumulative Prospect Theory. Therefore, the only advantage remaining of this axiom for application would be reducing the stages of risk between the various decisions. But, as discussed above this leads to some undesirable properties.

7 Data Description

For the estimations we use the SOEP dataset of the years 2000 to 2008. Only for the estimation of the empirical moments we take the data from the SAVE study 2007, because the SOEP survey does not contain enough detailed information about the wealth of households. In order to achieve a comparable basis between these two datasets, the weights provided by each survey are applied in the estimation procedures. These weights are, in both surveys, based on the German Micro Zensus, so that the datasets combined with the corresponding weights provide a representative sample of the German population.[56] The base year for the monetary variables in the datasets is 2008, i.e. the earlier years are transformed according to the yearly inflation.[57] For all estimations we use information on the household level. For those variables, which are not available on the household level, person level data is used to calculate the corresponding household information, e.g. for highest educational level in the household.[58] We use only German[59] households with a household head, a spouse, with kids or no kids and with dependent adults or no dependent adults. Financially dependent persons in the household are considered as kids if they are younger than 17, otherwise they are considered as dependent adults. Furthermore, we separate the datasets into three different educational groups based on the CASMIN-Classification, see Table C.2 in the Appendix and Brauns and Steinmann (1999). In the following we concentrate only on the lowest educational group, as this offers the most observations.

8 Estimation Procedure

In this section we will briefly describe the procedure of the Method of Simulated Moments estimation, which is illustrated in Figure 14. A detailed description of each step will be given in the following sections. The aim of the Method of Simulated Moments is to achieve estimates of the parameters α and r_{ref} - see bottom of Figure 14. These are obtained by minimization of an objective function, which measures the difference between empirical and simulated values. The empirical moments (left branch in Figure 14) in the objective function are simply obtained from the SAVE dataset, i.e. we estimate wealth to income ratios of 4 age cohorts. The simulated input of the objective function is a bit more involved (right branch in Figure 14). We start by estimating the

[56] For details of the calibration of the weights see Frick and Haisken-DeNew (2005) for SOEP and Börsch-Supan et al. (2009) for SAVE.

[57] The inflation values are taken from the Statistisches Bundesamt Deutschland.

[58] We determine that the highest educational level is given by the highest education of the household head or the spouse. We define it that way, as we assume that these two are responsible for the financial decisions in the household.

[59] A household is considered as German if the household head has German citizenship.

income, the number of kids and the number of dependent adults for the ages 20 to 90 from the SOEP dataset.

The results of these estimations and reasonable values for other parameters, like e.g. the interest rate or coefficient of relative risk aversion, are used to calibrate the model of Section 2. Then, the model is solved via Backwards Induction. The setup is designed in such a way that all possibilities, which can occur in the simulation are taken into account. From the solution of the model we get a decision matrix, which contains the optimal strategies at any age, i.e. 20 to 90, and any possible state a household could be in in the model.

Independent of that, we use the results of the income estimation to simulate 5000 individual income streams for the age 20 to 90. The difference of the income streams stems from the possible income shocks, which can occur each period. In particular, in the model we allow for 3 persistent and 5 transitory shocks. Each period and for each household there is a combination of these two kind of shocks.

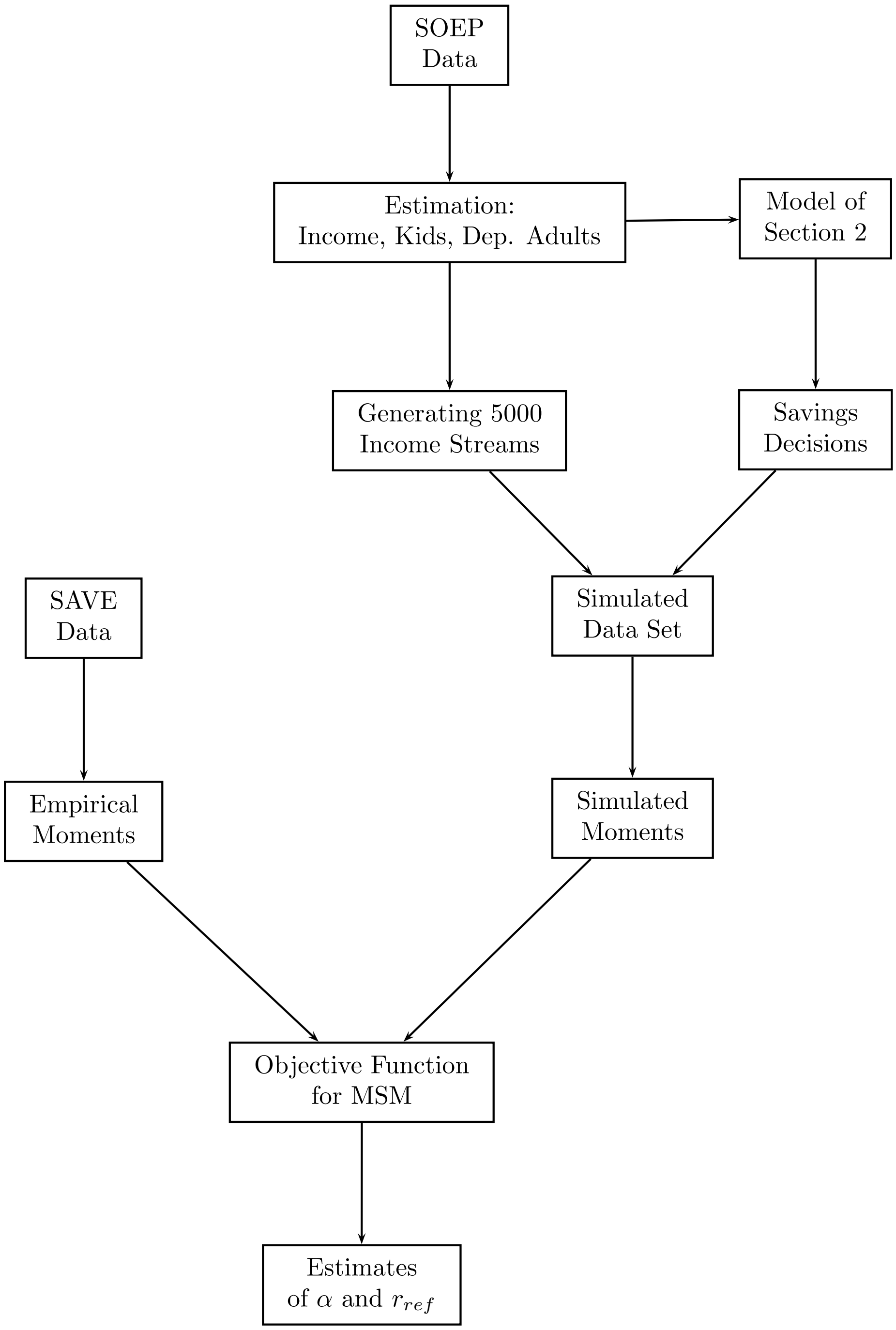
Then, the simulated income profile of each household is considered, in order to get the information which income shocks occurred in each period. The income shocks and the holdings of liquid and illiquid assets, which depend on the decisions of previous periods, determine the state in a certain period a household is in.

The decisions in the decisions matrix, which we obtained from the solution of our model are stored in the same way, i.e. by income shocks, asset holdings and age. Thus, we can look up the optimal decisions for each household.

This procedure yields a simulated dataset with 5000 observations over the ages 20 to 90. Then, this dataset is used to calculate the simulated moments, i.e. the wealth to income ratios of 4 age-cohorts.

Now, we can put the empirical and simulated moments into the objective function of the Method of Simulated Moments and estimate the parameters α and r_{ref}.

Figure 14: Estimation Procedure



9 Calibration

This and the next section deals with the calibration and estimation of necessary parameters for the simulation process. In the literature this is sometimes also called the first stage of the Method of Simulated Moments approach.

9.1 Parameters

Table 4 summarizes the parameters for the base case estimation. The values for the coefficient of relative risk aversion, the consumption flow, the credit limit, the discount rate and the interest rate for the illiquid assets are taken from Laibson et al. (2007). The interest rate for positive liquid assets is the average of the "Basiszins" of the Deutsche Bundesbank of the years 2000 to 2008. The retirement age T for the low educational group is the average of the households in the panel sample, which retire in the observed years from 2001 to 2008.

The survival rates are from the Statistisches Bundesamt Deutschland[60] and are displayed in Figure 15. We took the values for men[61] to proxy the survival rates for households. Recall, in our model we defined a household to have at least a head and a spouse, i.e. if one of these two dies then according to the definition in the model the household does not exist any longer, so to speak. Also the drop to zero in Figure 15 at the age of 90 is due to the assumption in the model that a household does not get older than 90 years. In general, the survival rates represent the probability of having reached a certain age and to live on for one year. This probability is near to 1 until it starts decreasing around the age of 60.

Note, this is the base case calibration. Later on we will conduct a robustness check, in order to analyze if changes in the parameter calibration have an influence on the Method of Simulated Moments estimator.

---

[60]See https://www.genesis.destatis.de.

[61]One could also use the data for women. But, we do not think that this would change the results, as there is not so much difference between the rates for women and men, only that women tend to live a bit longer in probabilities.

Figure 15: Survival rates

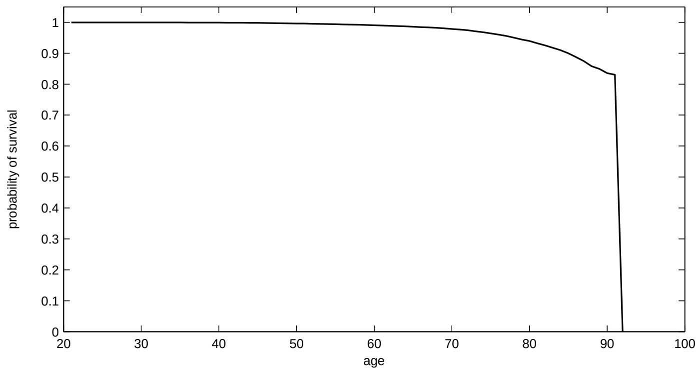


Table 4: Base case parameter calibration

coefficient of relative riskaversion ρ	2	-
consumption flow γ	0.05	-
discount rate δ	0.95	-
credit limit λ	0.318	(0.017)
interest rate positive liquid assets R^X	1.0241	(0.0135)
interest rate negative liquid assets R^X_{CC}	1.1	-
interest rate illiquid assets R^Z	1	-
retirement age T	66	(4.4071)

Available standard errors are given in parentheses.

9.2 Number of Kids and Dependent Adults

For the simulation of the household size we estimate the number of kids and financially dependent adults in a household over time.

The estimation model is selected by backward elimination. This means we start with a quite big model and eliminate all variables, which are not significant at the 5% level, i.e. if the t-value is smaller than 1.96 in absolut terms. This procedure yields the same model for kids and dependent adults. Thus, in Equation 1 x_{it} denotes either the number of kids or dependent adults in the household i in period t. The error process is assumed to be i.i.d.. In the estimation no fixed (or random) effects are included, although it might be reasonable to control for it. The reason why fixed (or random) effects are not considered is simply that it would increase the computation time of the simulation process enormously. In other words the individual effect of each household i is part of the error process in our estimations and is assumed to be zero in the simulation process.

$$x_{it} = \beta_0 * exp(\beta_1 * age_{it} - \beta_2 * age_{it}^2) + \epsilon_{it}^k \tag{1}$$

We estimate the above equation by Weighted Nonlinear Least-Squares, where the weights are given by the population weights from the SOEP survey. The parameter estimates for number of kids and dependent adults are displayed in Table 5. The estimated values are significant at the 5% level. Note that the parameter estimates of β_0 for both are not significantly different from 0, but if we set this factor equal to zero, then the number of kids and dependent adults is equal to 0 over all ages.[62] Due to this attribute of the specification, we use as a null hypothesis value for this parameter the value 1. This implies if the parameter estimate of β_0 is not significantly different from 1, then the next smaller model specification would be just the term of the exponential function. However, the estimates of β_0 are significantly different from 1, as the t-values for kids $t_k = (1.01e^{-5} - 1)/6.34e^{-6} = -1.58e^5$ and for dependent adults $t_a = (2.23e^{-12} - 1)/3.77e^{-12} = -2.65e^{11}$ are in absolute terms bigger than 1.96.

We also tested linear models, but these came up with negative values for young households. A negative number of kids or dependent adults is not

[62] For $\beta = 0$ we get: $x_{it} = 0 * exp(\beta_1 * age_{it} - \beta_2 * age_{it}^2) + \epsilon_{it}^k = 0$.

reasonable for the simulation, thus we dropped these specifications.

Figures 16 and 17 show the predicted profiles over the life cycle for kids and dependent adults in a representative household, respectively. These profiles are later on used in the simulation process for all simulated households, as we set the fixed effects of the individual households equal to 0 due to computation time.

Table 5: Parameter estimates for kids and adults

	kids	adults
β_0	$1.01e^{-5}$	$2.23e^{-12}$
	$(6.34e^{-6})$	$(3.77e^{-12})$
β_1	0.6536	1.076
	(0.032)	(0.068)
β_2	0.009	0.011
	$(4.08e^{-4})$	$(6.78e^{-4})$

The standard deviations are given in parentheses.

The number of kids peaks at the age around 38 and the number of dependent adults around 50. This pattern is partly depending on the definition of kids in the household. Recall, a person in the household is considered as a kid, if it is younger than 17 years. Thus, kids older than 17 who still stay in the household (for some years) are treated as dependent adults in the household.

10 Income Estimation

The estimation of income is separated in the two categories i.) when the household is in the workforce and ii.) when the household is retired. A household is considered to be in the working force if both, head and spouse, are working. On the other hand a household is considered to be in retirement if both, head and spouse, answered in the questionnaire that they are retired. The

Figure 16: Number of kids by age

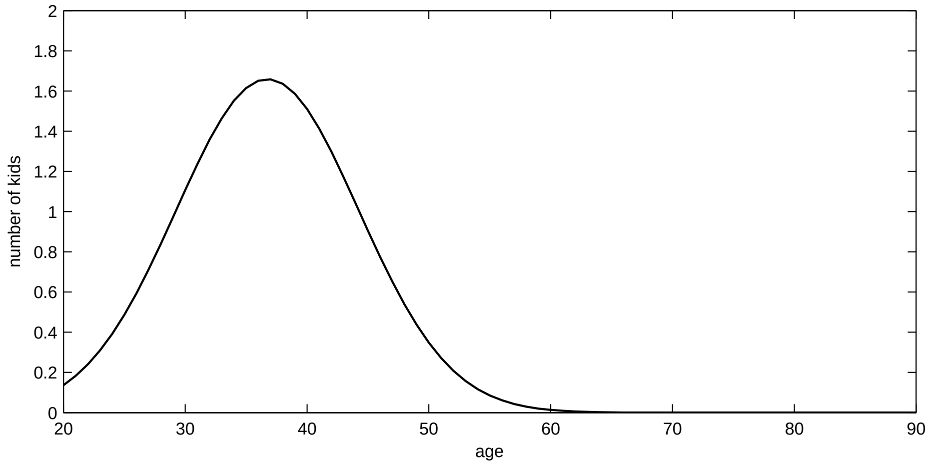
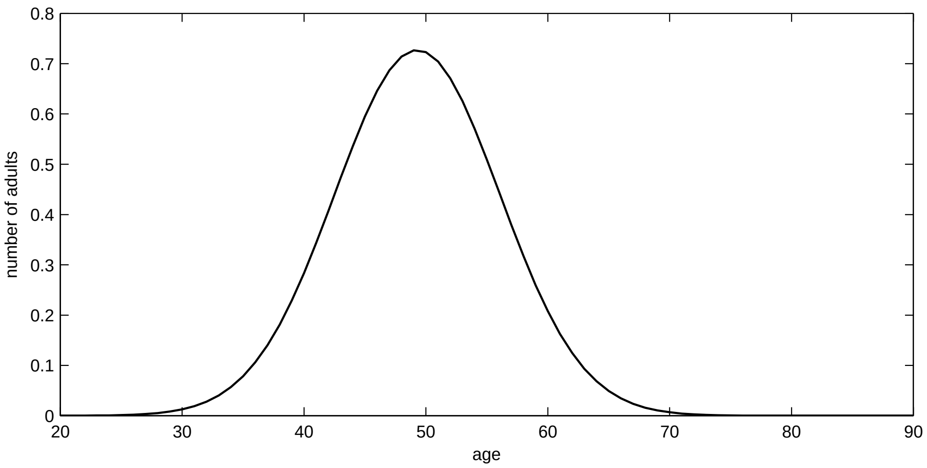


Figure 17: Number of dep. adults by age



households with either the head or the spouse retired were dropped from the sample, but were used to determine the average retirement age of each educational category.

The endogenous variable Y_t includes all regular after tax income from transfers and wages, inheritances, any kind of retirement pensions, and transfers from the government including Social Security. The equation for the estimation of the working households is given by

$$ln(Y_{it}) = y_{it} = \beta_0 + \beta_1 * age_{it} + \beta_2 * (age_{it}^2/100) \\ + \beta_3 * kids_{it} + \beta_4 * dep.adults_{it} + \xi_{it} \quad (2)$$

and is estimated by Weighted Least Squares. The weights are taken from the SOEP data survey. We also executed estimates with different polynomials of age (3. to 4. degree), but these were not significant at the 5% level.[63]

The abbreviations $kids$ and $dep.adults$ stand for number of kids and number of dependent adults in the household, respectively. Furthermore, included in Equation 2 are birth cohorts (10 year intervals) and it is assumed that the time effect is related to the business cycle, which is represented by the unemployment rate. This is done to rule out that time, age and cohort are perfectly correlated, as in Laibson et al. (2003). The results are displayed in Table C.3 in the Appendix.

The error term is defined by a fixed effect η_i, an AR(1) process u_{it} and a transitory shock v_{it}^T, i.e.

$$\xi_{it} = \eta_i + u_{it} + v_{it}^T = \eta_i + a \cdot u_{it-1} + \epsilon_{it} + v_{it}^T.$$

The fixed effect will be ignored in the simulation, like in the case for kids and dependent adults in the household, as it would also increase the computation time enormously. The estimation is only based on the lowest educational category, so some of the individual effect is captured. The crucial parameters for the simulation of the volatility or uncertainty/risk of income are a and σ_ϵ^2, the coefficient and variance of the error of the AR(1) process, and σ_v^2, the variance of the transitory shock. Following Laibson et al. (2007), we estimate these parameters by a weighted General Method of Moments approach. The theoretical moments for the estimation are the first eight autocovariances of $\Delta\xi$, labeled C_k for $k = 1, ..., 8$, and defined by

$$C_k = E(\Delta\xi_t, \Delta\xi_{t-k}).$$

The first differences of ξ are used to circumvent the estimation of η_i, the fixed

[63]This means the t-value, in absolute terms, was bigger than 1.96.

effect which is not used in the simulation. In particular, the first eight autocovariances are given by[64]

$$
\begin{aligned}
C_0 &= \frac{2\sigma_\epsilon^2}{1+\alpha} + 2\sigma_v^2 \\
C_1 &= \frac{-\sigma_\epsilon^2 \cdot (1-\alpha)}{1+\alpha} - \sigma_v^2 \\
&\vdots \\
C_k &= \frac{-\alpha^{k-1}\sigma_\epsilon^2 \cdot (1-\alpha)}{1+\alpha}.
\end{aligned}
$$

The objective of the Generalized Method of Moments is to minimize the following expression

$$
\begin{pmatrix} Cov(\Delta e_t, \Delta e_t) - C_0 \\ Cov(\Delta e_t, \Delta e_{t-1}) - C_1 \\ Cov(\Delta e_t, \Delta e_{t-2}) - C_2 \\ . \\ . \\ . \\ Cov(\Delta e_t, \Delta e_{t-k}) - C_k \end{pmatrix}' * W * \begin{pmatrix} Cov(\Delta e_t, \Delta e_t) - C_0 \\ Cov(\Delta e_t, \Delta e_{t-1}) - C_1 \\ Cov(\Delta e_t, \Delta e_{t-2}) - C_2 \\ . \\ . \\ . \\ Cov(\Delta e_t, \Delta e_{t-k}) - C_k \end{pmatrix}.
$$

where $Cov(\Delta\xi_t, \Delta\xi_{t-k})$ are the empirical autocovariances obtained by the estimation of the income process and W is the optimal weighting matrix, which is given by the variance-covariance matrix of the moment conditions.[65] Intuitively, the objective function measures the difference between the theoretical and empirical moments, which is to be minimized by the parameters a, σ_ϵ^2 and σ_v^2. Let us define the objective function, in short, by

$$
o(a, \sigma_\epsilon^2, \sigma_v^2) = m_{cov}(a, \sigma_\epsilon^2, \sigma_v^2)' \cdot W \cdot m_{cov}(a, \sigma_\epsilon^2, \sigma_v^2).
$$

The results of the GMM estimation are displayed in Table 6. We apply an overidentification test, called J-Test and proposed by Hansen (1982), as there

[64] For a detailed derivation of the autocovariances see Appendix F.

[65] In a two-step Generalized Method of Moments procedure one sets $W = I$ in the first step, as the variance-covariance matrix of the moment conditions is not known. Then, in a second step the estimated variance-covariance matrix of the first step is used as the optimal weighting matrix.

are more moment conditions than parameters to be estimated. The test statistic, given by

$$\psi(\hat{a}, \hat{\sigma_\epsilon^2}, \hat{\sigma_v^2}) = m_{cov}(\hat{a}, \hat{\sigma_\epsilon^2}, \hat{\sigma_v^2})' \cdot W_{opt} \cdot m_{cov}(\hat{a}, \hat{\sigma_\epsilon^2}, \hat{\sigma_v^2}),$$

and is chi-square distributed with $N_m - N_P$ degrees of freedom, if the model is correctly specified, where N_m is the number of moments conditions and N_P the number of estimated parameters.[66]

Table 6: Parameter estimates for income error term

a	0.7425 (0.2981)
σ_ϵ^2	0.023 (0.0081)
σ_v^2	0.0129 (0.0057)

The standard deviations are given in parentheses.

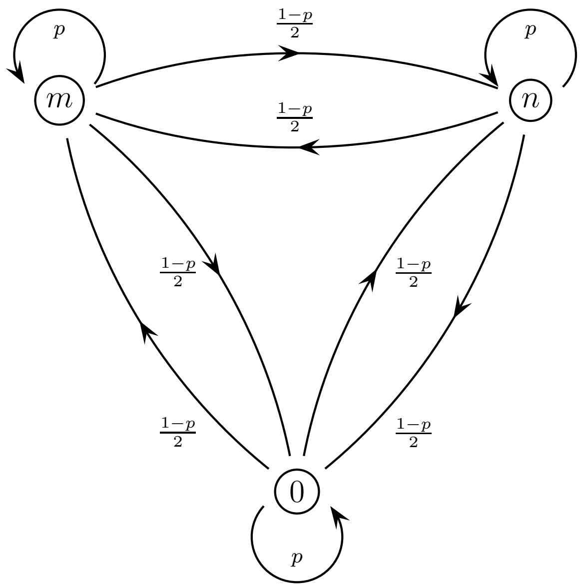
We have 5 degrees of freedom, as we use 8 moment conditions and estimate 3 parameters. The 5 percent critical value of the chi-square distribution is 12.83 and the value of the objective function is 0.0001. Thus, the null hypothesis that the model specification is valid can not be rejected at the 5% level.

Now, the AR(1) component of the income error process is approximated by a 3-state Markov process. This is done to reduce the computation time of the simulation procedure. Actually, an AR(1) process is a continuous Markov process and we represent this by a discrete Markov process with three possible states.

Figure 18 exhibits the 3-state Markov process, which is used in the simu-

[66]Note, this only holds in the case where the optimal weighting matrix W_{opt} is used.

Figure 18: Three state Markov process



lation. We assume a symmetric process, i.e. the two states m and n are equal in magnitude, labeled φ, and the third state is equal to 0. This implies that the 3-state Markov process applied here is defined by two parameters, namely the magnitude φ and the transition probability p. In order to illustrate the process, assume that you are in state m in Figure 18. In this state there are three possibilities

1. staying in state m with probability p,
2. changing to state 0 with probability $\frac{1-p}{2}$ or
3. changing to state n also with probability $\frac{1-p}{2}$.

The same principal holds for the two other states as well. For the approximation of the AR(1) process we generate 5000 time series of the process by using the estimates of a and σ_ϵ^2 from above, i.e. we take draws from N ~$(0, \sigma_\epsilon^2)$, the distribution of the error of the autoregressive process. One time series consists of 71 periods as this is the life cycle span of a household in the model.[67] These draws and a parameter combination of φ and p for the 3-state Markov process

[67]We assume that in the first period the time series starts at 0, i.e. $u_{i1} = 0 + \epsilon_{i1}$.

are used to calculate the difference between the AR(1) and the Markov process. The criterion for the distance is the mean squared error, given by

$$\frac{1}{71}\sum_{t=1}^{71}(x_t - m_t)^2,$$

where x_t is one series of the AR(1) process and m_t is a 3-state Markov process with a particular parameter combination. The same is done for the rest of the 5000 drawn time series of the autoregressive process. The sum of these mean squared errors is the objective function, given by

$$\frac{1}{5000}\sum_{j=1}^{5000}\frac{1}{71}\sum_{t=1}^{71}(x_{jt} - m_{jt})^2. \tag{3}$$

The parameter combination of φ and p which minimizes[68] the objective function, is the combination defining the 3-state Markov process that is used to approximate the AR(1) process of the income error. The minimization of equation 3 is repeated 1000 times, in order to get estimates of the standard deviations of the parameters. The mean of the parameter combinations and its corresponding standard deviations (stdx) are displayed in Table 7.

Table 7: Markov state and transition probability

	φ	p
mean	0.2088	0.7240
stdx	(0.0072)	(0.0239)

The standard deviations are given in parentheses.

Given these estimates the Markov process applied in the simulation is exhibited in Figure 19.

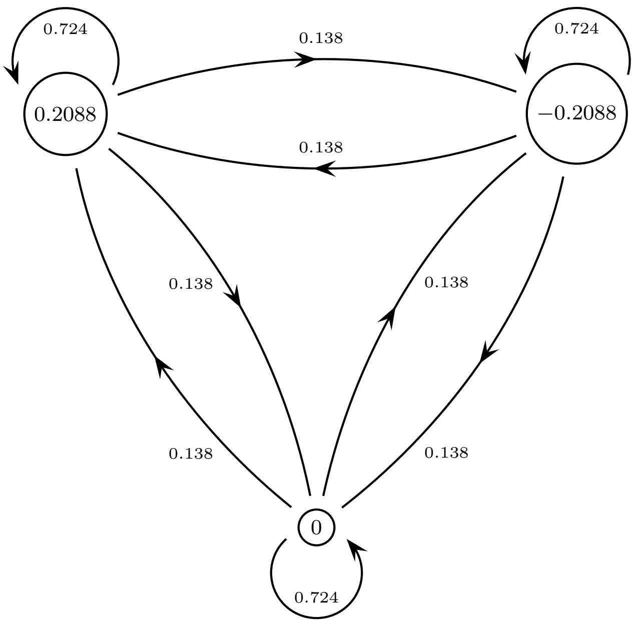
In order to get an impression of how the three state Markov process approximates a draw of the AR(1) process see Figure 20.

We also compared the three state Markov process used here with a two state Markov process[69], as in Laibson et al. (2007). Intuitively any increase of

[68] The minimization is executed by an algorithm - Nelder Mead Simplex standard algorithm within Matlab command "fminsearch".

[69] The two states are given by m and n - the state 0 is simply dropped.

Figure 19: Three state Markov process



the states in the Markov process yields a better approximation of the AR(1) process, as it is itself a continuous Markov process with infinite states. But, as mentioned above it would be to time consuming to use many states in the simulation procedure. Thus, there is a trade-off between a better approximation and the time needed to run the simulation. In order to check whether there is a significant improvement from a two state to a three state Markov process, we compared the mean squared error and its corresponding standard deviation of both processes.

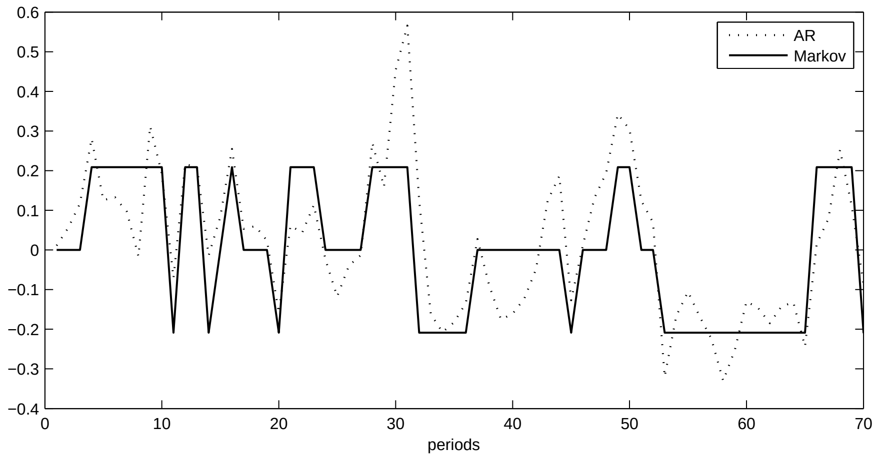
Table 8: Two vs. three state Markov process

	two state	three state
msqe	0.0274	0.0224
	(0.00042)	(0.00037)

The standard deviations are given in parentheses.

The standard deviations of the mean squared errors are estimated by, first, minimizing Equation 3 for both m being a two state and a three state Markov

Figure 20: AR(1) and Markov process

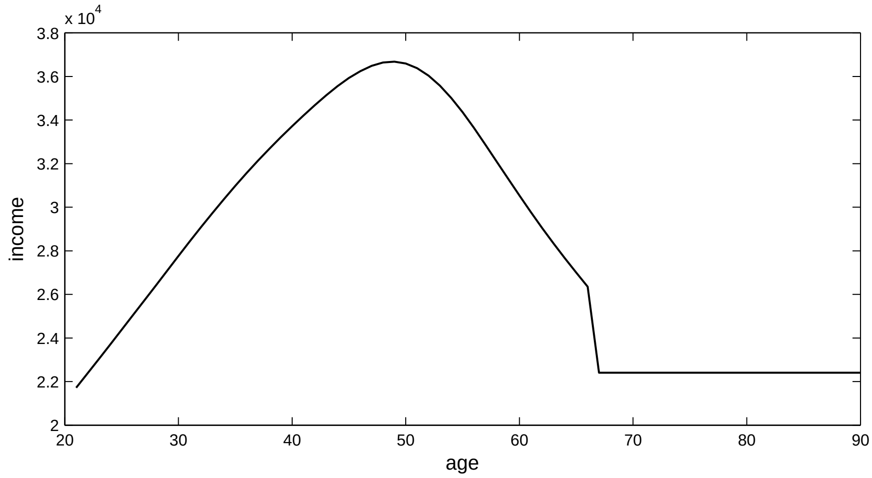


process. The obtained parameters by the minimization are then used to compute the mean squared errors of 100 additional draws of the AR(1) process. These 100 values of mean squared errors are finally used to get estimates of the mean of the mean squared error and the corresponding standard deviations of the two and three state Markov process, displayed in Table 8. The means and the standard deviations indicate that the three state Markov process is significantly better than the two state Markov process. We did not consider a further state, i.e. a four state Markov process, as three states already increased the simulation time enormously. Furthermore, in our opinion it was solely important to implement a state, which has a magnitude of 0, for a better approximation.

The estimation procedure just described is different to the approach of Laibson et al. (2003), who set $\varphi = \sqrt{\frac{\sigma_\epsilon^2}{1-a^2}}$ and $p = \frac{a+1}{2}$, and state that this matches the variance and autocovariance of u_t. However, taking the parameter estimates of a and σ_ϵ they found for the AR(1) process and comparing their calibration with the approach we use, we find that the estimation procedure provides a better fit, at least according to the mean squared error we use as a performance criterion.

Finally, we define the income process of the retired households simply by the average and the corresponding variance of the data, namely in logarithmic terms 10.0170 and 0.0637, respectively. The estimation of the retired income process like e.g. in Laibson et al. (2007) yields very bad results, i.e. the estimates are highly insignificant for all tested specifications. The income process

Figure 21: Average income process over the life cycle



over the life cycle (working and retired combined) for a representative household is displayed in Figure 21. These values are used in the simulation process as the age-specific income of a household. The overall income at a certain age is defined by this age-specific income, the Markov state a household is in and a transitory shock occurring at this age. A more detailed description will be given in the simulation section.

11 Empirical Moments - Wealth to Income Ratio

This section estimates the empirical moments for the Method of Simulated Moments procedure. These moments are the counterpart to the moments, which are generated by the simulation procedure. We use the wealth to income ratios for the age cohorts 20 to 29, 30 to 39, 40 to 49 and 50 to 59. One could chose also other moments, like e.g. Laibson et al. (2007), who use:

1. $\%Visa$ - fraction of households borrowing on any type of credit cards, i.e. paying interest on credit card debts.
2. $meanVisa$ - average outstanding credit card debt in relation to mean income of the age cohorts.
3. CY - marginal propensity of consumption to expected income changes.

4. $wealth$ - wealth to income ratio of the households aged 50 to 59.

The first two of these moments together with the fourth moment are a set to estimate (quasi-) hyperbolic discounting, which is the main objective in Laibson et al. (2007). However, we solely use exponential discounting and do not see a direct connection of the first two moments to our interest. The third one would have been of interest for our research, but the estimation attempts of the marginal propensity to consume with the available data did not come up with any useful results. This means, that nearly all coefficients were highly insignificant. Due to these aspects we defined as our empirical moments the wealth to income ratio of the age cohorts 20 to 29, 30 to 39, 40 to 49 and 50 to 59. We do not use older age cohorts, as in our model the illiquid assets (representing wealth) is constructed in such a way that it can not decrease. In reality, wealth is for example used to maintain a particular standard of living, especially if a household becomes retired.

Intuitively, the wealth to income ratio can be interpreted as the number of years, which wealth can compensate a complete loss of income. Thus, the ratio indicates the intensity of precautionary saving due to uncertain or risky events. As we deal in this thesis especially with the aspect how households act according to risky components, this ratio seems to be a good moment condition for our estimation. Recall, in the probability weighting function high probabilities are underestimated, i.e. that for example the chance to get quite old is underrated. This would imply that the wealth to income ratios should be rather low.

For the estimation of the wealth to income ratios we use cross sectional data from the SAVE study 2007. The data consists of 5 imputed datasets - imputation of datasets is a method to deal with nonresponse in a survey for variables of interest. This is done to overcome the drawback of too few observations, because the simplest way to handle nonresponse is to drop all observations with missing values.[70]

The regression equation for the wealth to income ratio (wr) is determined by

[70] For a description of the imputation method in the SAVE study, see Schunk (2007). The details about the estimation procedure with those 5 imputed datasets is shown in Appendix G.

$$wr = \beta_0 + \beta_1 \cdot kids + \beta_2 \cdot dep.adults + cohorts + \epsilon.$$

We use Weighted Least Squares for the estimation. Recall, the usage of the weights is important, as we have two different datasets in the Method of Simulated Moments approach. The term $cohorts$ represents 5 cohort dummies, with the oldest cohort being the reference group, and ϵ is a normally distributed error term.[71] The results of the (multiple-imputation) estimation are displayed in Table C.4 in the Appendix.

The constant, the cohort reference group[72], and the cohorts coefficients are highly significant. The results for number of kids and dependent adults are only nearly significantly different from 0 at the 10% level.[73] Nevertheless, we keep these variables, as we believe (from an economic point of view) that the number of kids and dependent adults influence the possibilty of a household to accumulate wealth.

These regression results are used to calculate the average wealth to income ratios of the 4 aforementioned age cohorts, which are shown in Table 9. The wealth to income ratios are increasing by age, i.e. wealth could compensate a loss of income over a longer time period the older a household gets. This could mean that the intensity of precautionary savings behavior increases with age.

The results for retired households are not displayed and used as moments here for two reasons. First, we believe that the wealth to income ratio is slowly decreasing after the household stops working, although there is a drop in income. But, wealth is also used to compensate this drop in income, which implies a continuous reduction of wealth over the years in retirement and thereby a decrease in the wealth to income ratio. However, the setup of the model excludes the possibility to use wealth (illiquid assets) for consumption implying that the wealth to income ratio can only decrease if income increases. Second, the estimation results predict that there is a small drop to 7.2851 for the age co-

[71] In contrast to Laibson et al. (2007) we do not use the unemployment rate as a proxy for the business cycle effects, as we only deal with cross sectional data.

[72] The reference group in the estimation is the oldest age cohort, but the results do not change if we change the reference group to the youngest cohort - this is due to the aspect that the dummies only represent relations to each other.

[73] The critical value of the onesided t-test for the 10 % level is 1.2816 - the t-values for number of kids and number of adults are 0.9812/0.7972=1.2309 and 1.5269/1.2649=1.2071.

hort 60 to 69, which is followed by an increase to 9.9312 in the last cohort, age 70 to 90. One can see from the data that this last result is due to some extreme outliers. One could control for this problem by smoothing or dropping the outliers for example. But in combination with the first argument we decided to omit these two age cohorts as empirical moments.

Table 9: Wealth to Income ratio by cohorts

	mean	stdx
age 20 to 29	2.8775	(0.2707)
age 30 to 39	3.4109	(0.2661)
age 40 to 49	4.1555	(0.1246)
age 50 to 59	7.7314	(0.3367)

The standard deviations are given in the stdx column.

The vector containing the empirical moments, i.e. the wealth to income ratios of the 4 age cohorts, is given by

$$m_{j_m} = \begin{pmatrix} 2.8775 \\ 3.4109 \\ 4.1555 \\ 7.7314 \end{pmatrix}.$$

This vector will later on be part of the objective function of the Method of Simulated Moments approach.

12 Simulation

This section deals with the details about the simulation process. First, there will be a description about how the consumption/savings decisions of households are derived. In this part the evaluation of risk is important, i.e. the general utility function and the probability weighting function of Prelec (1998)[74] from subsections 2.6 and 3.2 are used. Recall, it depends on the values of the parameters α and r_{ref} if Expected Utility Theory, Rank-Dependent Utility Theory or Cumulative Prospect Theory prevails. The second part shows the generation of the income streams for 5000 households, which are used for the Method of Simulated Moments estimation.

12.1 Simulation of Household Decisions

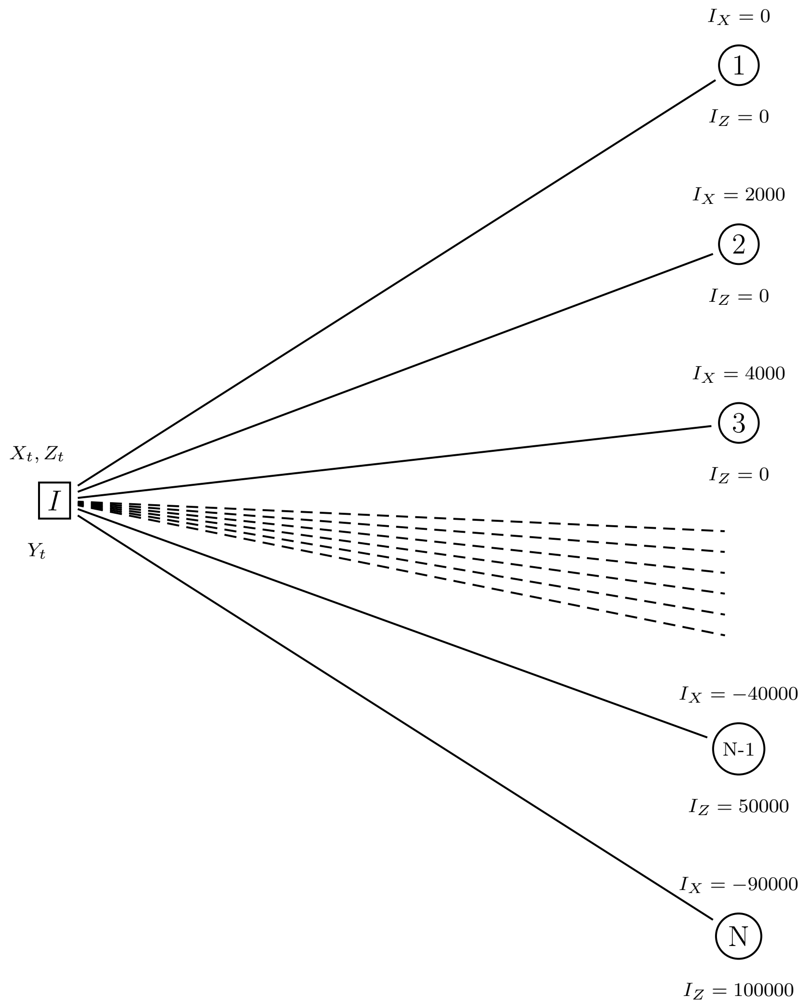
The setup for the simulation of the decisions of the households is constructed in such a way that all possible events and possible decisions that can occur or be made in the model are considered. This is done in order to circumvent that the decisions process has to be calculated for all 5000 households separately. Actually, the decisions are only calculated once in each iteration of the algorithmic search procedure. This means the decisions are recomputed in each run with new parameter values for α and r_{ref} until the objective function converges to a minimum. This will be explained in more detail later on.

Now, let us consider the setup of the model for the decisions, which can be imagined as a really big decision tree. This tree is too huge to display it entirely. Thus, we look just at parts of it, which will repeatedly occur in the whole tree - at least in the same pattern. Figure 22 shows what kind of decisions a household can make in each period t. It is assumed that a household has to make two decisions in each period simultaneously, that is how much money to invest in the liquid asset and how much in the illiquid asset. Recall, these two decisions are sufficient to determine the consumption of the current period - see static budget constraint in Section 2.

A household has to perceive in period t the state it is in, which is determined by the disposable income Y_t, the liquid asset holdings in t, X_t, and the illiquid asset holdings in t, Z_t. In Figure 22 this is summarized at rectangle I. In the model we set the maximum amount of liquid assets and illiquid assets equal to €400 000 and €1 000 000, respectively. These upper limits were chosen to restrict the possibilities that have to be computed. However, these values are

[74] We set $\beta = 1$ in the function of Prelec (1998), in order to reduce the number of parameters to be estimated.

Figure 22: Investment decisions each period



also not exceeded if one uses higher limits.

Recall, in decision trees the notation is that at each rectangle a decision has to be made and at each circle nature is moving, i.e. risk is resolved. This implies in the figure that after the decision is made a move of nature follows, which we will consider in a next step. In Figure 22 at point I a household has several options for its investment decisions, e.g. taking the branch leading to point N would mean that a household would withdraw €90 000 from the liquid asset holdings and invest €100 000 in the illiquid asset. Thus, there must be also a contribution of €10 000 from the disposable income to reach the €100 000 investment in the illiquid asset. This implies that for this option consumption would be equal to $Y_t - I_X - I_Z = Y_t - 10000$.

From Figure 22 one can see that the stepsize for the investments is €2 000 for the liquid and €50 000 for the illiquid asset. These relative high stepsizes are due to the aspect that the computation of the decisions in each iteration is not too time consuming . The high stepsize of the illiquid asset implies that the liquid asset has to be used as an intermediate step if a household wants to put money in the illiquid holdings. This is caused by the fact that the income (no matter what age) is not greater than €50 000. Note, we are dealing with disposable income, i.e. after taxes and other regular payments, and it is also assumed that at least €6 000 are used for basic consumption each year, including especially expenses for food, clothes etc..

Finally, a household has to make its decisions in each period, which depend on its state and the future decisions and possible moves of nature - this was explained in more detail in the example in Section 4. This implies that a household is in one of the points 1 to N, after choosing its investments. Then, nature is moving, i.e. in particular first the hazard of survival, then the risk of the Markov process and finally the risk of the transitory shock resolve. Assume that a household moved for example to point N in Figure 22. Actually, it does make no difference which point one takes, as the risky components are all the same at the same stage, i.e. in the same period. The resulting event tree of the risky elements is displayed in Figure 23. The timing of the occurrence of the risk in each stage[75] is assumed to be:

1. The risk of survival is the last event in each period - either the household lives on to the next period or not.

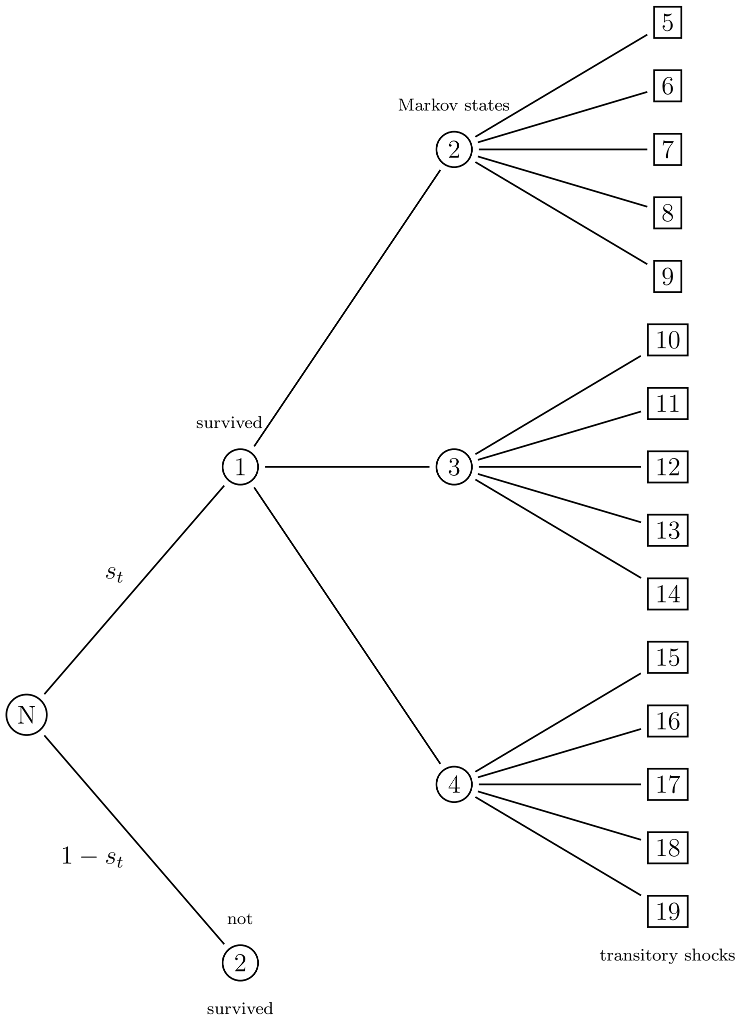
2. The very first thing happening at the beginning of each period is the resolution of the Markov state.

3. After that the transitory shock arises.

In order to give an impression of the model let us describe the whole decision tree. At the root of the tree there are three starting states (decision nodes), which share the same asset holdings, namely $X_0 = 0$ and $Z_0 = 0$, and have the same income but differ in the Markov state. The different Markov states can be interpreted as being in the

[75]Note, "stage" is not equal to "period" in our setting.

Figure 23: Risky elements in one stage



1. low[76] income group, i.e. Markov state 1

2. middle income group, i.e. Markov state 2

3. high income group, i.e. Markov state 3.

Then, the various investment decisions follow, as shown in Figure 22. After that the eventtree of the risky elements, Figure 23, is connected to the tree at each final node of Figure 22. The final nodes of all the event trees in the

[76] Note, a households does not stay in the group to which it initially belonged. There are probabilities each period that it moves to one of the other groups - see the Markov process in Figure 19.

combined tree are the docking points for the decision trees. This procedure continues until we have 70 stages, which represent the age from 20 to 90 and the corresponding decisions in each period. As described in Section 4 this tree is solved via Folding Back. The optimal decisions at each decision node in each period are stored. The vector of the state variables, which defines the state of a household at the beginning of each period, before anything happens, is given by

$$\Lambda_t = \left\{t, X_t, Z_t, \bar{Y}_t, m_{t-1}\right\}, \tag{4}$$

where t is the period, X_t and Z_t are one of the various combinations of liquid and illiquid assets held at the beginning of period t, $\bar{Y}_t$ is the age specific income and m_{t-1} is the Markov state a household was in in the previous period.

These states are followed by the resolution of the Markov state and a transitory shock in period t. Thus, the decision nodes in each period t, when the investment decisions concerning the liquid and illiquid assets are made, are uniquely defined by the following vector

$$\Lambda_t^d = \left\{t, X_t, Z_t, \bar{Y}_t, m_t, v_t\right\}. \tag{5}$$

Here, the vector is expanded by the Markov state and the transitory shock of period t - the Markov state of the previous period does no longer matter. After the realization of this state a household can make its investment decisions. At each of these nodes Λ_t^d the optimal decision is stored while Folding Back.
In order to give a rough impression about the number of states in the tree - t runs from 1 to 70, X from 1 to 200, Z from 1 to 20 and m from 1 to 3 (this loop runs two times, for the current and the previous Markov state) and v from 1 to 5. This yields altogether 12.6 million states where a decision has to be made. At each of these points there are up to 4 000 decision opportunities. In most cases the number of options is smaller, as the illiquid asset is defined to be not decreasing over time.

For completeness one should mention the resolution of the risk of survival, which is also part of the tree. However, this component has only one branch with subsequent nodes, i.e. it does not blow up the tree as the other elements.

12.2 Income Simulation

We generate, like Laibson et al. (2007), 5 000 independent income streams, i.e. the simulated dataset consists of 5 000 households. For the simulation of the income we use the estimation results and the Markov process from Section 10.

First of all the income in each period is determined by the age-specific income, displayed in Figure 21. The second part of the income is determined by the Markov state in the current period, which is depending on the Markov state of the previous period. This implies that the generation of the Markov states starts at period 0. At the beginning a household has, by assumption, an equal chance, i.e. 33.33 percent, to be in Markov state 1,2 or 3. The initial state of a household is determined by taking a random draw d_r out of the interval $[0; 1]$, if

- $d_r < 1/3$ then a household is in Markov state 1.
- $1/3 \leq d_r < 2/3$ then a household is in Markov state 2.
- $d_r \geq 2/3$ then a household is in Markov state 3.

The Markov states of the following periods are determined by the process defined in Section 10. This is also based on a random draw d_r out of the interval $[0; 1]$, but the categorization is now according to the transition probability $p = 0.724$. There are three different cases - if the previous state is

1. Markov state 1, then if
   - $d_r < 0.724$ then a household stays in state 1.
   - $0.724 \leq d_r < 0.862$ then a household moves to state 2.
   - $d_r \geq 0.862$ then a household moves to state 3.
2. Markov state 2, then if
   - $d_r < 0.138$ then a household moves to state 1.
   - $0.138 \leq d_r < 0.862$ then a household stays in state 2.
   - $d_r \geq 0.862$ then a household moves to state 3.
3. Markov state 3, then if

- $d_r < 0.138$ then a household moves to state 1.
- $0.138 \leq d_r < 0.276$ then a household moves to state 2.
- $d_r \geq 0.276$ then a household stays in state 3.

The random draw procedure for the Markov process is repeated 46 times, because we assumed that there are no persistent shocks in retirement.

Furthermore, there are transitory income shocks in each period. We allow for 5 different income shocks which can occur each year. The magnitude of the shocks is defined in terms of its estimated standard deviation from Section 10, namely $\sigma_v = \sqrt{0.0129} = 0.1136$ for working households and $\sigma_v^r = \sqrt{0.0637} = 0.2524$ for retired ones. In general, the 5 shocks are defined by $-5 \cdot \sigma_v$, $-2 \cdot \sigma_v$, 0, $2 \cdot \sigma_v$ and $5 \cdot \sigma_v$ while in the workforce and for retirement σ_v^r is used instead of σ_r.

The generation of the transitory shocks is based on the normal distribution assumption. The values of the probability density function for the defined deviations from the mean are for all kind of normal distributions the same - see Table 10. The values are interpreted as the relative frequency of occurrence.

Table 10: Values probability density function

$-5 \cdot \sigma_v$	$-2 \cdot \sigma_v$	0	$2 \cdot \sigma_v$	$5 \cdot \sigma_v$
0.039	0.4753	3.5118	0.4753	0.039

The values of the probability density function are now used to derive the probabilities of the transitory shocks to arise in each period. First, the values are normalized by dividing by the sum of all, which is $0.039 + 0.4753 + 3.5228 + 0.4753 + 0.039 = 4.5404$. This leads to the corresponding probabilities of the shocks, which are displayed in Table 11.

Table 11: Probabilities of transitory shocks

$-5 \cdot \sigma_v$	$-2 \cdot \sigma_v$	0	$2 \cdot \sigma_v$	$5 \cdot \sigma_v$
0.0086	0.1047	0.7735	0.1047	0.0086

Now, in a last step, the transitory shock which occurs in each period is identified by drawing a random number out of the interval $[0; 1]$. The intervals which relate the random number to a particular shock are given in Table 12.

Table 12: Intervals for transitory shocks

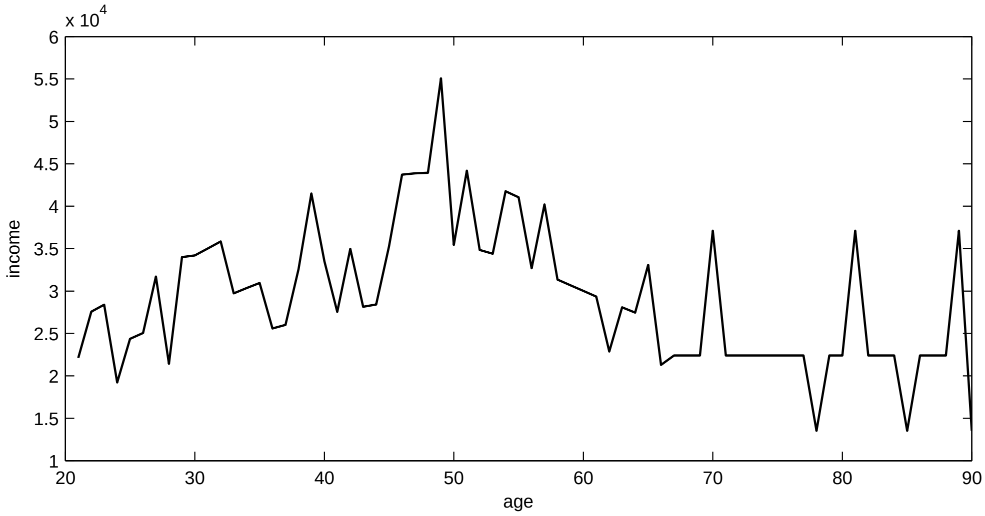
bound	$-5 \cdot \sigma_v$	$-2 \cdot \sigma_v$	0	$2 \cdot \sigma_v$	$5 \cdot \sigma_v$
lower	0	0.0086	0.1133	0.8867	0.9914
upper	0.0086	0.1133	0.8867	0.9914	1

For example assume that the draw of the random number for the transitory shock in period t is 0.9213. This implies, as $0.8867 < 0.9213 \leq 0.9914$, that in period t a positive shock of 2 times the standard deviation occurs. Finally, to give an impression of how a generated income stream of one household looks like, see Figure 24.

Recall, we generate such income streams for 5 000 different households, which form the simulated dataset. For a last illustration of the simulation process of the income, consider the peak in Figure 24 at age 49. Here, this simulated household has a disposable income of €55 062. The components, which are still logarithmized from the income estimation, yielding this outcome are

1. age-specific income $e^{10.4802}$ =€35 604,
2. Markov state $e^{10.4802+0.2088} - e^{10.4802}$ =€8 267,

Figure 24: Example Income Stream



3. transitory shock $e^{10.4802+0.2088+2\cdot 0.1136} - e^{10.4802+0.2088}$ =€11 191.

Note, we do not simulate the survival of the households. We assume that each household lives until the age of 90. The main reason why we proceed that way is to obtain a balanced simulated panel data set. Also, with a large number of simulated observations the inclusion of survival would not change the results for the simulated wealth to income ratios, especially in the way we estimate those. The important aspect is that it is part of the decision process, i.e. the households are aware of the fact that there is a chance to get 90 years old.

12.3 Combination

This subsection deals with the procedure how the two simulation steps, of subsections 11.1 and 11.2, are combined to get the dataset of 5 000 households. For every household we start at period 0. The liquid and illiquid asset holdings are, by assumption, equal to zero at the beginning. The first step is to determine the initial Markov state. Then, at the beginning at period 1 (age 20) the Markov state and the transitory shock of period 1 are simulated, as described in the section above. Thus, we get the first state Λ_1^d where a decision has to be made. In order to get the optimal decision at this point we take the decision of this state, which we saved while the Folding Back procedure of the overall tree. Now, we can calculate the liquid and illiquid asset holdings at the end of period 1. The investment decisions and the disposable income given, we can calculate the consumption of period 1. According to the definition of the timing, the resolution of survival would be next, but as explained above we omit this part and assume that the household exists until the age of 90. Thus, the household survives the first period and the state at the very beginning of period 2 is given by vector Λ_2.

Then, the Markov state and the transitory shock of period 2 are simulated. This leads again to a decision node Λ_2^d. Here, we again look for the optimal decision, which we have saved while Folding Back. This is continued until period 70 is reached, i.e. the age of 90.

We repeat this procedure 5000 times to get our simulated dataset for the determination of the simulated moments for the Method of Simulated Moments approach. The simulated moments are denoted by m_{J_S} and defined as

$$m_{J_S} = \begin{pmatrix} m_{J_S}^1 \\ m_{J_S}^2 \\ m_{J_S}^3 \\ m_{J_S}^4 \end{pmatrix},$$

where $m_{J_S}^r$, for $r = 1, 2, 3, 4$, are the simulated wealth to income ratios of 4 age cohorts from the simulated dataset. The wealth to income ratios are estimated in the same way and for the same age cohorts as the empirical moments in Section 11.

13 MSM Estimation

The Method of Simulated Moments[77] approach estimates the parameters of interest in the model. We follow the approach of Laibson et al. (2007) and Gourinchas and Parker (2002). The parameter vector, which will be estimated is given by

$$\theta = (\alpha, r_{ref}).$$

[77]The Method of Simulated Moments goes back to McFadden (1989), Pakes and Pollard (1989) and Duffie and Singleton (1993). See Stern (1997) for a review of simulation-based estimation techniques.

Recall, the parameter α belongs to the probability weighting function of Rank-Dependent Utility and Cumulative Prospect Theory, respectively. The parameter r_{ref} is a reference point and exhibits the distinction of Rank-Dependent Utility and Cumulative Prospect Theory, at least in the way it is applied here. The null hypothesis in our estimation is Expected Utility Theory. For $\alpha = 1$ the probability weighting function is the identity function, i.e. as already shown above

$$\pi(p) = \exp(-(-\ln(p))^{\alpha}) = \exp(-(-\ln(p))^{1}) = p.$$

We define the reference point r_{ref} as a fraction of the age-specific income in the estimation model. Recall that $r_{ref} = 0$ and $\alpha = 1$ implies Expected Utility Theory.

See Figure 25 for an illustration of possible reference point curves. The dash-dotted line for $r_{ref} = 0$ implies that for the evaluation of outcomes a reference point is not incorporated, i.e. either Rank-Dependent Utility, if $\alpha \neq 1$, or Expected Utility Theory, if $\alpha = 1$, are used.

Before we come to the actual estimation, let us repeat and formulate some definitions.[78] Let m_{J_m} be the vector of empirical moments, which was determined in Section 11, where J_m is the number of observation used for the estimation. The vector $m(\theta, \chi)$ consists of the theoretical population moment counterparts to m_{j_m} and $m_{J_S}(\theta, \chi)$ is the vector of simulated moments, where J_S is the number of simulated households. The parameter χ contains the parameters, which were calibrated and estimated in Sections 9 and 10. In the following we will refer to these sections as the first stage estimation or calibration phase.[79]

Let us now define the two relevant moment conditions

$$g(\theta, \chi) = [m(\theta, \chi) - m_{J_m}]$$

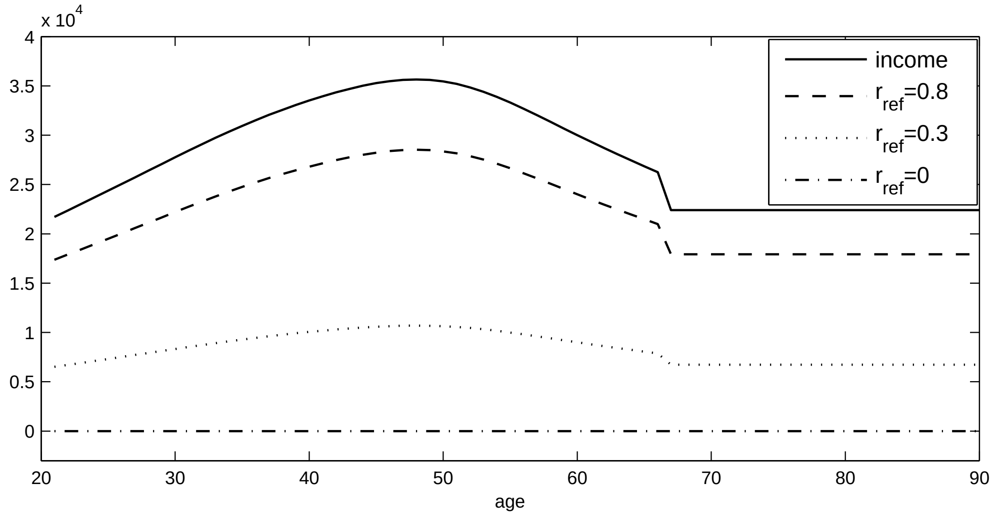
and

$$g_{J_S}(\theta, \chi) = [m_{J_S}(\theta, \chi) - m_{J_m}] .$$

[78]The notation is as in Laibson et al. (2007).

[79]A list summarizing these parameters is displayed in Table C.6 in the Appendix.

Figure 25: Example Reference Curves



For J_m and J_S going to infinity it must hold that

$$g(\theta_0, \chi_0) = g_{J_S}(\theta_0, \chi_0) = 0,$$

where (θ_0, χ_0) is the true parameter vector.

The objective function of the method of moments estimation is defined as

$$q(\theta, \chi) = g_{J_S}(\theta, \chi)' \times W \times g_{J_S}(\theta, \chi) \tag{6}$$

where W is a positive definite weighting matrix, with the size $N \times N$ and N representing the number of moments.

The function $q(\theta, \chi)$ has to be minimized with respect to the parameter vector θ. Recall, the vector χ was already estimated in the first stage of the Method of Simulated Moments approach, i.e. in the calibration sections, and is treated as constant in this step. Thus, the parameter estimate can be defined as

$$\hat{\theta} = \arg \min_{\theta} \ q(\theta, \hat{\chi}). \tag{7}$$

This estimator is consistent and asymptotically normally distributed if certain regularity conditions, see Pakes and Pollard (1989), are satisfied.

The minimization of the objective function is executed by an algorithm.[80] The algorithm continues to iterate, i.e. calculates the value of $q(\theta, \hat{\chi})$ with a new parameter combination α and r_{ref}, until the objective function converges to a minimum.[81]

Let Σ_g be the variance matrix of the population moments and Ω_g be the variance matrix of the second stage moments condition[82], which is defined by

$$\Omega_g = E\left[g(\theta_0, \chi_0)g(\theta_0, \chi_0)'\right].$$

At this point, one is normally also interested in the variance matrix of $\hat{\theta}$ to obtain the standard errors of $\hat{\alpha}$ and $\hat{r}_{ref}$. These are usually used to analyze if the parameter estimates are significantly different from the values which define the null hypothesis. However, the usage of the standard errors for significance checks is only reasonable if the estimator is (asymptotically) normally distributed. But, this is not the case in our setting, i.e. we can not proof that the regularity conditions from Pakes and Pollard (1989) for asymptotic normality are satisfied. Due to this we do not derive the variance matrix of $\hat{\theta}$. We will deal with the implications of this aspect in the next section.[83]

The proceeding of the estimation is as in a two-step Generalized Method of Moments estimation. This means, in a first run the weighting matrix W is the identity matrix. In the second run the inverse of the estimated variance matrix of the moments, $\hat{\Omega}_g$, of the first run is used as the optimal weighting matrix. The usage of the inverse of the estimated variance matrix implies that those conditions, where the distance between the empirical and simulated moments is relatively high, are weighted relatively low.

Finally, the model we use for the estimation is said to be overidentified, as we use more moments than parameters to be estimated. We use the J-Test as for

[80]We use as Laibson et al. (2007) the Nelder-Mead Simplex algorithm, which is the default in the "fminsearch" command in Matlab. Laibson et al. (2007) state that this algorithm is slower but more robust than derivative-based methods, and they prefer it because of the nonconvexities in quasi hyperbolic policy functions. In our setup there also prevail nonconvexities, due to the usage of Rank-Dependent Utility and Cumulative Prospect Theory.

[81]As one can not be sure with this method to have reached a global minimum, we repeat the estimation with different starting values.

[82]In contrast to Laibson et al. (2007) it holds that $\Omega_g = \Sigma_g / J_m$, as we use the same number of observations to derive the empirical moments.

[83]See Laibson et al. (2007) for a description how the variance matrix of the Method of Simulated Moments estimator can be derived.

the Generalized Method of Moments estimation in Section 10 for the Method of Simulated Moments setup - see also Laibson et al. (2007) and Gourinchas and Parker (2002). If the model is correct, than

$$\xi(\hat{\theta}, \hat{\chi}) = g_{J_S}(\hat{\theta}, \hat{\chi})' \cdot W_{opt} \cdot g_{J_S}(\hat{\theta}, \hat{\chi})$$

will be chi-squared distributed with $N - N_\theta$ degrees of freedom, where N is the number of moments and N_θ is the number of parameters. This test statistic is equal to $q(\hat{\theta}, \hat{\chi})$ if the optimal weighting matrix W_{opt} is used, which here is $\hat{\Omega}_g^{-1}$. In our case we have 2 degrees of freedom and the critical value is $\chi^2_{0.05} = 5.99$. Thus, if $q(\hat{\theta}, \hat{\chi})$ is smaller than 5.99 than the model is valid at a significance level of 5 percent.

14 Results and Hypothesis Testing

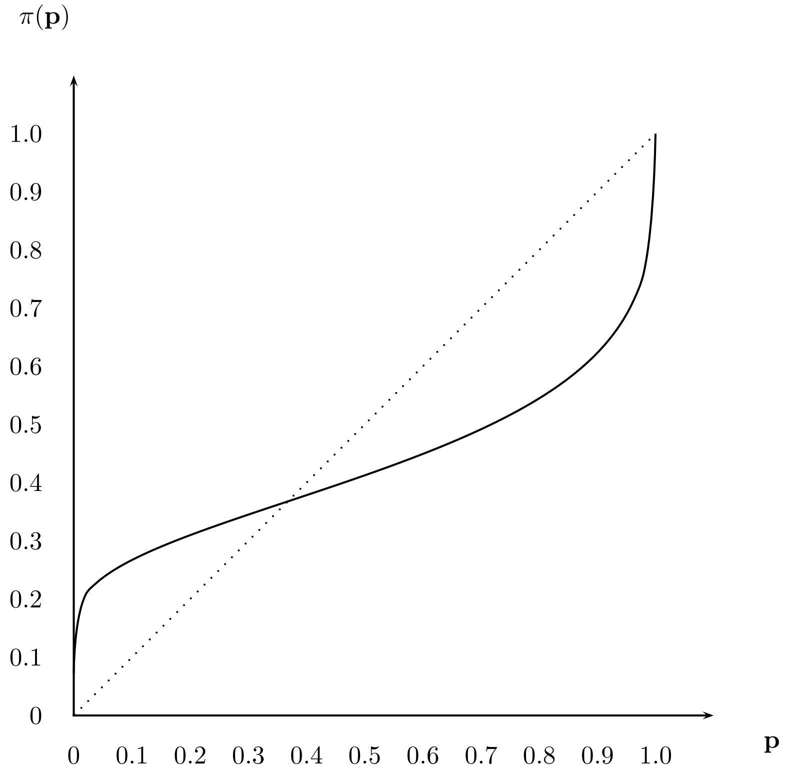
The results of the Method of Simulated Moments estimation are displayed in Table 13. The value of α is different from 1, which is the value of the null hypothesis. Recall, the probability function from Prelec (1998) is equal to the identity function if the curvature parameter $\alpha = 1$. Furthermore, the coefficient for the reference curve is different from 0, which is the case in Expected Utility Theory and Rank-Dependent Utility Theory. Thus, our estimates suggest for now that Cumulative Prospect Theory is the evaluation principle, which explains the data in use best. Figures 26 and 27 show the probability function and the reference curve, respectively, with the estimated parameters.

Table 13: MSM Estimation Results

	α	r_{ref}
coef	0.3328	0.1105

However, as mentioned above, we could not show that the regularity conditions, listed in Pakes and Pollard (1989), are satisfied. This implies that we have no indication that the estimator of the Method of Simulated Moments estimation is asymptotically normally distributed. That is why we can not use

Figure 26: Probability function with estimated α



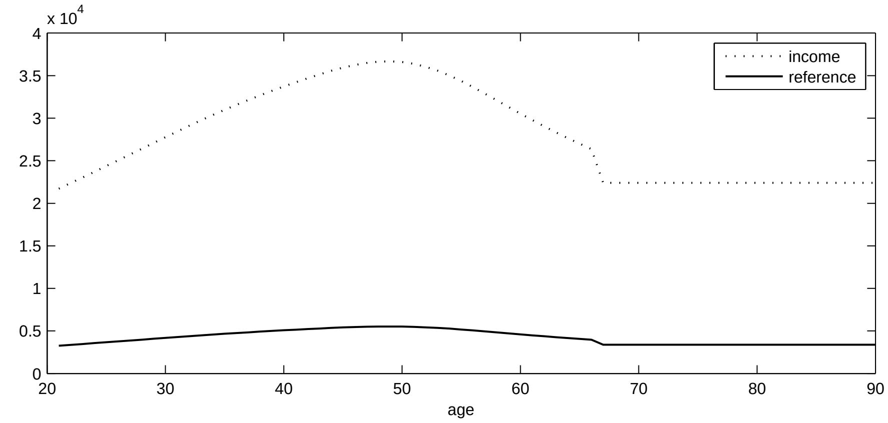
the standard errors to check whether the estimates of α and r_{ref} are **significantly** different from their null hypothesis values.

Due to this fact, we apply a bootstrapping[84] approach. We use the model of Section 2 to generate 10 000 artificial datasets under the null hypothesis, that is Expected Utility Theory. Hence, we set $\alpha = 1$ and $r_{ref} = 0$ in the model. We proceed that way as we want to know whether our estimates are significantly different from the values determined under Expected Utility Theory. In order to analyze that we are interested in the distribution of the estimates under the null hypothesis, i.e. the values for $\hat{\alpha}$ and $\hat{r}_{ref}$ which are generated when Expected Utility Theory is the prevailing system in our model. This means if the results of Table 13 lay in the tails of this distribution, then the estimates are significantly different from the null hypothesis values.

The difference of the bootstrap datasets stems from the different seeds, which are used to simulate the transitory shocks and the Markov process for

[84] The idea of bootstrap methods goes back to Efron (1979).

Figure 27: Age-specific income and reference curve



the persistent shocks. For each dataset i (for i=1,2,...,10 000) the wealth to income ratio for the 4 age cohorts are estimated, as described in Section 11. The bootstrap moments are denoted by

$$m_b^i(\theta_b, \hat{\chi}) = \begin{pmatrix} m_b^{i\,1} \\ m_b^{i\,2} \\ m_b^{i\,3} \\ m_b^{i\,4} \end{pmatrix},$$

where

$$\theta_b = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

i.e. with $\alpha = 1$ and $r_{ref} = 0$, and $\hat{\chi}$ is the parameter vector from the first stage of the Method of Simulated Moments estimation. Recall, the simulated moments were defined by

$$m_{J_S}(\theta, \hat{\chi}) = \begin{pmatrix} m_{J_S}^1 \\ m_{J_S}^2 \\ m_{J_S}^3 \\ m_{J_S}^4 \end{pmatrix}.$$

The bootstrap moment conditions are determined by

$$g_b^i(\theta, \hat{\chi}) = [m_{J_S}(\theta, \hat{\chi}) - m_b^i(\theta_b, \hat{\chi})],$$

where the index i stands again for one of the 10 000 generated datasets. The objective functions, which measure the distance between the bootstrap and the simulated moments, are given by

$$q_b^i(\theta, \hat{\chi}) = g_b^i(\theta, \hat{\chi})' \times W \times g_b^i(\theta, \hat{\chi}). \quad (8)$$

Now, we minimize these functions with respect to θ. This yields 10 000 estimates of α and r_{ref}. As the estimation procedure, which was described in the previous sections, is very time consuming we will not use an algorithm to find the minimum of the objective functions. Instead, we will perform a grid search. This means we will for both parameters α and r_{ref} solely concentrate on the interval [0;1], with a stepsize of 0.01. This implies the sets $s_\alpha = s_{r_{ref}} = \{0, 0.01, 0.02, 0.03, ..., 1\}$. A value smaller than 0 is not reasonable for both parameters, at least from a theoretical point of view. First, a negative reference point is not according to the concept of Cumulative Prospect Theory and, furthermore, in our setting $r_{ref} < 0$ would imply that the reference curve falls when the age-specific income increases. Second, for $\alpha < 0$ the probability weighting function turns its end points and smaller probabilities are perceived as higher than bigger probabilities - see Figure B.2 in the Appendix. Similar arguments apply for values bigger than 1. A reference point greater than the age-specific income might be reasonable for some ages for some households, but not over the whole life cycle and especially not over all households in average. The probability function again changes its features if α is bigger than 1. In this case low probabilities are underestimated and high probabilities are overestimated - see Figure B.3. This contradicts the concept of Cumulative Prospect Theory and Rank-Dependent Utility Theory. Due to these aspects, we restrict our search on the interval [0;1] for both parameters. This yields 10 201 combinations of the two parameters, as the stepsize for both is 0.01. For each of these combination pairs the simulated wealth to income ratios are calculated.

In a next step the distances between the simulated moments of all combinations and the moments of the 1st (i=1) dataset from the bootstrapping are calculated by the objective functions 8. The combination pair α and r_{ref} of the 10 201 possibilities, which yields the smallest value of $q_b^1(\theta, \hat{\chi})$ is the estimator of the first round, which uses the identity matrix as a weighting matrix. For the second run, we take the inverse of the estimated variance-covariance matrix of the moment conditions of the first run as the weighting matrix W in Equation 8.

This procedure is executed for all the artificial datasets from the bootstrapping approach. Thus, we finally have 10 000 estimates of the parameters α and r_{ref}. A first impression of the distribution of the parameter estimates can be obtained by looking at the histograms, see Figures 28 and 29.

Now, we focus on the question whether the estimated parameters α and r_{ref} are significantly different from the values of the null hypothesis. For that reason we look at the percentiles of the distributions of the parameters. For α we are interested in the threshold, which declares that 99 % of the distribution lies right of it, as the value of the null hypothesis is equal to 1 . This threshold is 0.63 and, hence, is bigger than 0.3328. This implies that α is significantly different from 1 at a significance level of 1%. In order to check the significance of r_{ref}, we look at the threshold, where 99% of the observations lie left of it. This yields the interval [0;0.32]. The estimate $r_{ref} = 0.1105$ is in this interval and is therefore not significantly different from 0 at the level of 1%. The same holds for a significance level of 10%, where the interval is given by [0;0.42].

Hence, the final result of the significance check is that Rank-Dependent

Figure 28: Distribution of α under Expected Utility Theory

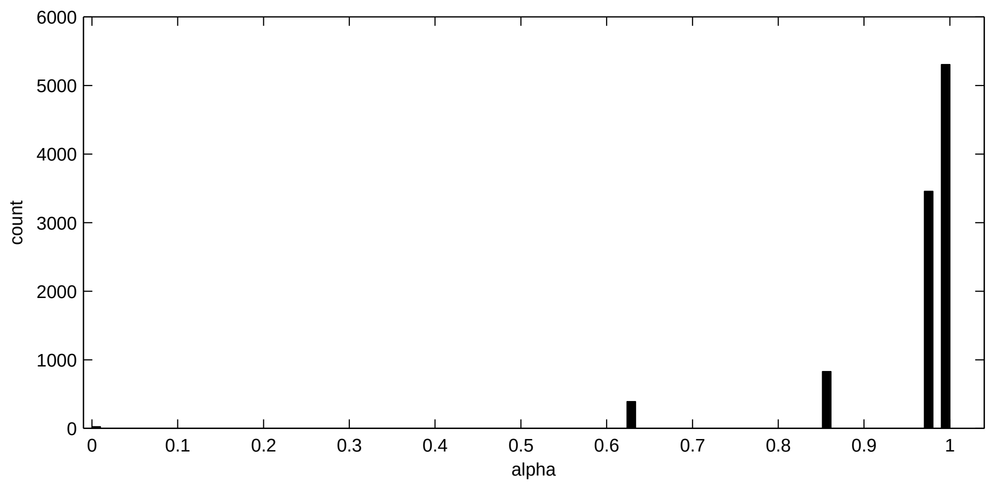
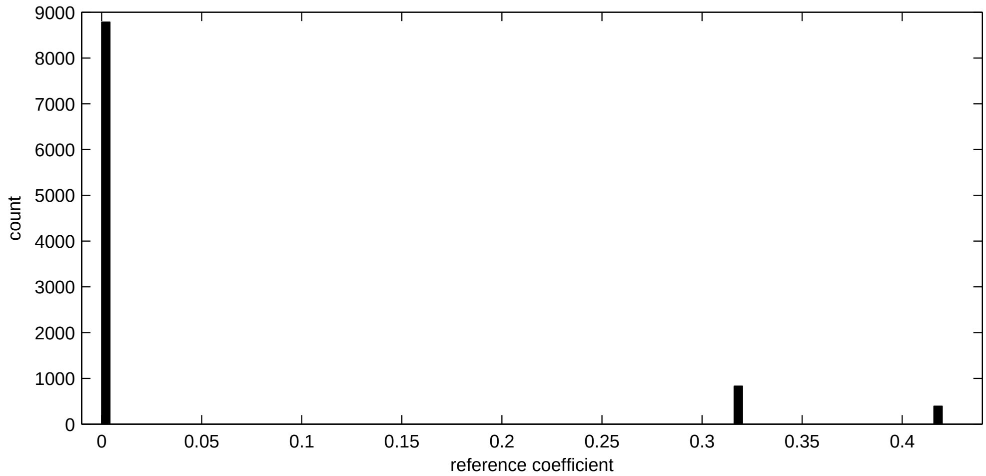


Figure 29: Distribution of r_{ref} under Expected Utility Theory



Utility Theory is the risk evaluation system recommended by our estimation procedure, and not as suggested above Cumulative Prospect Theory.
At this point recall, we restricted our search on the interval [0;1]. Thus, there are probably values beyond the boundaries 0 and 1, which belong to the distribution of the parameter estimates under the null hypothesis. Though, these are in some cases not reasonable from a theoretical point of view. Nevertheless, this implies that the critical values of the percentiles could change, if one does not stick to the interval [0;1]. However, we believe that these possible changes would not be in favor of the null hypotheses. This is due to the aspect that for example estimates of α would more likely be to the right of 1 than to the left of 0, as the bootstrap datasets are generated under the null hypothesis with $\alpha = 1$. For the percentiles this means that the critical value for α would be pushed to the right and for r_{ref} to the left.

Note that one should also take into the account the variance of the first stage parameter vector $\hat{\chi}$, in order to simulate the distributions of $\hat{\alpha}$ and $\hat{r}_{ref}$. However, this is at the moment way beyond the computing capabilities we have.

Finally, as mentioned above, we perform a J-Test, because we use more moments than parameters to be estimated, i.e. the model is overidentified. The value of the objective function $q(\hat{\theta}, \hat{\chi}) = 5.1812$ and thereby smaller than the critical value of the chi-square distribution with two degrees of freedom, which is 5.99. This implies that our the specification is valid at the 5% level.

15 Interpretation

The result that $r_{ref} = 0.1105$ would indicate that a reference point is used to evaluate outcomes. But, as the estimate is not significantly different from 0, we come to the conclusion that households do not link the evaluation of their consumption to the development of their age-specific income. This means the carriers of value or utility are the absolute terms of consumption.

The parameter estimate for the probability function suggests some attitudes of the households how they evaluate risk. These attitudes partly depend on particular situations a household is in. The following interpretations are made under the assumption that households act according to Rank-Dependent Utility Theory with the estimated parameter. In order to give an impression of the deviations from Expected Utility Theory, we use it as a base case for the interpretation. Table 14 displays, in the first line the probabilities of the 5 transitory shocks in the model and in the second line the transformed probabilities used to evaluate the risk of these shocks in each period. From comparing the values we see that the lowest probabilities (0.0086) are overestimated and the highest probability (0.7735) is underestimated. For the two moderate shocks (2 times the standard deviation) the transformed values go in different directions. The probability for the negative shock is overestimated, whereas it is underestimated for the positive one. However, the most remarkable result is that the transformation for the extreme shocks is over 20 times higher than its chance of occurring. Furthermore, the distribution is no longer symmetric.

Table 14: Probabilities and weights of transitory shocks

$-5 \cdot \sigma_v$	$-2 \cdot \sigma_v$	0	$2 \cdot \sigma_v$	$5 \cdot \sigma_v$
0.0086	0.1047	0.7735	0.1047	0.0086
0.1853	0.2045	0.3365	0.0874	0.1863

Let us now consider the transition weights of the Markov process. The initial and transformed probabilities for the different states are shown in Table 15.

The first column lists the initial state a household is in. The row pairs "initial" and "trans" give the probabilities from the Markov process and the weights used by the evaluation process of Rank-Dependent Utility Theory. The first row represents the states a household could move to or stay in.

Table 15: Weights for change in Markov states

	low	zero	high
low - initial	0.724	0.138	0.138
low - trans	0.6630	0.0520	0.2850
zero - initial	0.138	0.724	0.138
zero - trans	0.4114	0.3036	0.2850
high - initial	0.138	0.138	0.724
high - trans	0.4114	0.0852	0.5034

If a household is in the low income state (1st row pair) then the probability of staying is underestimated. Hence, the general chance of a change to a better state is highly weighted, especially a move to the high income level. One could interpret this as a quite optimistic attitude by the households in this situation. In contrast to that, if a household is in the high income state (3rd row pair), the values suggest a quite pessimistic tendency. The chance of staying in the same state is here even more underestimated than in the case of being initially in the low income situation. The lowest weight for the probability to stay in the same state we find if a household is in the middle income position (2nd row pair). Thus, the probability of a change is overestimated, with again a pessimistic tendency, as the value of a move to the low income state is higher than the one for the high income state.

Now, let us turn to the risk of survival. Here, also the aspect that high probabilities are underestimated prevails. For example a survival rate of 0.9994 (age 20) yields a weight of 0.9188. Considering the survival rates separately this does not seem to have so much influence. But, over the whole life cycle this leads to quite big differences. The weight of the contribution of the 90th year at the age of 20, in the model, is equal to 1.55e-7, although the conditional probability of living from 20 to 90 is 0.1573. Hence, this can be interpreted in that way that young households intensely underestimate the probability of getting quite old.

We believe that the interpretations about the risk evaluation of the simulated households remain valid, even if some aspects in the model setup are changed. For example if the number of transitory shocks in each period is increased. A detailed analysis may be subject to future work.

16 Robustness Check

This section checks the robustness of the results of Section 14. We reestimate $\hat{\alpha}$ and $\hat{r}_{ref}$ with different values for the coefficient of relative risk aversion ρ, the consumption flow parameter γ, and the interest rate for negative liquid asset holdings R_X^{CC}. We apply the same bootstrap method as in Section 14 as we still can not assume that the estimator is asymptotically normally distributed. In contrast to the approach above, we reduce the stepsize in the interval [0;1] to 0.025 for the grid search, due to computation time.

We set the coefficient of relative risk aversion ρ equal 1 and 3, respectively. Laibson et al. (2007) state that in most life cycle consumption models the coefficient of relative risk aversion lies between 0.5 and 5, but there is no generally accepted value. Nevertheless, we use these two values, in order to get an impression of the reaction of the estimates for α and r_{ref}. The estimates, displayed in Table 16, are quite close to those we obtained in the base case, for both alternative values of ρ.

The bootstrap approach indicates that $\hat{\alpha}$ is different from 1 at a significance level of 1%, with $\rho = 1$. In contrast to the base case, here the estimate of r_{ref} is also sginificantly different from its null hypothesis value at the 1% level. According to this, the estimation procedure with $\rho = 1$ suggests that Cumulative Prospect Theory is the prevailing risk evaluation theory in the model. In the case with $\rho = 3$ the parameter estimate of α is also significantly different from 1 at the 1% level. But, $\hat{r}_{ref}$ is not significantly different from 0 at the 1%, 5% or

10% level. Thus, this setting comes to the same conclusion as the base case.

For the consumption flow, which is generated by the illiquid asset, we set the parameter γ equal to 0.03 and 0.07. Recall, the illiquid holdings in the model can be interpreted as home equity. The return on home equity is not easy to determine, as there are various aspects, which one has to take into account, like capital gains, use-value, maintenance and taxes.[85] Furthermore, it is surely relevant if a houseowner lives in the house or gets paid a rent. Here, we orientate again on Laibson et al. (2007), who set γ equal to 0.0338 and 0.0659. The changes of the parameter estimates of α and r_{ref} are negligible. The results of the bootstrap indicate, as in the base case, that $\hat{\alpha}$ is significantly different from 1 and $\hat{r}_{ref}$ is not significantly different from 0, for both calibration cases.

Now, we look at changes of the interest rate for credits, namely R_X^{CC}. We set this rate equal to 8% and 12%. The intuition behind these changes is to vary the price to borrow money. If getting money is cheaper, then households have an incentive to borrow more money, especially in younger years, and vice versa. We find that the estimates of α and r_{ref} nearly stay the same as in the base case. The bootstrap approach shows that the parameter estimates are both significantly different from their null hypothesis values at the 1% level. This implies that these two calibration cases also suggest that Cumulative Prospect Theory fits the data best.

Note, it is also possible to simultaneously estimate the parameters, considered in this section, together with $\hat{\alpha}$ and $\hat{r}_{ref}$. For example Laibson et al. (2007) do so for the coefficient of relative aversion, as the sensitivity of their parameter estimates is quite high to changes of the coefficient of relative risk aversion. However, our robustness checks do not indicate any high sensitivities.

We perform an overidentification test, i.e. a J-Test as above, for all different cases and find that the specifications are all valid at the 5% level.

The robustness check shows that the estimates of α and r_{ref} are not sensitive to changes of the coefficient of relative risk aversion, the consumption flow parameter or the interest rate for negative, liquid asset holdings. We find for some cases that the significance of the results has changed in comparison to the base case. In these cases Cumulative Prospect Theory is preferred, in-

[85] See Flavin and Yamashita (2002).

stead of Rank-Dependent Utility Theory. This is a shift more away from the null hypothesis, namely Expected Utility Theory, as this implies that also a reference curve is used to evaluate risky situations. However, the grid we used for the bootstrap approach in the base case is tighter, therefore we give that result more weight and stay with the conclusion that Rank-Dependent Utility Theory is the risk evaluation system prevailing in our model.

Table 16: Estimation results robustness check

	α	r_{ref}
$\rho = 0.99$	0.3229	0.1143
$\rho = 3$	0.3328	0.1103
$\gamma = 0.03$	0.3308	0.1090
$\gamma = 0.07$	0.3327	0.1005
$R_X^{CC} = 0.08$	0.3328	0.1102
$R_X^{CC} = 0.12$	0.3328	0.1103

17 Extensions

In the setup of the model and the simulation we introduced some restrictions, which were mainly due to the aspect to reduce the computation time of the Method of Simulated Moments approach. Relaxing these would be one way to extend the model or to test if our results are robust to such changes. However, there are also some other ideas which can be applied to the setup of the model or the simulation procedure.[86]

[86] As we use nearly the same model as Laibson et al. (2007), some of the aspects mentioned in this section can also be found in their paper.

17.1 Relaxing Restrictions

A quite unrealistic assumption in the design of the model concerns the features of the illiquid asset. It is assumed that it is completely illiquid over the lifetime of a household. This implies that it can not be used to compensate any kind of shock which occurs. Moreover, it can not be used to increase the disposable income when a household is retired. A way of relaxing this restriction is to permit that these holdings can be liquidated, but in connection with some transaction costs and perhaps a particular time delay. This would be necessary to maintain a difference between liquid and illiquid assets.

One could also think of separating the illiquid asset, e.g. in illiquid wealth and pension plans. The pension plans would then generate additional income for a household if it is in retirement, without causing any transaction costs.

Furthermore, the interest rate of the illiquid asset is set equal to 1. This is mainly due to the point that we considered wealth as some sort of home equity. But, doing so one could also account for the possibility to borrow against this property.

Another possibility would be to increase the number of transitory shocks or the states for the Markov process. But, one would face a direct trade-off between obtaining a better approximation of the income process versus a substantial increase in the computation time of the Method of Simulated Moments estimation. The same holds for the stepsizes we have chosen for the liquid and illiquid assets. Recall, we determined that the liquid holdings can increase or decrease in steps of €2 000 and the illiquid asset can only increase by steps of €50 000. We do not believe that a reduction in the stepsize of the liquid asset would change our results. In contrast to that, the stepsize of the illiquid asset is pretty high. This implies that in the model the accumulation of wealth could be delayed by some periods, as a household has to save money in the liquid holdings first if it has decided to invest in the illiquid asset. In combination with the wealth to income ratios we have chosen as the moments, this could have an influence on the results. However, this would also extremely increase the computation time. For example, if we cut down the stepsize of the illiquid asset to €10 000, the computing time would increase by more than the factor 5.

Another idea, which we also did not consider due to the same argument, is another determination of the time unit. In reality households do not only make one decision in a year concerning their investments and especially consumption. Thus, a richer model could use months, weeks or even days to simulate the decisions of a household.

17.2 Options

For the estimation we used the wealth to income ratios of the age cohorts 20 to 29, 30 to 39, 40 to 49 and 50 to 59, as the moments in Method of Simulated Moments procedure. There are a lot of other potential moments one could choose, like the marginal propensity to consume[87], consumption to income ratio for different age cohorts, the fraction of wealth that is illiquid, etc.. The choice of the moments, as pointed out in the theory of the Generalized Method of Moments approach, is important for the estimation results. But, if the model is correctly specified by the chosen moments, then a change of these would not alter the estimated parameter values. Nevertheless, this could be a test of the model and we leave that to future work.

The way in which we use the theory of Kahneman and Tversky (1992) deviates in some aspects from their original paper. We did this in order to achieve equality between Rank-Dependent Utility and Cumulative Prospect Theory. This made it possible to nest all three theories in the same model and simulation process, respectively. The main deviation is that we do not use different probability functions for gains and losses. This could perhaps alter the estimation results, but therefore a different setting in the model would be necessary.

According to Cumulative Prospect Theory one could also use another definition for the reference point or points, which categorizes outcomes into gains and losses. One possibility would be a concept from the Habit Formation or Habit Persistence literature[88] , which determines the reference point in terms of past consumption. However, as long as we solve our model via Backwards Induction this would imply some serious computational burdens.

[87] As mentioned, we tried to estimate the propensity to consume from our available data, but did not come up with any useful results.

[88] See e.g. Duesenberry (1949), Pollak (1970) or Constantinides (1990).

17.3 Future Research

The model could be extended by allowing the possibility of quasi-hyperbolic discounting, as in Laibson et al. (2007). Here, the question arises if the two concepts, preferences like Rank-Dependent Utility or Cumulative Prospect Theory and hyperbolic discounting, do influence each other, i.e. that they are highly correlated. For example, using hyperbolic discounting is a way to capture the fact that individuals do not save sufficiently for old age. The probability function of Rank-Dependent Utility or Cumulative Prospect Theory[89] has the feature to underestimate high probabilities. This means that households like in our model underestimate the probability of getting really old. This implies that a household also does not save sufficiently for old age, due to the characteristics of the probability function. However, one could also examine if such an ambiguous effect exists.

Another possibility for future research is to account for heterogeneity in preferences. If one considers the experiments, like those of Kahneman and Tversky (1979) described in Section 3.1, the results indicate that solely a majority of the participants, but not all, violate Expected Utility Theory. This implies that there is also a fraction, which acts according to the standard theory in economics. Thus, our model could be extended by a further parameter, which separates the households into those following Expected Utility Theory and those having Rank-Dependent Utility or Cumulative Prospect Theory preferences. Such a parameter could also be estimated in the Method of Simulated Moments procedure. One could also think of intra-household heterogeneity. The idea is that households could change their attitudes over time. For example a household could learn that its way of evaluating the future is wrong and therefore changes its preferences. However, before going in that direction one should first have a look at some data if there is any evidence for that kind of heterogeneity.

Finally, an aspect, which also Laibson et al. (2007) mention, is the possibility that people are acting naively. Recall from Section 4, if preferences like Rank-Dependent Utility or Cumulative Prospect Theory prevail and the model is not solved via Backwards Induction, but with the usage of the Re-

[89] Recall, the probability functions of these theories are equal in our setting, but differ a little bit in the original papers of Quiggin (1982) and Kahneman and Tversky (1992).

duction of Compound Lotteries Axiom, this leads to dynamically inconsistent decisions. The assumption that individuals can solve a model, like the one presented here, by Folding Back probably might be beyond their computational capabilities. Thus, it could be reasonable to assume that individuals do not apply Folding Back. This, as stated, would imply dynamically inconsistent preferences. Strotz (1955), Akerlof (1991) and O'Donoghue and Rabin (1999a, 1999b) suggest that people with that kind of preferences could act naively, i.e. they believe that later selfs behave according to the optimal initial plan they set up earlier. Hence, one could assume this kind of behavior for the solution of the model. It could also be possible to permit both ways, i.e. a fraction of households using Folding Back and the others acting naive - as above heterogeneous households in the simulation. In general, the objective of any of these ideas about heterogeneity should be if these are able to explain reality better than others, i.e. if they fit the available data better or not.

18 Conclusion

The model in this thesis is designed in such a way, that two parameters indicate which risk evaluation theory is used, i.e. Expected Utility Theory, Rank-Dependent Utility Theory or Cumulative Prospect Theory. For the solution of our model we find that Folding Back should be used in the case of Rank-Dependent Utility or Cumulative Prospect Theory, in order to circumvent dynamic inconsistencies. Furthermore, we show that the critique given in the literature about the usage of Folding Back in combination with Non-Expected Utility theories does not apply, if one does not assume the Reduction of Compound Lotteries Axiom, at least for the two theories we use.

The Method of Simulated Moments procedure provides an estimate of the curvature parameter α of the probability weighting function equal to 0.3328. For the reference parameter r_{ref} we get an estimate of 0.1105. The estimates of the two parameters α and r_{ref} indicate that Cumulative Prospect Theory is the risk evaluation model, which matches the real life data from the SAVE and SOEP survey best. However, we could not show that the parameter estimates are asymptotically normally distributed.

Due to this we apply a bootstrap method, in order to check whether the estimated values are significantly different from the values of the null hypothesis, which is Expected Utility Theory. We find that the estimate of α is significantly different from 1 at the 1% level. The estimated value of r_{ref} is not significantly different from 0, even at the 10% level. This leads to the conclusion that Rank-Dependent Utility Theory is the risk evaluation system recommended by our model. This implies that the households in our setup do not define consumption in terms of gains and losses due to a reference, but they overestimate low probabilities and underestimate high probabilities. These results are also robust to changes in the calibration. Nevertheless, future work will be required to confirm these estimation results with other models and/or data.

References

Allais, Maurice (1953), "Le Comportement de l'Homme Rational devant le Risque, Critique des Postulats et Axiomes de l'Ecole Americaine", Econometrica, 21, pp. 503-546.

Akerlof, George (1991), "Procrastination and Obedience", American Economic Review, 81(2), pp. 1-19.

Barnes, Jim D. and James E. Reinmuth (1976), "Comparing Imputed and Actual Utility Functions in a Competitive Bidding Setting", Decision Sciences, 7, pp. 801-812.

Blundell, Richard, Martin Browning, and Costas Meghir (1994), "Consumer Demand and the Life-Cycle Allocation of Household Expenditures", Review of Economic Studies, 61, pp.57-80.

Börsch-Supan, Axel, Michaela Coppola, Lothar Essig, Angelika Eymann and Daniel Schunk (2009), "The German SAVE Study - Design and Results", `http://www.mea.uni-mannheim.de/fileadmin/files/polstudies/3aferngy0iaowiys_MEA_Study_6.pdf`, (last access: 12.01.2011).

Brauns, Hidlegard and Susanne Steinmann (1999), "Educational Reform in France, West-Germany and United Kingdom", ZUMA-Nachrichten,44, 23: pp. 7-44.

Carroll, Christopher (1992), "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence", Brooking Papers on Economic Activity, 1992(2), pp. 61-156.

Carroll, Christopher (1997), "Buffer-Stock and the Life Cycle/Permanent Income Hypothesis", Quarterly Journal of Economics, 112(1), pp. 1-55.

Chew, Soo Hong (1983), "A Generalization of the Quasilinear Mean with Applications to the Measurement of Inequality and Decision Theory Resolving the Allais Paradox", Econometrica, 51(4), pp. 1065-1092.

Chew, Soo Hong, Larry G. Epstein and Uzi Segal (1991) , "Mixture Symmetry and Quadratic Utility", Econometrica, 59(1), pp. 139-163.

Chew, Soo Hong and Kenneth MacCrimmon (1979), "Alpha-Nu Choice Theory: A Generalization of Expected Utility Theory", Working Paper No. 669, University of British Columbia.

Chew, Soo Hong and William Waller (1986), "Empirical Tests of Weighted Utility Theory", Journal of Mathematical Psychology, 30(1), pp. 55-72.

Cohen, Michele, Jean-Yves Jaffray, and Tanios Said (1987), "Experimental Comparisons of Individual Behavior under Risk and under Uncertainty for Gains and Losses", Organizational Behavior and Human Decision Processes, 39, pp. 1-22.

Constantinides, Georg M. (1990), "Habit Formation: A Resolution of the Equity Premium Puzzle", Journal of Political Economy, 98(3), pp. 519-543.

Deaton, Angus (1991), "Saving and Liquidity Constraints", Econometrica, 59(5), pp. 1221-1248.

Duesenberry, James S. (1949), "Income, Saving, and the Theory of Consumer Behavior", Cambridge Massachusetts: Harvard University Press.

Duffie, Darrell and Ken Singleton (1993), "Simulated Moments Estimation of Markov Models of Asset Prices", Econometrica, 61(4), pp. 929-952.

Edwards, Ward (1955), "The Prediction of Decisions Among Bets", Journal of Experimental Psychology, 50(3), pp. 201-214.

Edwards, Ward (1962), "Subjective Probabilities Inferred from Decisions", Psychological Review, 69(2), pp. 109-135.

Efron, Bradley (1979), "Bootstrap Methods: Another Look at the Jackknife", The Annals of Statistics, 7(1), pp. 1-26.

Ellsberg, Daniel (1961), "Risk, Ambiguity and the Savage Axiom", Quarterly Journal of Economics, 75(4), pp. 643-669.

Epstein, Larry G. (1992), "Behavior under Risk: Recent Developments in Theory and Application", in J.J. Laffont (Ed.) *Advances in Economic Theory II*, Cambridge University Press, New York, pp. 1-63.

Fishburn, Peter C. (1983), "Transitive Measurable Utility", Journal of Economic Theory, 31(2), pp. 293-317.

Fishburn, Peter C. and Gary A. Kochenberger (1979), "Two-Piece von Neumann-Morgenstern Utility Functions", Decision Sciences, 10(4), pp.503-518.

Flavin, Marjorie and Takashi Yamashita (2002), “Owner-Occupied Housing and the Composition of the Household Portfolio”, American Economic Review, 92(1), pp. 345-362.

Fox, Craig R. and Amos Tversky (1995), "Weighing Risk and Uncertainty", Psychological Review, 102(2), pp. 269.283.

Frick, Joachim R. and John P. Haisken-DeNew (2005), "Desktop Companion to the German Socio-Economic Panel (SOEP)", `www.diw.de/documents/dokumentenarchiv/17/diw_01.c.38951.de/dtc.409713.pdf`, (last access: 12.01.2011).

Gourinchas, Pierre-Olivier and Jonathan Parker (2002), "Consumption over the Life Cycle", Econometrica, 70(1), pp. 47-89.

Grayson, Jackson C. (1969), "Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators", Cambridge, Massachusetts: Graduate School of Business, Harvard University.

Green, Jerry and Bruno Jullien (1988), "Ordinal Independence in Non-Linear Utility Theory", Journal of Risk and Uncertainty, 1(4), pp. 355-387.

Grether, David M. and Charles R. Plott (1979), "Economic Theory of Choice and the Preference Reversal Phenomenon", The American Economic Review, 69(4), pp. 623-638.

Hagen, Ole (1979), "Toward a Positive Theory of Preferences under Risk", in Allais and Hagen (Ed.) *Expected Utility Hypothesis and the Allais Paradox*, Dordrecht: D. Reidel Publishing Co., pp. 271-302.

Halter , Albert N. and Gerald W. Dean (1971), "Decisions under Uncertainty", Cincinnati: South Western Publishing Co..

Hammond, Peter J. (1988), "Orderly Decision Making: A Comment on Professor Seidenfeld", Economics and Philosophy, 4(2), pp. 292-297.

Hammond, Peter J. (1988), "Consequentialist Foundations for Expected Utility", Theory and Decision, 25, pp. 25-78.

Hansen, Lars P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", Econometrica, 50, pp. 1029-54.

Hersey, John C. and Pamela J.H. Schoemaker (1980), "Prospect Theory's Reflections Hypothesis: A Critical Examination", Organizational Behavior and Human Performance, 25, pp. 395-418.

Hey, John D. (1984), "The Economics of Optimism and Pessimism: A Definition and some Applications", Kyklos, 37(2), pp. 181-205.

Hilton, Ronald (1990), "Failure of Blackwell's Theorem under Machina's Generalization of Expected Utility Analysis without the Independence Axiom", Journal of Economic Behavior and Organization, 13(2), pp. 233-244.

Holler, Manfred (1983), "Do Economics Students Choose Rationally? A Research Note", Social Science Information, 22(4-5), pp. 623-630.

Hubbard, Glenn, Jonathan Skinner and Stephan Zeldes (1994), "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving", Carnegie-Rochester Conference Series on Public Policy, 40, pp. 59-126.

Hubbard, Glenn, Jonathan Skinner and Stephan Zeldes (1997), "Precautionary Saving and Social Insurance", Journal of Political Economy, 103(2), pp. 369-399.

Jaffray, Jean-Yves and Thomas D. Nielsen (2006), "Dynamic Decision Making without Expected Utility: An Operational Approach", European Journal of Operational Research, 169(1), pp. 226-246.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem", Journal of Political Economy, 98(6), pp. 1325-1348.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk", Econometrica, 47, pp. 263-291.

Kahneman, Daniel and Amos Tversky (1981), "The Framing of Decisions and the Psychology of Choice", Science, 211, pp. 453-458.

Kahneman, Daniel and Amos Tversky (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty", Journal of Risk and Uncertainty, 5, pp. 297-323.

Kamae, Teturo, Ulrich Krengel and George L. O'Brien (1977), "Stochastic Inequalities on Partially Ordered Spaces", Annals of Probability, 5, pp. 899-912.

Karmarkar, Uday (1978), "Subjectively Weighted Utility: A Descriptive Intention of the Expected Utility Model", Organizational Behavior and Human Performance, 21(1), pp. 61-72.

Karmarkar, Uday (1979), "Subjectively Weighted Utility and the Allais Paradox", Organizational Behavior and Human Performance, 24(1), pp. 67-72.

Keasey, Kevin (1984), "Regret Theory and Information: A Note", Economic Journal, 94(375), pp. 645-649.

Keeney, Ralph and Robert Winkler (1985),"Evaluating Decision Strategies for Equity of Public Risk", Operations Research, 33(5), pp. 955-970.

Kreps, David M. and Evan L. Porteus (1975), "Temporal Resolution of Uncertainty and Dynamic Choice Theory", Econometrica, 46, pp. 185-200.

Laibson, David (1997), "Golden Eggs and Hyperbolic Discounting", Quarterly Journal of Economics, 62(2), pp. 443-478.

Laibson, David, Andrea Repetto, and Jeremy Tobacman (2003), "A Debt Puzzle", in Aghion, Frydman, Stiglitz and Woodford (Ed.) *Knowledge, Information, and Expectations in Modern Economics: In Honor of Edmund S. Phelps*, Princeto:, Princeton University Press.

Laibson, David, Andrea Repetto, and Jeremy Tobacman (2007), "Estimating Discount Functions with Consumption Choices over the Life Cycle", Economics Series Working Papers 341, University of Oxford.

LaValle, H. Irving and Kenneth R. Wapman (1986),"Rolling Back Decision Trees Requires the Independence Axiom", Management Science, 32(3), pp. 382-385.

Loomes, Graham and Robert Sugden (1984), "Regret Theory and Information: A Reply", Economic Journal, 94(375), pp. 649-650.

MacCrimmon, Kenneth and Stig Larsson (1979), "Utility Theory: Axioms versus Paradoxes", in Allais and Hagen (Ed.) *Expected Utility Hypothesis and the Allais Paradox*, Dordrecht: D. Reidel Publishing Co., pp. 333-409.

Machina, Mark J. (1989), "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty", Journal of Economic Literature, 27, pp. 1622-1668.

McFadden, Daniel (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration", Econometrica, 57(5), pp. 995-1026.

Markowitz, Harry M. (1959), "The Utility of Wealth", Journal of Political Economy, 60, pp. 151-158.

O'Donoghue, Ted and Matthew Rabin (1999a), "Doing It Now or Later", American Economic Review, 89(1), pp. 103-124.

O'Donoghue, Ted and Matthew Rabin (1999b), "Incentives for Procrastinators", Quarterly Journal of Economics, 114(3), pp. 769-816.

Pakes, Ariel and David Pollard (1989), "Simulation and Asymptotics of Optimization Estimators", Econometrica, 57(5), pp. 1027-57.

Payne, John W., Dan J. Laughhunn and Roy Crum (1981), "Aspiration Level Effects and Risky Behavior", Management Science, 27, pp. 958-958.

Phelps, Edmund S. and Robert A. Pollak (1968), "On Second-best National Saving and Game Equilibrium Growth", Review of Economic Studies, 35, pp. 185-199.

Pollak, Robert A. (1968), "Consistent Planing", Review of Economic Studies, 35, pp. 201-208.

Pollak, Robert A. (1970), "Habit Formation and Dynamic Demand Functions", Journal of Political Economy, 78(4), pp. 745-763.

Prelec, Drazen (1998), "The Probability Weighting Function", Econometrica, 66(3), pp.497-527.

Quiggin, John (1982), "A Theory of Anticipated Utility", Journal of Economic Behavior and Organization, 3, pp. 323-343.

Ronen, Joshua (1971), "Some Effects on Sequential Aggregation in Accounting on Decision Making", Journal of Accounting Research, 9(2), pp. 307-332.

Rubin, Donald B. (1987), "Multiple Imputation for Nonresponse in Surveys", New York: Wiley.

Sarin, Rakesh and Peter P. Wakker (1998), "Dynamic Choice and NonExpected Utility", Journal of Risk and Uncertainty, 17, pp. 87-119.

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

Schmidt, Ulrich, Chris Starmer and Robert Sugden (2008), "Third-Generation Prospect Theory", Journal of Risk and Uncertainty, 36, pp. 203-223.

Schunk, Daniel (2007), "A Markov Chain Monte Carlo Multiple Imputation Procedure for Dealing with Item Nonresponse in the German SAVE Survey", MEA Discussionpaper 121-07.

Segal, Uzi (1984), "Nonlinear Decision Weights with the Independence Axiom", UCLA Economics Working Paper 353, UCLA Department of Economics.

Segal, Uzi (1990),"Two-Stage Lotteries without the Reduction Axiom", Econometrica, 58(2), pp. 349-377.

Snowball, Doug and Cliff Brown (1979), "Decision Making Involving Sequential Events: Some Effects of Disaggregated Data and Dispositions Toward", Decision Sciences, 10(4), pp. 527-545.

Starmer, Chris and Robert Sugden (1991), "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation", American Economic Review, 81(4), pp. 971-978.

Starmer, Chris (2000), "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice Under Risk", Journal of Economic Literature, 38, pp. 332-382.

Stern, Steven (1997), "Simulation-Based Estimation", Journal of Economic Literature, 35(4), pp. 2006-2039.

Strotz, Robert H. (1956), "Myopia and Inconsistency in Dynamic Utility Maximization", Review of Economic Studies, 23, pp.165-180.

Swalm, Ralph O. (1966), "Utility Theory - Insights into Risk Taking", Harvard Business Review, 44, pp. 123-136.

Tversky, Amos (1975), "A Critique of Expected Utility Theory: Descriptive and Normative Considerations", Erkenntnis, 9(2), pp.163-173.

Tversky, Amos and Peter P. Wakker (1995), "Risk Attitudes and Decision Weights", Econometrica, 63(6), pp. 1255-1280.

Yaqub, Muhammad Zafar, Gökhan Saz and Dildar Hussain (2009), "A Meta Analysis of Empirical Evidence on Expected Utility Theory", European Journal of Economics, Finance and Administrative Sciences, 15, pp. 117-133.

Wakker, Peter P. (1988), "Nonexpected Utility as Aversion to Information", Journal of Behavioral Decision Making, 1(3), pp 169-75.

Wehrung, Donald A. (1989), "Risk Taking over Gains and Losses: A Study of Oil Executives", Annals of Operation Research, 19, pp. 119-139.

Williams, C. Arthur (1966), "Attitudes toward Speculative Risks as an Indicator of Attitudes toward Pure Risks", Journal of Risk and Insurance, 33, pp. 577-586.

Appendix

A Violation of First-Order Stochastic Dominance

Kahneman and Tversky (1979) state that $\pi(p) < p$, except for low probabilities. However, this feature can exhibit violations of dominance with probabilities of the form $1/n$. The following example stems from Quiggin (1982). Suppose, the outcome X is received with certainty and is compared to a lottery, where all outcomes are of the form $X + x_i$ for $i = 1, ..., n$, with x being a random variable $0 < x \leq \epsilon$, and each outcome with probability $1/n$. If $\pi(1/n) < 1/n$ and for a sufficiently small ϵ it can be that

$$U(X) > \sum_{i=1}^{n} \pi(1/n) U(X + x_i).$$

This violates dominance, as the lottery always yields X plus something positive (even if it is very small), which is apparently greater than X from the certain option.

B Additional Figures

Figure B.1: Prelec's Function for diverse α

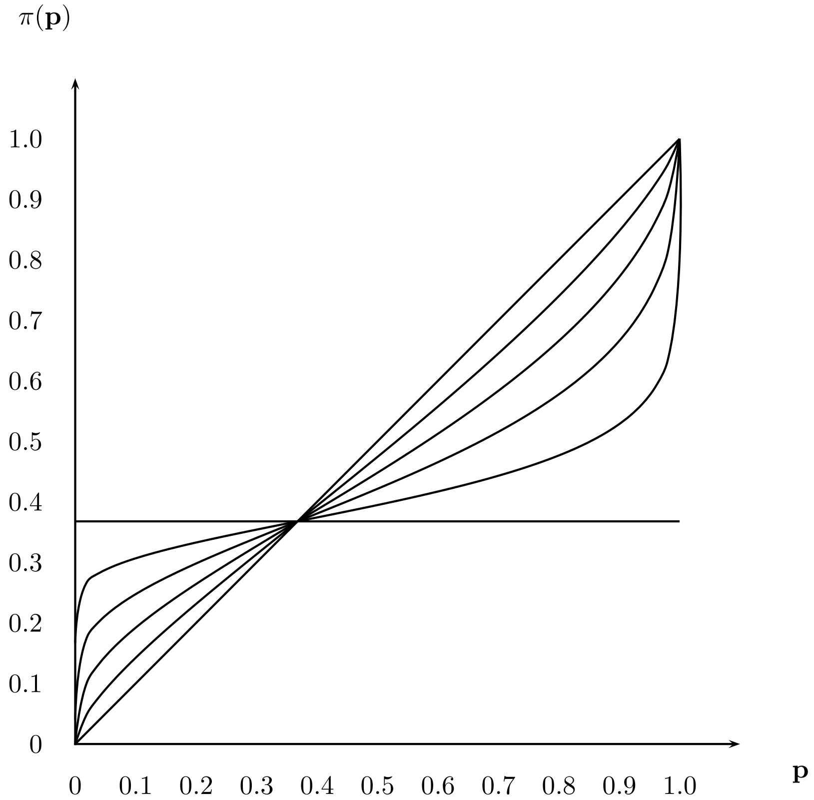


Figure B.2: Probability Function with $\alpha = -0.5$

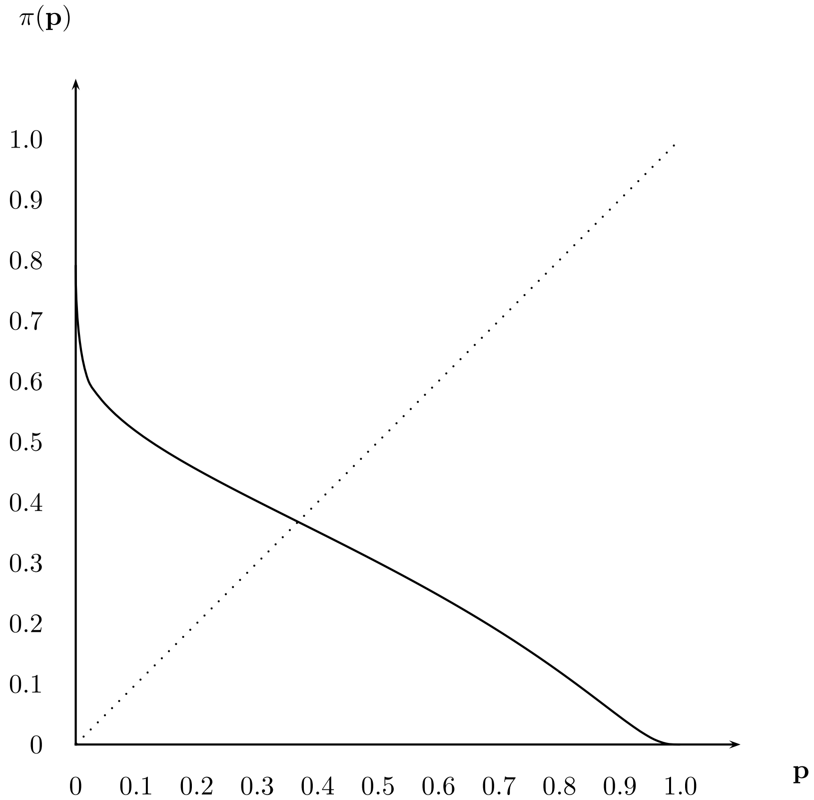
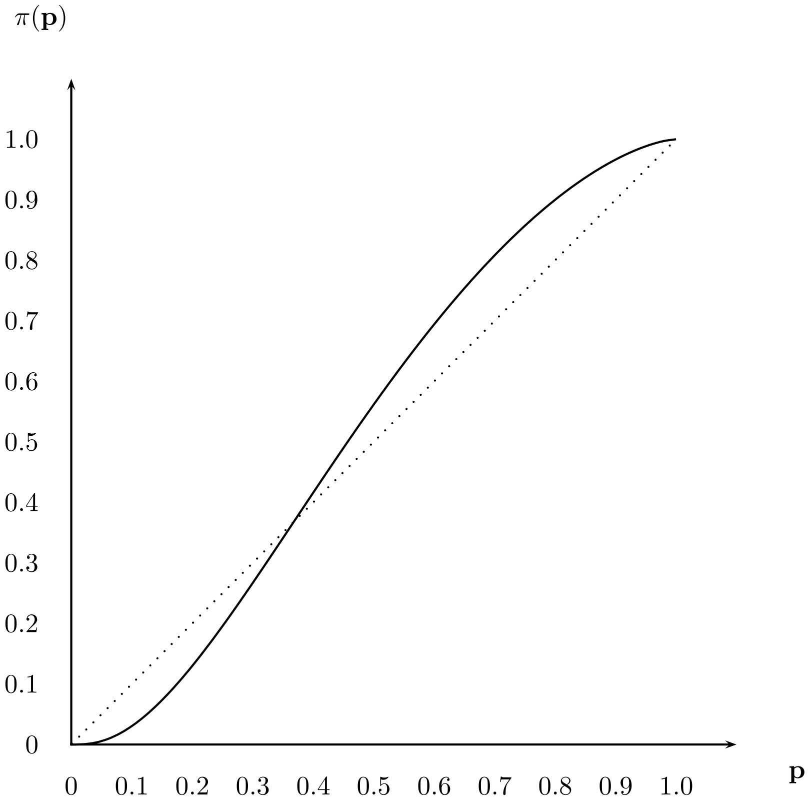


Figure B.3: Probability Function with $\alpha = 1.5$



C Additional Tables

Table C.1: Non-Expected Utility Models

Prospect Theory	Ward Edwards (1955, 1962), Kahneman and Tversky (1979)
Subjectively Weighted Utility	Uday Karmarkar (1978, 1979)
Weighted Utility	Chew and MacCrimmon (1979), Chew (1983), Fishburn (1983)
Anticipated Utility	John Quiggin (1982)
Optimism/Pessimism	John Hey (1984)
Ordinal Independence	Segal (1984), Green and Jullien (1988)
General Quadratic	Chew, Epstein and Segal (1991)
Cumulative Prospect Theory	Kahneman and Tversky (1992)
3rd Generation Prospect Theory	Schmidt, Starmer and Sugden (2008)

This list, except for the last two, is from Machina (1989).

Table C.2: Three educational categories

low education	inadequately completed general elementary school basic vocational qualification	kein Abschluss Hauptschulabschluss ohne berufliche Ausbildung Hauptschulabschluss mit beruflicher Ausb.
middle education	intermediate vocational intermediate general qualification general maturity certificate vocational maturity certificate	Mittlere Reife ohne berufliche Ausbildung Mittlere Reife mit beruflicher Ausbildung Fachabitur/Abitur ohne berufliche Ausbildung Fachabitur/Abitur mit beruflicher Ausbildung
high education	lower tertiary education higher tertiary education	Fachhochschulabschluss Hochschulabsschluss

Table C.3: Parameter estimates for income error term

	value	stdx
$constant$	9.3965	(0.1418)
age	0.0505	(0.0061)
$age^2/100$	-0.0546	(0.006)
$kids$	0.036	(0.0066)
$dep.adults$	0.1455	(0.0069)
$unemploymentrate$	0.071	(0.004)
$cohort2$	-0.3063	(0.0519)
$cohort3$	-0.2674	(0.061)
$cohort4$	-0.4189	(0.0681)
$cohort5$	-0.3109	(0.0706)
$cohort6$	-0.3139	(0.0747)

Table C.4: Wealth to Income ratios with cohorts

	value	stdx
constant	9.9300	(0.9740)
kids	0.9813	(0.7972)
dep.adults	1.5269	(1.2649)
cohort2	-2.7693	(1.2206)
cohort3	-3.1256	(1.1874)
cohort4	-7.5563	(1.3439)
cohort5	-8.1138	(1.3981)
cohort6	-7.5333	(1.9017)

Table C.5: Parameter Vector χ - Part I

	value	stdx
coefficient of relative riskaversion ρ	2	-
consumption flow γ	0.05	-
discount rate δ	0.95	-
credit limit λ	0.318	(0.017)
interest rate pos. liquid assets R_X	1.0241	(0.0135)
interest rate neg. liquid assets R_X^{CC}	1.1	-
Kids estimation		
β_1	0.25	(0.002)
β_2	0.89	(0.321)
β_3	0.12	(0.102)
Dependent Adults estimation		
β_1	0.34	(0.055)
β_2	0.1	(0.03)
β_3	0.44	(0.78)

Table C.6: Parameter Vector χ - Part II

	value	stdx
income working		
$constant$	9.3965	(0.1418)
age	0.0505	(0.0061)
$age^2/100$	-0.0546	(0.006)
$kids$	0.036	(0.0066)
$dep.adults$	0.1455	(0.0069)
$unemploymentrate$	0.071	(0.004)
$cohort2$	-0.3063	(0.0519)
$cohort3$	-0.2674	(0.061)
$cohort4$	-0.4189	(0.0681)
$cohort5$	-0.3109	(0.0706)
$cohort6$	-0.3139	(0.0747)
income error term		
α	0.7425	(0.2981)
σ_ϵ^2	0.023	(0.0081)
σ_v^2	0.0129	(0.0057)
income retired[a]		
mean	10.017	(0.0637)

[a] For income in retirement both, the mean and the standard deviation, are used for the simulation.

D Proof: Rank-Dependent Utility fits the Alternative Stochastic Dominance Definition

In contrast to the example in the text we use a more general setting. In particular, we replace the given probabilities by s for the probability of survival, m for the probability to end up in the high income class, and p for the chance of receiving a positive income shock:

$$L_a = \left(\left[(A; p|B; 1-p); m|(C; p|D; 1-p); 1-m \right]; s|B_u; 1-s \right)$$

$$L_b = \left(\left[(E; p|F; 1-p); m|(G; p|H; 1-p); 1-m \right]; s|B_d; 1-s \right).$$

Recall, the two simple lotteries in squared brackets are already ordered, i.e. the value of the left lottery is bigger (high income class) than the value of the right lottery (low income class). The definition of stochastic dominance for the three-stage lotteries arising in our model states that L_a stochastically dominates L_b if the following conditions are true:

C-1: (A;p | B;1-p) first-order-stochastically dominates (E;p | F;1-p)

C-2: (C;p | D;1-p) first-order-stochastically dominates (G;p | H;1-p)

C-3: B_u first-order-stochastically dominates B_d

and L_a strictly stochastically dominates L_b, if at least one of these conditions satisfies strict First-Order-Stochastic Dominance.

It is known that Rank-Dependent Utility satisfies First-Order-Stochastic Dominance - see for example Machina (1989). Hence, the conditions C-1 and C-2 imply that V_{AB}, the value of lottery (A;p | B;1-p), is at least as big as V_{EF}, or for strict First-Order-Stochastic Dominance bigger than V_{EF} - the same holds for V_{CD} and V_{GH}. Now, consider the two-stage lotteries in the squared brackets [(A;p | B;1-p);m | (C;p | D;1-p);1-m] and [(E;p | F;1-p);m | (G;p | H;1-p);1-m], and label these W and Q, respectively. There are two possible relations, given the conditions above, for the two-stage lotteries W and Q according to the definition of Two-Stage-Stochastic Dominance - ("fosd" is for first-order-stochastically dominates, "tssd" for two-stage-stochastically dominates, and V_W and V_Q are the values of W and Q from Rank-Dependent Utility provided by Folding

Back.):

- W strictly tssd Q, if
  - (A;p | B;1-p) strictly fosd (E;p | F;1-p) and (C;p | D;1-p) strictly fosd (G;p | H;1-p) or
  - (A;p | B;1-p) strictly fosd (E;p | F;1-p) and (C;p | D;1-p) fosd (G;p | H;1-p) or
  - (A;p | B;1-p) fosd (E;p | F;1-p) and (C;p | D;1-p) strictly fosd (G;p | H;1-p)

  all implying $V_W > V_Q$.
- W tssd Q, if (A;p | B;1-p) fosd (E;p | F;1-p) and (C;p | D;1-p) fosd (G;p | H;1-p) - implying $V_W \geq V_Q$.

The implications for the relation of V_W and V_Q are simply derived from the combinations of first-order-stochastic dominance of the two lottery pairs. Remember, the simple lotteries are already ordered, then according to Rank-Dependent Utility the values of the two-stage lotteries W and Q (with V(A;p | B;1-p) being the value of lottery (A;p | B;1-p) etc.) are:

$$V_W = (1 - \pi(m)) \cdot V(C; p|D; 1 - p) + \pi(m) \cdot V(A; p|B; 1 - p)$$

$$V_Q = (1 - \pi(m)) \cdot V(G; p|H; 1 - p) + \pi(m) \cdot V(E; p|F; 1 - p).$$

The weights from the cumulative weighting functions are identical for both equations and the relations between the simple lotteries are clarified by the conditions C-1 and C-2 from the definition. That is, if L_a stochastically dominates L_b, then by condition C-2 $V(C; p|D; 1 - p) \geq V(G; p|H; 1 - p)$ and by condition C-1 $V(A; p|B; 1 - p) \geq V(E; p|F; 1 - p)$. This implies that

$$V_W \geq V_Q. \tag{D.1}$$

Now, consider the remaining first stage of the lotteries L_a and L_b:

$$L_a = (V_W; s|B_u; 1-s)$$

$$L_b = (V_Q; s|B_d; 1-s)\,.$$

From condition 3 it is known that $B_u \geq B_d$, which implies $V_{B_u} \geq V_{B_d}$. According to the ordering required for the evaluation with Rank-Dependent Utility preferences there are two cases for each lottery L_a and L_b - see Table D.1.

Table D.1: Possible Rankings

L_a	L_b
$a_1 : V_W \geq V_{B_u}$	$b_1 : V_Q \geq V_{B_d}$
$a_2 : V_W < V_{B_u}$	$b_2 : V_Q < V_{B_d}$

Altogether 4 possible ranking combinations for the comparison of the two lotteries under Rank-Dependent Utility can occur - (a_1, b_1), (a_1, b_2), (a_2, b_1) and (a_2, b_2). In the setting of the model there are never more than these 4 possibilities, as the first-stage of the three-stage lotteries is the risk of either survival or death of the household.

Now, it is necessary to show that each evaluation by Rank-Dependent Utility of these possibilities yields $V_{L_a} \geq V_{L_b}$, as it was assumed that L_a stochastically dominates L_b. In order to proof this, assume the opposite, i.e. $V_{L_a} < V_{L_b}$ and check if this can be true. Note, in some cases it can be that V_W is equal to V_{B_u} and/or V_Q is equal to V_{B_d}. For these constellations the outcomes are ranked in the evaluation, although they are equal, as it suits best for the proof. This is permitted as both outcomes in the lottery are equal and the cumulative probabilities by definition sum up to 1.

Case 1: $V_W \geq V_{B_u}$ and $V_Q \geq V_{B_d}$ - Given these rankings evaluate the lotteries L_a and L_b by Rank-Dependent Utility:

$$(1-\pi(s)) \cdot V_{B_u} + \pi(s) \cdot V_W < (1-\pi(s)) \cdot V_{B_d} + \pi(s) \cdot V_Q.$$

The cumulative weighting functions are identical on both sides, therefore one can rewrite the inequality. This leads to

$$(1-\pi(s))\cdot(V_{B_u}-V_{B_d})+\pi(s)\cdot(V_W-V_Q)<0.$$

It was shown above that $V_W \geq V_Q$ by equation E.1. C-3 in the definition states that $V_{B_u} \geq V_{B_d}$ and the definition of the probability weighting function determines that $\pi(s) \in [0;1]$ for $s \in [0;1]$. Hence, the above inequality can not be true.

Case 2: $V_W \geq V_{B_u}$ and $V_Q < V_{B_d}$

$$(1-\pi(s))\cdot V_{B_u}+\pi(s)\cdot V_W<(1-\pi(1-s))\cdot V_Q+\pi(1-s)\cdot V_{B_d}.$$

One can rewrite by the assumptions of this case - $V_W = V_{B_u}+\epsilon$ and $V_{B_d} = V_Q+\eta$, with $\epsilon, \eta \geq 0$. This yields

$$\begin{aligned}
&(1-\pi(s))\cdot V_{B_u}+\pi(s)\cdot(V_{B_u}+\epsilon)<(1-\pi(1-s))\cdot V_Q+\pi(1-s)\cdot(V_Q+\eta)\\
\Rightarrow &V_{B_u}-\pi(s)\cdot V_{B_u}+\pi(s)\cdot V_{B_u}+\pi(s)\cdot\epsilon\\
&<V_Q-\pi(1-s)\cdot V_Q+\pi(1-s)\cdot V_Q+\pi(1-s)\cdot\eta\\
\Rightarrow &V_{B_u}+\pi(s)\cdot\epsilon<V_Q+\pi(1-s)\cdot\eta\\
\Rightarrow &\pi(s)\cdot\epsilon<V_Q+\pi(1-s)\cdot\eta-V_{B_u}.
\end{aligned}$$

We know $\pi(s)\cdot\epsilon \geq 0$, as $\pi(s) \in [0;1]$ for $s \in [0;1]$ and $\epsilon \geq 0$. For the right hand side it holds that $V_Q + \pi(1-s)\cdot\eta \leq V_{B_d}$, because of the substitution $V_{B_d} = V_Q + \eta$ and $\pi(s) \in [0;1]$. In combination with the implication from C-3, namely $V_{B_u} \geq V_{B_d}$, this leads to the conclusion that $V_Q + \pi(1-s)\cdot\eta - V_{B_u} \leq 0$. Hence, the resulting inequality

$$0 \leq \pi(s)\cdot\epsilon < V_Q+\pi(1-s)\cdot\eta-V_{B_u} \leq 0$$

is not true.

Case 3: $V_W < V_{B_u}$ and $V_Q \geq V_{B_d}$

$$(1-\pi(1-s))\cdot V_W+\pi(s)\cdot V_{B_u}<(1-\pi(s))\cdot V_{B_d}+\pi(s)\cdot V_Q.$$

Similar to Case 2 set $V_{B_u} = V_W + \epsilon$ and $V_Q = V_{B_d} + \eta$, with $\epsilon, \eta \geq 0$:

$$
\begin{aligned}
&(1-\pi(1-s)) \cdot V_W + \pi(1-s) \cdot (V_W + \epsilon) < (1-\pi(s)) \cdot V_{B_d} + \pi(s) \cdot (V_{B_d} + \eta) \\
\Rightarrow & V_W - \pi(1-s) \cdot V_W + \pi(1-s) \cdot V_W + \pi(1-s) \cdot \epsilon \\
&< V_{B_d} - \pi(s) \cdot V_{B_d} + \pi(s) \cdot V_{B_d} + \pi(s) \cdot \eta \\
\Rightarrow & V_W + \pi(1-s) \cdot \epsilon < V_{B_d} + \pi(s) \cdot \eta \\
\Rightarrow & \pi(1-s) \cdot \epsilon < V_{B_d} + \pi(s) \cdot \eta - V_W.
\end{aligned}
$$

For the left hand side holds again $\pi(1-s) \cdot \epsilon \geq 0$, as $\pi(1-s) \in [0;1]$ for $s \in [0;1]$ and $\epsilon \geq 0$. From $\pi(s) \in [0;1]$ and the substitution $V_Q = V_{B_d} + \eta$ it follows that $V_{B_d} + \pi(s) \cdot \eta \leq V_Q$. Together with equation E.1 one gets the upcoming inequality, which is obviously not true

$$0 \leq \pi(1-s) \cdot \epsilon < V_{B_d} + \pi(s) \cdot \eta - V_W \leq 0.$$

Case 4: $V_W < V_{B_u}$ and $V_Q < V_{B_d}$

$$(1-\pi(1-s)) \cdot V_W + \pi(1-s) \cdot V_{B_u} < (1-\pi(1-s)) \cdot V_Q + \pi(1-s) \cdot V_{B_d}.$$

Like in Case 1 the cumulative weighting functions are identical - yielding

$$(1-\pi(1-s)) \cdot (V_W - V_Q) + \pi(1-s) \cdot (V_{B_u} - V_{B_d}) < 0.$$

From Equation E.1 it is known that $V_W \geq V_Q$. Condition C-3 implies that $V_{B_u} \geq V_{B_d}$ and it is known that $\pi(1-s) \in [0;1]$ for $s \in [0;1]$. Therefore, the above inequality can not be true.

Conclusion: The cases 1-4 show that Rank-Dependent Utility preferences select the three-stage lottery, which is stochastically dominant, at least by the definition introduced in this thesis.

E Rank-Dependent Utility and the Independence Axiom

Let L_X, L_Y and L_Z be the values of the lotteries X, Y and Z. The preference $X \succ (\succeq) Y$ implies for the evaluation with Rank-Dependent Utility that $L_X > (\geq) L_Y$. The probability function is denoted by π and exhibits the same features as shown in Section 3. Note, the evaluation process applied here does not use the Reduction of Compound Lotteries Axiom, but Folding Back.

Part 1: The first statement of the Independence Axiom is

$$X \succ Y \Leftrightarrow (X; p|Z; 1-p) \succ (Y; p|Z; 1-p). \tag{E.1}$$

To show that the above equation holds in general under Rank-Dependent Utility preferences, we first consider the left to right direction of statement E.1, namely

$$X \succ Y \Rightarrow (X; p|Z; 1-p) \succ (Y; p|Z; 1-p),$$

and check if this is true.

There are five cases for the ranking of the three lotteries X, Y and Z, under the condition $X \succ Y$. Some of these cases can be summarized to one case, as these share the same patterns in the evaluation process

1. $Z \prec Y \prec X$ or $Y \prec X \prec Z$
2. $Y \sim Z \prec X$ or $Y \prec Z \sim X$
3. $Y \prec Z \prec X$.

Case 1: The rankings in the two-stage lotteries $(X; p|Z; 1-p)$ and $(Y; p|Z; 1-p)$ are $L_Z < L_X$ and $L_Z < L_Y$. This leads under Rank-Dependent Utility to

$$(1-\pi(p)) \cdot L_Z + \pi(p) \cdot L_X > (1-\pi(p)) \cdot L_Z + \pi(p) \cdot L_Y$$
$$L_X > L_Y.$$

The above inequation is true, as the preference $X \succ Y$ implies $L_X > L_Y$. The argumentation for the case $Y \prec X \prec Z$ is in principle the same.

Case 2: The ranking in the two-stage lottery $(X;p|Z;1-p)$ is $L_Z < L_X$. For the lottery $(Y;p|Z;1-p)$ there is actually no ranking, as $Y \sim Z$ implies $L_Y = L_Z$. Recall the intuition of the evaluation procedure under Rank-Dependent Utility preferences for each outcome x - the cumulative weighting function for x is defined by the difference of *"the chance of winning an outcome at least as good as x"* and *"the chance of winning an outcome strictly better than x"*. This implies for $(Y;p|Z;1-p)$ that *"the chance of winning an outcome at least as good as Z"* is 1 and *"the chance of winning an outcome strictly better than Z"* is 0 - the same holds if one starts with Y. This implies that the lottery $(Y;p|Z;1-p)$ can be written as $(Y;1)$ or $(Z;1)$, because of the indifference of Y and Z, which also implies that $L_Y = L_Z$. For the expression $(X;p|Z;1-p) \succ (Y;p|Z;1-p)$ this yields

$$
\begin{aligned}
(1-\pi(p)) \cdot L_Y + \pi(p) \cdot L_X &> (1-\pi(0)) \cdot L_Y \\
(1-\pi(p)) \cdot L_Y + \pi(p) \cdot L_X &> L_Y \\
\pi(p) \cdot L_X - \pi(p) \cdot L_Y &> 0 \\
L_X &> L_Y.
\end{aligned}
$$

According to the ranking above, we know that this inequation is true. The same argument holds for $Y \prec Z \sim X$.

Case 3: This case is responsible for the fact that Rank-Dependent Utility is not compatible with the Independence Axiom, if the Reduction of Compound Lotteries Axiom is used. The rankings in the two-stage lotteries $(X;p|Z;1-p)$ and $(Y;p|Z;1-p)$ are $L_Z < L_X$ and $L_Z > L_Y$, respectively. With Folding Back and without the Reduction of Compound Lotteries Axiom this leads to

$$
\begin{aligned}
(1-\pi(p)) \cdot L_Z + \pi(p) \cdot L_X &> (1-\pi(1-p)) \cdot L_Y + \pi(1-p) \cdot L_Z \\
L_Z - \pi(p) \cdot L_Z + \pi(p) \cdot L_X &> L_Y - \pi(1-p) \cdot L_Y + \pi(1-p) \cdot L_Z \\
L_Z + \pi(p) \cdot (L_X - L_Z) &> L_Y - \pi(1-p) \cdot (L_Y - L_Z) \\
\pi(p) \cdot (L_X - L_Z) &> L_Y - L_Z - \pi(1-p) \cdot (L_Y - L_Z) \\
\pi(p) \cdot (L_X - L_Z) &> (1-\pi(1-p)) \cdot (L_Y - L_Z).
\end{aligned}
$$

The left hand side is greater than 0, as $0 < \pi(p) < 1$ and $L_X - L_Z > 0$. It also holds that $0 < \pi(1-p) < 1$ for all $p \in]0;1[$. This implies $(1-\pi(1-p)) > 0$. The above assumed preference $Y \prec Z$ results in $L_Y < L_Z$ or $L_Y - L_Z < 0$. Hence, the right hand side is negative and therefore the above inequation is true.

Now, we consider the other direction of statement E.1, namely

$$X \succ Y \Leftarrow (X; p|Z; 1-p) \succ (Y; p|Z; 1-p), \tag{E.2}$$

and check if this is also true. The difference to the approach above is the information one starts with, i.e. one knows it holds that

$$(X; p|Z; 1-p) \succ (Y; p|Z; 1-p). \tag{E.3}$$

In order to evaluate the compound lotteries in E.3, we have to make assumptions about the relation of X and Y, although we do not know which is true. There are three possibilities, namely

A: $X \prec Y$

B : $X \sim Y$

C: $X \succ Y$.

Now, we check whether the preference E.3, which we know is true, holds under A or B. For the case that neither A nor B is the correct assumption, then we know that $X \succ Y$ must be true.

Case A: For $X \prec Y$ we get the following three ranking cases (with different patterns in the evaluation) for the three lotteries X, Y and Z, namely

1. $X \prec Y \prec Z$ or $Z \prec X \prec Y$
2. $X \sim Z \prec Y$ or $X \prec Z \sim Y$
3. $X \prec Z \prec Y$.

Case 1: Here, we again only consider $X \prec Y \prec Z$, as the other case has the same patterns. Evaluating E.3 according to Rank-Dependent Utility yields

$$(1-\pi(1-p)) \cdot L_X + \pi(1-p) \cdot L_Z > (1-\pi(1-p)) \cdot L_Y + \pi(1-p) \cdot L_Z$$
$$L_X > L_Y.$$

We know that the above inequation must hold, therefore the assumption $X \prec Y$ can not be true in this case.

Case 2: Here, we also consider only $X \sim Z \prec Y$. The relation $X \sim Z$ implies that $L_X = L_Z$, which yields

$$\begin{aligned}
L_X &> (1-\pi(p)) \cdot L_X + \pi(p) \cdot L_Y \\
L_X &> L_X - \pi(p) \cdot L_X + \pi(p) \cdot L_Y \\
L_X &> L_Y.
\end{aligned}$$

Again, we know that the above inequation must hold, therefore the assumption $X \prec Y$ can not be true in this case.

Case 3:

$$\begin{aligned}
(1-\pi(1-p)) \cdot L_X + \pi(1-p) \cdot L_Z &> (1-\pi(p)) \cdot L_Z + \pi(p) \cdot L_Y \\
L_X - \pi(1-p) \cdot L_X + \pi(1-p) \cdot L_Z &> L_Z - \pi(p) \cdot L_Z + \pi(p) \cdot L_Y \\
L_X - L_Z - \pi(1-p) \cdot L_X + \pi(1-p) \cdot L_Z &> \pi(p) \cdot (L_Y - L_Z) \\
L_X - L_Z - \pi(1-p) \cdot (L_X - L_Z) &> \pi(p) \cdot (L_Y - L_Z) \\
(1-\pi(1-p)) \cdot (L_X - L_Z) &> \pi(p) \cdot (L_Y - L_Z).
\end{aligned}$$

We know that $(1-\pi(1-p)) > 0$ and that $(L_X - L_Z) < 0$ by the ranking assumption, therefore the left hand side is smaller than 0. The right hand side is positive, as $\pi(p) > 0$ and $(L_Y - L_Z) > 0$. This implies that the above inequation can not be true under the assumptions of this case. However, we know that the inequation is true, therefore we get to the conclusion that the assumption $X < Y$ is not correct in this case

Conclusion Case A: All three subcases come up with wrong expressions, therefore the assumption $X < Y$ for the evaluation of the compound lotteries can not be true, if preference E.3 is true.

Case B: There are three different possibilities for the ranking of the three lotteries, which can be summarized in the two cases

1. $X \sim Y \prec Z$ or $Z \prec X \sim Y$
2. $X \sim Y \sim Z$.

Case 1: We consider $X \sim Y \prec Z$, which yields

$$(1-\pi(1-p)) \cdot L_X + \pi(1-p) \cdot L_Z > (1-\pi(1-p)) \cdot L_X + \pi(1-p) \cdot L_Z$$
$$L_X > L_Y.$$

The assumption $X \sim Y$ contradicts the above inequation and therefore can not be true.

Case 2: The indifference implies that we can write $(X;p|Z;1-p)$ as $(X;1)$, and $(X;p|Z;1-p)$ as $(Y;1)$. This leads to

$$L_X > L_Y.$$

This implies that $X \sim Y$ can not be true.

Conclusion Case B: In both cases we find contradictions, that is why we conclude, as above, that the assumption $X \sim Y$ can not be true, as long as E.3 is true.

Conclusion: We showed that the assumptions $X \prec Y$ and $X \sim Y$ disagree with preference E.3, which we know is true. We also know that one of the assumptions A, B and C must hold. Due to this we conclude that $X \succ Y$ must be true if $(X;p|Z;1-p) \succ (Y;p|Z;1-p)$ prevails. This implies that statement E.2 must be true.

Part 2: The other part of the Independence Axiom states that

$$X \sim Y \Leftrightarrow (X;p|Z;1-p) \sim (Y;p|Z;1-p). \tag{E.4}$$

We again start with the left to right direction, namely

$$X \sim Y \Rightarrow (X;p|Z;1-p) \sim (Y;p|Z;1-p).$$

Under the condition that X is indifferent to Y, there are two possibilities with different patterns for the ordering of all three lotteries, namely

1. $X \sim Y \prec Z$ or $Z \prec X \sim Y$.
2. $X \sim Y \sim Z$

Case 1: We only consider the case $X \sim Y \prec Z$. This leads to

$$(1-\pi(1-p)) \cdot L_X + \pi(1-p) \cdot L_Z = (1-\pi(1-p)) \cdot L_Y + \pi(1-p) \cdot L_Z$$
$$\pi(1-p) \cdot L_X = \pi(1-p) \cdot L_Y.$$

The equality holds, as $X \sim Y$ implies that $L_X = L_Y$.

Case 2: The indifference between all three lotteries implies that $(X; p|Z; 1-p)$ can be written as $(X; 1)$, and $(Y; p|Z; 1-p)$ as $(Y; 1)$. This leads to

$$L_X = L_Y.$$

This implies that $X \sim Y$, which is the initial condition.

The proof in the other direction of E.4 is in principle the same, as in Part 1. We know that

$$(X; p|Z; 1-p) \sim (Y; p|Z; 1-p) \tag{E.5}$$

is true. In order to evaluate the compound lotteries, we have to make assumptions about the relation of X and Y. As above there the three possibilities

A: $X \prec Y$

B : $X \succ Y$

C: $X \sim Y$.

One has to check if A and B can be true under E.5, as we are interested whether $X \sim Y$ is the correct assumption.

Case A: We have the rankings

1. $X \prec Y \prec Z$ or $Z \prec X \prec Y$
2. $X \sim Z \prec Y$ or $X \prec Z \sim Y$
3. $X \prec Z \prec Y$.

Case 1: We consider $X \prec Y \prec Z$. Evaluating E.5 according to Rank-Dependent

Utility yields

$$(1-\pi(1-p))\cdot L_X+\pi(1-p)\cdot L_Z\sim(1-\pi(1-p))\cdot L_Y+\pi(1-p)\cdot L_Z$$
$$L_X\sim L_Y.$$

We know that the above inequation must hold, therefore the assumption $X \prec Y$ can not be true in this case.

Case 2: We consider $X \sim Z \prec Y$. The relation $X \sim Z$ implies that $L_X = L_Z$, which yields

$$\begin{aligned}
&L_X\sim(1-\pi(p))\cdot L_X+\pi(p)\cdot L_Y\\
&L_X\sim L_X-\pi(p)\cdot L_X+\pi(p)\cdot L_Y\\
&L_X\sim L_Y.
\end{aligned}$$

Again, we know that the above inequation must hold, therefore the assumption $X \prec Y$ can not be true in this case.

Case 3:

$$\begin{aligned}
(1-\pi(1-p))\cdot L_X+\pi(1-p)\cdot L_Z&\sim(1-\pi(p))\cdot L_Z+\pi(p)\cdot L_Y\\
L_X-\pi(1-p)\cdot L_X+\pi(1-p)\cdot L_Z&\sim L_Z-\pi(p)\cdot L_Z+\pi(p)\cdot L_Y\\
L_X-L_Z-\pi(1-p)\cdot L_X+\pi(1-p)\cdot L_Z&\sim\pi(p)\cdot(L_Y-L_Z)\\
L_X-L_Z-\pi(1-p)\cdot(L_X-L_Z)&\sim\pi(p)\cdot(L_Y-L_Z)\\
(1-\pi(1-p))\cdot(L_X-L_Z)&\sim\pi(p)\cdot(L_Y-L_Z)
\end{aligned}$$

We know that $(1-\pi(1-p)) > 0$ and that $(L_X - L_Z) < 0$ by the ranking assumption, therefore the left hand side is smaller than 0. The right hand side is positive, as $\pi(p) > 0$ and $(L_Y - L_Z) > 0$. This implies that the above indifference does not hold under the assumption of this case. However, we know that it holds, therefore we conclude that $X \prec Y$ can not be correct in this case

Conclusion Case A: We have seen that the assumption $X \prec Y$ can not be true if indifference E.5 is true.

Case B: The rankings are

1. $Z \prec Y \prec X$ or $Y \prec X \prec Z$

2. $Y \sim Z \prec X$ or $Y \prec Z \sim X$

3. $Y \prec Z \prec X$.

Case 1: We consider $Z \prec Y \prec X$. This yields

$$\begin{aligned} (1-\pi(p)) \cdot L_Z + \pi(p) \cdot L_X &\sim (1-\pi(p)) \cdot L_Z + \pi(p) \cdot L_Y \\ L_X &\sim L_Y. \end{aligned}$$

The assumption $X \succ Y$ contradicts this indifference and therefore can not be true in this case.

Case 2: We consider $Y \sim Z \prec X$. Recall, the indifference implies that $L_Y = L_Z$ and yields

$$\begin{aligned} (1-\pi(p)) \cdot L_Y + \pi(p) \cdot L_X &\sim L_Y \\ (1-\pi(p)) \cdot L_Y + \pi(p) \cdot L_X &\sim L_Y \\ \pi(p) \cdot L_X - \pi(p) \cdot L_Y &\sim 0 \\ L_X &\sim L_Y. \end{aligned}$$

This indifference is not satisfied under the assumption $X \succ Y$. This implies that the assumption is also not true in this case.

Case 3:

$$\begin{aligned} (1-\pi(p)) \cdot L_Z + \pi(p) \cdot L_X &\sim (1-\pi(1-p)) \cdot L_Y + \pi(1-p) \cdot L_Z \\ L_Z - \pi(p) \cdot L_Z + \pi(p) \cdot L_X &\sim L_Y - \pi(1-p) \cdot L_Y + \pi(1-p) \cdot L_Z \\ L_Z + \pi(p) \cdot (L_X - L_Z) &\sim L_Y - \pi(1-p) \cdot (L_Y - L_Z) \\ \pi(p) \cdot (L_X - L_Z) &\sim L_Y - L_Z - \pi(1-p) \cdot (L_Y - L_Z) \\ \pi(p) \cdot (L_X - L_Z) &\sim (1-\pi(1-p)) \cdot (L_Y - L_Z). \end{aligned}$$

The left hand side is greater than 0, as $0 < \pi(p) < 1$ and $L_X - L_Z > 0$. We know that $(1-\pi(1-p)) > 0$. The above assumed preference $Y \prec Z$ results in $L_Y < L_Z$ or $L_Y - L_Z < 0$. Hence, the right hand side is negative and therefore the above indifference is not true under the assumed relation. However, we know that the indifference is true. Due to this we conclude that the assumption $X \succ Y$ can not be true here.

Conclusion Case C: We get contradictions for all three rankings of Case B. This implies that the assumption X can not be true under the indifference E.5.

Conclusion: We showed that under both assumptions A and B the indifference $(X;p|Z;1-p) \sim (Y;p|Z;1-p)$ does not hold, although we know that it is true. Due to these findings we conclude that the correct assumption in this case is $X \sim Y$. This implies that the statement

$$X \sim Y \Leftarrow (X;p|Z;1-p) \sim (Y;p|Z;1-p)$$

is true.

Summary: Part 1 and 2 with the corresponding subcases show that Rank-Dependent Utility satisfies the Independence Axiom, if Folding Back is applied instead of the Reduction of Compound Lotteries Axiom. However, the proof considers solely lotteries with two lotteries as outcomes, but it can be shown with the same procedure that this is also true for lotteries with three, four etc. lotteries as outcomes.

F Empirical Moments for GMM Estimation of the Income AR(1) Error Process

The error in t of the income estimation consists of an individual effect v_i, an AR(1) process and a transitory shock ν_{it}:

$$e_{it} = v_i + u_{it} + \nu_{it} = v_i + \alpha u_{it-1} + \epsilon_{it} + \nu_{it}.$$

In matrix notation for all $i = 1, ..., N$ this reduces to

$$e_t = v + u_t + \nu_t = v + \alpha u_{t-1} + \epsilon_t + \nu_t,$$

where v, u, ν and ϵ are all column vectors of the size $N \times 1$ and α is a scalar. For the upcoming transformation define σ_ν^2 as the variance of the transitory shock ν, and σ_ϵ^2 as the variance of ϵ - both are independent and identically distributed random variables. The parameters σ_v^2, $\sigma_{\stackrel{2}{\epsilon}}$ and α are estimated via GMM (Generalized Method of Moments) by minimizing the objective function

$$\begin{pmatrix} Cov(\Delta e_t, \Delta e_t) - C_0 \\ Cov(\Delta e_t, \Delta e_{t-1}) - C_1 \\ Cov(\Delta e_t, \Delta e_{t-2}) - C_2 \\ . \\ . \\ . \\ Cov(\Delta e_t, \Delta e_{t-k}) - C_k \end{pmatrix}' * W * \begin{pmatrix} Cov(\Delta e_t, \Delta e_t) - C_0 \\ Cov(\Delta e_t, \Delta e_{t-1}) - C_1 \\ Cov(\Delta e_t, \Delta e_{t-2}) - C_2 \\ . \\ . \\ . \\ Cov(\Delta e_t, \Delta e_{t-k}) - C_k \end{pmatrix}.$$

The theoretical covariances C_k for $k = 0, ..., 6$, which are used in the estimation, are obtained by some transformations. Here, the example for $Cov(\Delta e_t, \Delta e_{t-1})$ is considered - the remaining covariances follow in principal the same transformations:

$$\begin{aligned} C_1 &= Cov(\Delta e_t, \Delta e_{t-1}) \\ &= E\left[(\Delta e_t - E(\Delta e_t))(\Delta e_{t-1} - E(\Delta e_{t-1}))'\right]. \end{aligned}$$

In order to simplify the above equation regard the expectation of Δe_t:

$$
\begin{aligned}
E(\Delta e_t) &= E(v + u_t + \nu_t - v - u_{t-1} - \nu_{t-1}) \\
&= E(\alpha u_{t-1} + \epsilon_t + \nu_t - u_{t-1} - \nu_{t-1}) \\
&= (\alpha - 1)E(u_{t-1}) + E(\nu_t) - E(\nu_{t-1}).
\end{aligned}
$$

The assumption that ν is normally distributed implies that $E(\nu_t) = E(\nu_{t-1}) = 0$ etc.. This leads to

$$
E(\Delta e_t) = (\alpha - 1)E(u_{t-1}).
$$

Now, consider the expectation of u_{t-1} and assume as a starting value for the AR(1) process $u_{t-n} = 0$ with $n \to \infty$ (this assumption states that u_{t-n} is no longer a vector of random variables and implies that $E(u_{t-n}) = 0$) and recall that $E(\epsilon_t) = 0$:

$$
\begin{aligned}
E(u_{t-1}) = E(\alpha u_{t-2} + \epsilon_{t-1}) &= E(\alpha(\alpha u_{t-3} + \epsilon_{t-2}) + \epsilon_{t-1}) \\
= \ldots &= E\left(\alpha^{n-1} u_{t-n} + \sum_{j=1}^{n-1} \alpha^{j-1} \epsilon_{t-j}\right) \\
&= \alpha^{n-1} E(u_{t-n}) + \sum_{j=1}^{n-1} \alpha^{j-1} E(\epsilon_{t-j}) = 0.
\end{aligned}
$$

Hence, the definition of the covariance C_1 is given by:

$$
\begin{aligned}
C_1 &= E(\Delta e_t \Delta e'_{t-1}) \\
&= E((v + u_t + \nu_t - v - u_{t-1} - \nu_{t-1})(u_{t-1} + \nu_{t-1} - u_{t-2} - \nu_{t-2})') \\
&= E\left[((\alpha - 1)u_{t-1} + \epsilon_t + \nu_t - \nu_{t-1})((\alpha - 1) \cdot u_{t-2} + \epsilon_{t-1} + \nu_{t-1} - \nu_{t-2})'\right] \\
&= E[(\alpha - 1)^2 u_{t-1} u'_{t-2} + (\alpha - 1) u_{t-1} \epsilon'_{t-1} + (\alpha - 1) u_{t-1} \nu'_{t-1} \\
&\quad - (\alpha - 1) u_{t-1} \nu'_{t-2} \\
&\quad + \epsilon_t (\alpha - 1) u'_{t-2} + \epsilon_t \epsilon'_{t-1} + \epsilon_t \nu'_{t-1} - \epsilon_t \nu'_{t-2} \\
&\quad + \nu_t (\alpha - 1) u'_{t-2} + \nu_t \epsilon'_{t-1} + \nu_t \nu'_{t-1} - \nu_t \nu'_{t-2} \\
&\quad - \nu_{t-1} (\alpha - 1) u'_{t-2} - \nu_{t-1} \epsilon'_{t-1} - \nu_{t-1} \nu'_{t-1} + \nu_{t-1} \nu'_{t-2}].
\end{aligned}
$$

Now, applying two properties of the expectation operator ($E(X+Y) = E(X) + E(Y)$ and $E(aX) = aE(X)$ - where X and Y are random variables and a is a constant) in combination with the assumptions that $E(\epsilon_{t-i} \epsilon'_{t-j}) = 0, \forall i \neq j$,

$E(\nu_{t-i}\nu'_{t-j}) = 0, \forall i \neq j$ and $E(\epsilon_{t-i}\nu'_{t-j}) = 0, \forall i, j$ reduces the above equation. These assumptions are sufficient as u_{t-1} can be written in terms of ϵ, namely $u_{t-1} = \epsilon_{t-1} + \alpha\epsilon_{t-2} + \ldots + \alpha^{n-1}\epsilon_{t-n}$ - as above for $E(u_{t-1})$ with $u_{t-n} = 0$ as a starting point:

$$\begin{aligned}
C_1 &= (\alpha-1)^2 E(u_{t-1}u'_{t-2}) + (\alpha-1)E(u_{t-1}\epsilon'_{t-1}) - E(\nu_{t-1}\nu'_{t-1}) \\
&= (\alpha-1)^2 E\left((\epsilon_{t-1} + \alpha\epsilon_{t-2} + \ldots + \alpha^{n-1}\epsilon_{t-n})(\epsilon_{t-2} + \alpha\epsilon_{t-3} + \ldots + \alpha^{n-2}\epsilon_{t-n})'\right) \\
&\quad + (\alpha-1)E\left((\epsilon_{t-1} + \alpha\epsilon_{t-2} + \ldots + \alpha^{n-1}\epsilon_{t-n})\epsilon'_{t-1}\right) - E(\nu_{t-1}\nu'_{t-1}) \\
&= (\alpha-1)^2\left(\alpha E(\epsilon_{t-2}\epsilon'_{t-2}) + \alpha^3 E(\epsilon_{t-3}\epsilon'_{t-3}) + \ldots + \alpha^{2n-3}E(\epsilon_{t-n}\epsilon'_{t-n})\right) \\
&\quad + (\alpha-1)E(\epsilon_{t-1}\epsilon'_{t-1}) - E(\nu_{t-1}\nu'_{t-1}).
\end{aligned}$$

By definition $\sigma_\epsilon^2 I_{NT} = E(\epsilon_{t-i}\epsilon'_{t-i}), \forall i$ and $\sigma_\nu^2 I_{NT} = E(\nu_{t-1}\nu'_{t-1})$:

$$\begin{aligned}
C_1 &= (\alpha-1)^2(\alpha + \alpha^3 + \ldots + \alpha^{2n-3})\sigma_\epsilon^2 I_{NT} + (\alpha-1)\sigma_\epsilon^2 I_{NT} - \sigma_\nu^2 I_{NT} \\
&= \left[(\alpha-1)^2(\alpha + \alpha^3 + \ldots + \alpha^{2n-3})\sigma_\epsilon^2 + (\alpha-1)\sigma_\epsilon^2 - \sigma_\nu^2\right] I_{NT}.
\end{aligned}$$

The following part just deals with the term in the squared brackets, as this contains the parameters to be estimated in the GMM procedure.
Excursion: Convergence of a geometric series

$$\begin{aligned}
s &= (\alpha + \alpha^3 + \ldots + \alpha^{2n-3}) \\
s - \alpha^2 s &= (\alpha + \alpha^3 + \ldots + \alpha^{2n-3}) - (\alpha^3 + \alpha^5 + \ldots + \alpha^{2n-1}) \\
(1-\alpha^2)s &= \alpha - \alpha^{2n-1} \\
s &= \frac{\alpha - \alpha^{2n-1}}{(1-\alpha^2)}.
\end{aligned}$$

This converges to

$$s = \frac{\alpha}{(1-\alpha^2)},$$

as $n \to \infty$ implies that $\alpha^{2n-1} \to 0$ for $0 < \alpha < 1$, which is the stationarity condition of the AR(1) process. Now, the covariance C_1 is determined by:

$$\begin{aligned} C_1 &= (\alpha-1)^2 \frac{\alpha}{(1-\alpha^2)} \sigma_\epsilon^2 + (\alpha-1)\sigma_\epsilon^2 - \sigma_\nu^2 \\ &= \frac{(\alpha-1)(-1)(1-\alpha)\alpha}{(1-\alpha)(1+\alpha)} \sigma_\epsilon^2 + (\alpha-1)\sigma_\epsilon^2 - \sigma_\nu^2 \\ &= \frac{\alpha-\alpha^2}{1+\alpha}\sigma_\epsilon^2 + \frac{\alpha^2-1}{1+\alpha}\sigma_\epsilon^2 - \sigma_\nu^2 \\ &= \frac{\alpha-1}{1+\alpha}\sigma_\epsilon^2 - \sigma_\nu^2. \end{aligned}$$

G Estimation of the 5 imputed SAVE Datasets

The procedure how the estimation results of the 5 imputed datasets from the SAVE study are combined goes back to Rubin (1987) and is also described, for the SAVE data, in Schunk (2007). Each dataset is used for a separate estimation, i.e. one gets 5 results for each coefficient and its corresponding variance. Let $\hat{\theta}_m$ be the estimated coefficient of the m ($m = 1, ..., 5$) imputed dataset, and $\hat{\sigma^2}_m$ the related standard deviation. The overall estimate $\bar{\Theta}$ is simply the average of the estimates from the 5 imputed datasets:

$$\bar{\Theta} = \frac{1}{5} \sum_{m=1}^{5} \hat{\Theta}_m.$$

For the overall variance of the parameter of interest, following Rubin (1987), one first has to compute the within-imputation variance, which is defined by

$$\bar{\sigma^2} = \frac{1}{5} \sum_{m=1}^{5} \hat{\sigma^2}_m.$$

Second, the between-imputation variance must be calculated, given by

$$\bar{b} = \frac{1}{5-1} \sum_{m=1}^{5} (\hat{\Theta}_m - \bar{\Theta})^2.$$

Finally, the overall variance of the multiple-imputation method is determined by

$$\sigma^2 = \bar{\sigma^2} + (1 + \frac{1}{5}) \cdot \bar{b}.$$

and the standard deviation is, of course, $\sigma = \sqrt{\sigma^2}$.

H Abstract

English Version:

This thesis deals with the question if Rank-Dependent Utility or Cumulative Prospect Theory, belonging to the so called Non-Expected Utility models, are better in explaining real life data on consumption and savings decisions than Expected Utility Theory. We use a consumption/savings model and the Method of Simulated Moments to estimate two parameters, which distinguish Expected Utility, Rank-Dependent Utility and Cumulative Prospect Theory in our setting. For the solution of the model we propose a method to which the conventional critique on Non-Expected Utility in dynamic settings does not apply. Our main finding is that Rank-Dependent Utility Theory is the theory which fits the data best.

Deutsche Version:

Diese Dissertation beschäftigt sich mit der Frage, ob die Rank-Dependent Utility Theorie oder die Cumulative Prospect Theorie, welche zu den sogenannten Nicht-Erwartungsnutzen Theorien gehören, reale Daten über Konsum- und Sparentscheidungen besser abbilden können als die Erwartungsnutzentheorie. Wir verwenden ein Konsum/Spar Modell and die Methode der Simulierten Momente um zwei Parameter zu schätzen, welche in unserem Aufbau die Rank-Dependent Utility Theory, die Cumulative Prospect Theorie und die Erwartungsnutzentheorie unterscheiden. Wir schlagen für das Auflösen des Modells eine Methode vor, für welche die Kritik hinsichtlich dem Gebrauch von Nicht-Erwartungsnutzen Theorien in dynamischen Modellen nicht zutreffend ist. Das Endergebnis ist, dass die Rank-Dependent Utility Theorie die Daten am besten abbilden kann.

I Curriculum Vitae

Torsten Jürgen Lisson

Date of Birth: 08.01.1981
Place of Birth: Bogen/Straubing, Germany
Email: torsten.lisson@gmail.com

Academic Education

- PhD Program at the University of Vienna, 2006 to present.
  - Dissertation: Non-Expected Utility vs. Expected Utility Theory in Consumption/Savings Decisions over the Life Cycle
  - Supervision: Univ.-Prof. Dipl.-Ing. Dr. Gerhard Sorger, Univ.-Prof. Dipl.-Ing. Robert Kunst
- Summer School in Applied Macroeconometrics at the University of Salento, Lecce, Italy, July 2009.
- Diploma in Economics at the Ludwig-Maximilian-University in Munich, 2001 to 2006.
  - Diploma Thesis: The Role of Rules of Thumb in Savings Decisions of German Households
  - Supervision: Prof. Dr. Joachim Winter

Research Interests

- Behavioral Economics
- Decision Theory

- Simulation-based Estimation

References

Univ.-Prof. Dipl.-Ing. Dr. Gerhard Sorger (1st Supervisor - Dissertation)
Department of Economics
University of Vienna
Hohenstaufengasse 9
A-1010 Vienna
Office: +43 (0) 1 4277 37443
Email: gerhard.sorger@univie.ac.at
Homepage: homepage.univie.ac.at/gerhard.sorger

Univ.-Prof. Dipl.-Ing. Dr. Robert Kunst (2nd Supervisor - Dissertation)
Department of Economics
University of Vienna
Brünner Straße 70
A-1210 Vienna
Office: +43 (0) 1 4277 37479
Email: robert.kunst@univie.ac.at
Homepage: homepage.univie.ac.at/robert.kunst

Prof. Dr. Joachim Winter (Supervisor - Diploma Thesis)
Department of Economics
Ludwig-Maximilians-University Munich
Ludwigstraße 28 RG
D-80539 München
Office: +49 (0) 89 2180 2459
Email: winter@lmu.de
Homepage: empwi.vwl.lmu.de

Non-Academic Work

- BMW Museum Munich,Germany, Management and Sales, 2004 to 2006.
- Hypo-Vereins Bank Munich, Germany, Working Student, 2004.
- Siemens Munich, Internship Repair Center, 2003.
- Stadtsparkasse Munich, Internship Section Stock Trading, 1999.

Social Service

- Johanniter Allershausen, Mobile Nursing Service, 2000 to 2001.