Empirical analysis of time preferences and risk aversion

Tu, Q.

Publication date:
2005

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Tu, Q. (2005). *Empirical analysis of time preferences and risk aversion*. CentER, Center for Economic Research.

Empirical Analysis of Time Preferences and Risk Aversion

Empirical Analysis of Time Preferences and Risk Aversion

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof. dr. F.A. van der Duyn Schouten, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op

woensdag, 18 mei 2005 om 10.15 uur

door

Qin Tu

geboren op 7 maart 1966 te Jiangxi, China.

PROMOTOR:	prof. dr. Arthur van Soest
CO-PROMOTOR:	dr. Bas Donkers

学而时习之, 不亦说乎?

Confucius

Acknowledgments

I am grateful to numerous "peers" who have contributed towards shaping this thesis. This thesis contains most of my work during my four years study at Tilburg University.

I was often asked why I chose the Netherlands to study when I talked with Dutch people. Actually, I made the decision to studying at Tilburg University by chance five years ago. I had been working at the Institute of World Economics and Politics, in the Chinese Academy of Social Sciences (CASS) for more than seven years before I got an opportunity to visit Tilburg University at the beginning of the year 2000. It was a "Joint Educational Program of Economics" between Tilburg University, CASS and the Chinese Ministry of Education, supported by Dutch government. I spent half year in Tilburg for the program, starting from January 2000. Frankly speaking, I had never heard of Tilburg University before I applied for this program, which subsequently led me to pursue my Ph.D. study in Tilburg. The main reason was that before the year 2000, Dutch universities didn't enter the Chinese high-educational market, and only those top universities in U.S. and U.K., like Harvard and Cambridge, were well known in China at that time. I had a very good time at Tilburg University, so after having stayed there for half a year, I realized that Tilburg University was a very nice place to study, with nice faculty and good facilities. I was lucky enough to meet Professor Bertrand Melenberg, who was my supervisor when I visited Tilburg. With his kind help, I learned a lot about econometrics and found an opportunity to stay in Tilburg for another four years.

At the outset, I would like to express my appreciation to Professor Arthur van Soest for his advice during my doctoral research endeavor for the past four years. As my supervisor, he has constantly encouraged me to pursue my own research ideas, and spent a lot of time discussing my research. His observations and comments helped me to establish the overall direction of my research and to move forward with investigations in depth. I thank Dr. Bas Donkers, who is my co-promotor; without his common-sense, knowledge and perceptiveness, I would never have finished my thesis. I would like to thank Professor Bertrand Melenberg again; I could get his help at any time when I needed it. Just like an excellent supervisor, he read everything so thoroughly, and spent lot of time discussing my work with me. I am very glad that he is one of the committee members.

Special thanks go to the other members of my thesis committee, Professor Jan Potters and Professor Rob Alessie, for managing to read the whole thing so thoroughly, and for some useful comments.

I appreciate the financial support of NWO (Netherlands Organisation for Scientific Research), CentER and the Department of Econometrics and Operations Research, and NAKE, which gave me chances to present my work at conferences, and to attend workshops and summer schools. I would also like to thank administrative and technical staff members of the university who have been kind enough to advise and help in their respective roles. I would also like to thank Marcel Das, Vera Toepoel and all the team at CentERdata for collecting the data which I used in Chapter 4.

I would like to thank all the Ph.D. students and my friends in Tilburg, with special mention to Charles, Steffan, Pierre-Carl, Chendi, Ju Yuan, Hendri, Youwei, Wang Yue, Wang Ting, Vera, Sabine, Danto, Manyi, Yeping, and Chen Ying. Steffan and Pierre-Carl taught me how to ski in France; and together with Chendi and Wang Yue spent a lot of time with me jogging in the small forest near the campus. I hope I could find some time to do this with my friends again in the future. With the help of all my friends, I could still survive after more than four years of hard work.

Last, but not least, I would like to dedicate this thesis to my family, my wife Weiqing and my daughter Weiyu, for their love, patience, courage and understanding. Over more than four years, they allowed me to spend most of my time on this thesis.

Thanks you all, my friends.

Qin Tu

December 2004

Contents

Chapter 1

Introduction

1.1 Summary

Individual decision making is one of the corner stones of economics. Many decisions of economic agents involve trade-offs between different uncertain outcomes and/or between present and future utility. For example, individuals and households have to decide on housing and consumption of durables, saving and portfolio choice, insurances and pension schemes, and household consumption over life cycle. Therefore, many economic theories and models are based on decision making under uncertainty in an intertemporal setting.

After the discounted utility (DU) model and expected utility (EU) model were introduced, their simple and elegant structures quickly made them become the "standard" models of individual decision making in intertemporal choice and under uncertainty, respectively. In the more than fifty years since then, however, many findings contradicting these models (so-called "anomalies") were found in empirical and experimental studies. This made behavioral economics more and more relevant in individual decision making, particularly after the seminal paper of Kahneman and Tversky (1979). Many concepts in behavioral economics and traditional economic theory, such as preference parameters, discount rates, loss aversion, reference points, and risk aversion are all closely linked with individual decision making. The survey paper of Camerer and Loewenstein (2003) gives a good introduction of behavioral economics and how it changes the traditional way of modeling individual decision making. With taking advantage of loss aversion and reference points, some financial phenomena can also be better understood. This is what behavioral finance is doing, and it has recently become a major alternative approach to study individual decision making in financial markets, the traditional view is based upon expected utility maximization. See Barberis and Thaler (2003) for a comprehensive survey of behavioral finance.

The most important concepts in behavioral economics are loss aversion and reference points. With the help of these two concepts, Kahneman and Tversky (1979) presented a reference dependent model for decision making under uncertainty: prospect theory. In 1991, they proposed a similar reference dependent model for riskless choice. The key idea of their reference dependent models is a value function in which gains and losses are defined on deviations from a reference point, rather than on total wealth for utility function as in the traditional theory. The results of experimental studies show that the value function is steeper for losses than for gains. This implies people are loss averse: the disutility of a loss is greater than the utility associated with the same amount of gain. The value function will have a kink at the origin, the reference point. The value function is also assumed to have a diminishing sensitivity: the marginal value of an additional unit both in

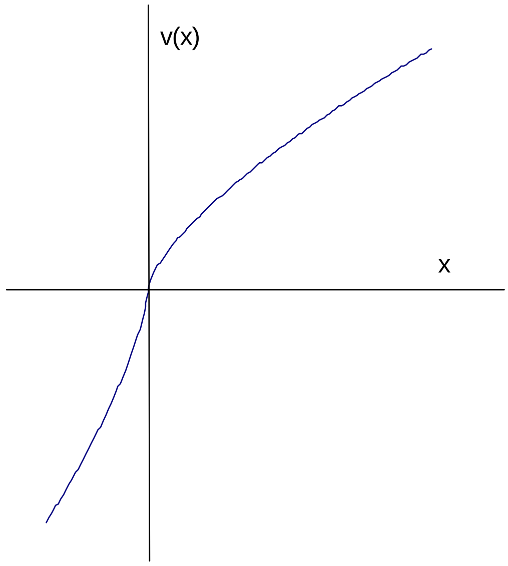


Figure 1.1: Value function

the domain of gains and in the domain of losses is decreasing. See Figure 1.1 for a typical value function in reference dependent models, this kind of value function is quite different from the utility function in the traditional EU and DU models.

Both loss aversion and reference points are concepts that originate from psychology. Many traditional economists are still not convinced of the usefulness of reference points. Reference points are closely linked with loss aversion, and experimental results suggest that they are a useful baseline when people reframe a result as a gain or a loss. Tversky and Kahneman (1991) claimed that the reference point usually corresponds to the decision maker's current position, and it can also be influenced by aspirations, expectations, norms, and social comparisons. More research is needed to investigate how reference points vary across a population, and how they relate to other preference parameters and individual decision making.

Different from the EU model, which is the traditional theory of individual decision making under uncertainty, an appealing analogue in behavioral economics is cumulative prospect theory (CPT) developed by Tversky and Kahneman (1992). In CPT, not only the value function is different from the utility function in the EU model, but also the way probabilities are used differs. Kahneman and Tversky (1979) first introduced a probability weighting

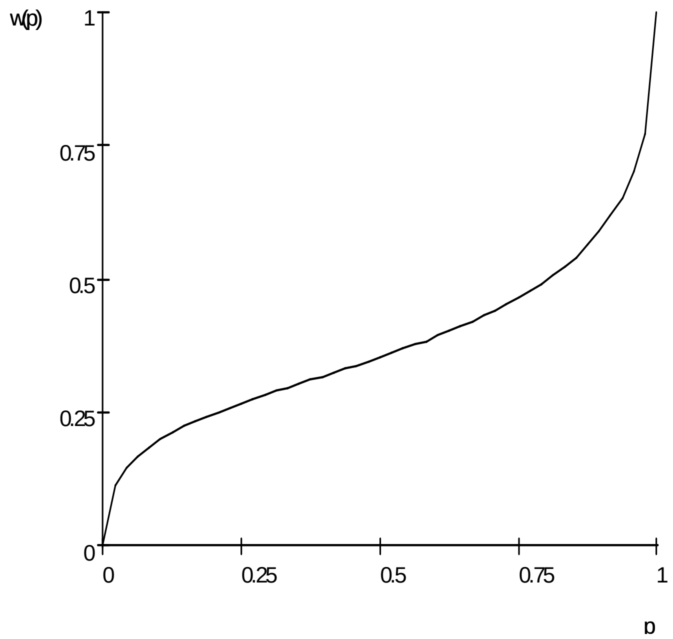


Figure 1.2: Probability weighting function

function, which is a nonlinear transformation of probabilities into "decision weights", $p \to w(p)$, for decision making under uncertainty, because individuals do not treat probabilities linearly. Instead of taking the mathematical expectation of utility where possible utility outcomes are weighted with the probabilities, Kahneman and Tversky propose to use transformed probabilities based upon the nonlinear weighting function. Many experimental findings showed that non-linear probability weights were needed for both very small and big probabilities. Normally, the probability weighting functions are inverse S-shaped both for gains and losses. People tend to overweight small probabilities and underweight large probabilities. This implies that people are risk-seeking in small probabilities for gains and in high probabilities for losses, and risk averse in high probabilities for gains and in small probabilities for losses. Figure 1.2 is an example of an inverse S-shaped weighting function.

In the traditional DU model, the preference of trade-offs between outcomes (either gains or losses) right now and in the future is described entirely by a single discount rate, that is constant for all situations and across all individuals. This implies, in the DU model, that the amount one is willing to pay for postponing a payment of \$100 for one month should be about the same as the amount of compensation one demands for receiving \$100 one month later. But this is not supported by experimental studies. Research in experimental economics has found a number of systematic deviations from the DU model, see Frederick, Loewenstein and O'Donoghue (2002). The most famous ones are the *sign effect,* meaning that gains are discounted more than losses; the *delay-speedup asymmetry*, indicating that different discount rates are used depending on whether a gain (or a loss) is delayed or speeded up; the *hyperbolic discounting*, showing that the discount rate over two periods differs from the product of the two corresponding one period rates; and the *magnitude effect*, implying that small outcomes are discounted more than large ones. Therefore, understanding the way in which people make their decisions and how preferences, behavioral rules, and decision strategies vary with socioeconomic characteristics is crucial for policy making and policy analysis.

It is not easy to use loss aversion and reference points in an empirical study for individual decision making, because reference points and the coefficient of loss aversion are not directly observed. Discrete choice models with reduced forms are still very useful for the study of individual decision making. Housing tenure choice and residential mobility are two closely linked and important decisions faced by households in their life cycles. It should be a good exercise to use a traditional discrete choice model (multinomial probit model) to investigate how these two choices are linked, and how socioeconomic variables affect households decision making.

1.2 Two Examples

The literature suggests that when measuring rates of time preference, four scenarios should be distinguished that lead to markedly different discount rates: delay of gains, delay of losses, speedup of gains and speedup of losses. Consequently, it is crucial to find out what kinds of discount rates are used in order to understand individual's behavior. For example, Warner and Pleeter (2001) estimated the discount rate from a social experiment in which the U.S. Department of Defense offered separatees the choice between two separation benefit packages: a lump-sum separation benefit and an annuity. Because annuities are more common than lump sum payments in practice, people might be more likely to use an annuity as the baseline when they make their decision. If people use the annuity as their reference point, the lump sum payment should be considered as a speedup of gain. In contrast, if people would use the lump sum payment as their reference point, the annuity should be considered as a delay of gain. It seems plausible that most people will use the discount rate of speedup of gains to compare two alternatives when they make their decision because of the popularity of annuity. But actually, their result implied that the vast majority of personnel had discount rates of at least 18 percent. From the results of previous studies, it is obvious that the discount rate of delay of gains is much bigger than that of speedup of gains. This indicates that most of personnel set the lump sum payment as their reference point and employed the discount rate of delay of gains to make their decision, rather than the annuity as their reference points and speedup of gains as their discount rates. This example also shows that the way in which people frame questions have big effects on their decision making.

Another example of a policy issue where the rate of time preference is of vital importance, is giving incentives for purchasing energy saving appliances. In a seminal paper, Hausman (1979) analyzed such a decision. He used data on both the purchase and the utilization of room air conditioners to estimate individual discount rates, based upon the trade-offs between capital cost and operating costs. Purchasing an energy saving appliance normally needs pay quite big amount more for the appliance only once, but then pays less energy bill periodically. This kind of decision making can be considered as the choice between a lump sum loss and periodical losses. If people set the energy efficient appliance as their baseline, then buying the energy inefficient one becomes a delay of losses, people probably use their discount rate of delay of losses to make their decision. On the contrary, if the energy inefficient appliance is their reference point, then buying the energy efficient one becomes speedup of losses; people may use their discount rate of speedup of losses to make their decision. The high discount rates found by Hausman (1979) suggest that for most households this second situation is the relevant setting.

These two examples reveal that it is very important to understand how people choose from different discount rates when they make their decisions, and what is the interaction between the discount rate and other preference parameters, in order to predict individual behaviors well and to analyze them.

1.3 Contents of the Thesis

This thesis is based on four independent papers. Some notations may differ in each chapter. The first three papers are about the interactions among preference parameters, such as loss aversion, reference points, discount rate, risk aversion, and how demographic and socioeconomic variables affect them.

Chapter 2 is a paper about time preference of gains and losses. Specially designed experiments but also field studies have revealed that people tend to discount gains more than losses and that there are differences in the ways delays and speed-ups of payments are discounted. Such different discount rates might have important implications for the analysis of various economic policies, making a better understanding of them of interest. Using a representative household panel survey, the implied discount rates for four different scenarios are analyzed: delay of gains, delay of losses, speed-up of gains, and speed-up of losses. First, the existing literature on the relationship between discount rates and other individual characteristics is summarized. Then the discount rates are linked to frequently observed demographic variables, like gender and age, but also to subjective variables, such as price expectations. Many of these variables significantly affect individual discount rates, and, more importantly, these variables affect the discount rates in different ways. Such differences permit us to generate scenario specific discount rates for each individual. Unobserved heterogeneity, which explains a substantial part of the variation in the reported discount rates, is allowed in the model. Interestingly, both unobserved heterogeneity and the remaining error terms appear to be positively correlated across the four scenarios. The observed relationships can be used to better understand and predict the behavior of households for policy evaluation.

In order to better understand time preference of gains and losses in four different scenarios, a structural model which incorporates loss aversion and reference points for intertemporal choice based on the insights of Loewenstein's (1988) reference point model is presented in Chapter 3. Data from a Dutch representative household panel survey of the years 1997-2002 is used, containing rich information on individual time preferences and other characteristics. A non-linear random coefficients model with panel data is employed to jointly estimate the reference points of delay and speedup, the coefficient of loss aversion and the discount rate. The result shows that on average the reference point of delay is larger than the reference point of speedup, consistent with the hypothesis of Loewenstein; the mean coefficient of loss aversion is around two, similar to other findings, showing that the disutility of a loss is as twice large as the utility associated with the same amount of gain; females are more loss averse than males, and high education and age make people less loss-averse; high educated or older people are also more patient.

In Chapter 4, a joint estimation of loss aversion and probability weighting function with a power utility function is presented for decision making under uncertainty. Cumulative prospect theory is becoming a dominant way to model individual decision-making under risk. Tversky and Kahneman estimated a model with a power utility function with loss aversion and a two-part power function as the weighting function for probabilities. In this study, we use a similar structure to model individual decision making. The five main parameters of interest are the powers in the value functions for gains and losses, the loss aversion parameter, and the coefficients of the weighting functions for gains and losses. To model heterogeneity across the population, the empirical model treats these as random coefficients, depending on observed demographics and unobserved characteristics. The data we use stem from a survey which is a representative of the Dutch population, with seven questions about one or two bets. Our results show that on average powers of value functions are 0.68 and 0.73 for gains and losses respectively, females have smaller power for gains than males, implying that females are more risk averse in domain of gains. The average coefficient of loss aversion is 3.1; on average, the coefficients of weighting functions are 1.0 and 0.59 for gains and losses respectively.

The last paper, Chapter 5, is quite different in content from the other three papers, but uses similar econometric techniques and models. The chapter is about how to model household's mobility (moving decision) and housing tenure choice (decision of renting versus owning) jointly, using a multinomial probit model with panel data. Account is taken of the fact that a change of housing tenure can only be observed when the household moves. The models are estimated by the method of maximum simulated likelihood, emphasizing the importance of properly accounting for the initial conditions problem. The estimation is based on unbalanced panel data from the CentER Panel, 1994-2003. Negative state dependence in the moving decision is found. Owners are less likely to move to either another owner-occupied or a rented home than renters, which can be explained from the much higher moving costs for owners.

1.4 Further Research

In this thesis, there are three papers about the interactions among preference parameters, such as the discount rate, loss aversion, reference points, and risk aversion, and one paper about estimating household decision making using a discrete choice model with a reduced form. There are still many questions that need further research.

The problem of dynamic inconsistency is quite difficult to understand. The currently popular way to explain dynamic inconsistency assumes that the discount rate is a decreasing function of the time period involved. This is hyperbolic discounting. It might be extremely interesting to investigate whether the structural model in chapter 3, which is a behavioral model with loss aversion and reference points for intertemporal choice, could be one of the explanations of dynamic inconsistency. Another question which was found to be important and in need of further research is to find how to use preference parameters: discount rate, loss aversion, reference points, and risk aversion in an often used reduced form model like in chapter 5. With help of these concepts of behavioral economics, maybe we can better predict and understand household moving decisions and tenure choice.

Chapter 2

The Time Preference of Gains and Losses

2.1 Introduction

Time preference plays a crucial role in large parts of economics. Indeed, economic agents, when dealing with intertemporal choices are confronted with the task to compare economic quantities related to the present and future periods. As early as 1834, John Rae discussed the concept of time preference quite clearly in his book "*The Sociological Theory of Capital*". Samuelson introduced the discounted utility model in 1937, as a way to model time preference and, more generally, intertemporal choice. The simple and elegant structure of this model made it popular straight away. Since then it has been dominating in intertemporal choice modelling. In modern microeconomic theory of individual and household behavior, people are assumed to maximize their lifetime discounted utility given some economic and other constraints, resulting in decisions on their consumption, saving, and investment behavior. Moreover, economic theories like those dealing with asset pricing and economic growth generally include intertemporal tradeoffs by means of time discounting. So, indeed, time preference modelled by means of time discounting is one of the cornerstones of economic analysis. A recent overview of the history of and studies on time preference is given by Frederick, Loewenstein and O'Donoghue (2002).

Time preference has important implications for many aspects of public policy and individual economic behavior. Policy makers with a good understanding of how households decide on intertemporal tradeoffs are able to design policies that are better accepted by the public. For example, the U.S. Department of Defense offered separatees the choice between two separation benefit packages: a lump-sum separation benefit and an annuity. Over half of the officers and over 90 percent of the enlisted personnel took the lump-sum payment. Offering this choice was not only welfare improving for separatees, but also saved \$1.7 billion in separation costs (Warner and Pleeter, 2001). Another policy issue where the rate of time preference is of vital importance, is giving incentives for purchasing energy saving appliances. In a seminal paper, Hausman (1979) analyzed such a decision. He used data on both the purchase and the utilization of room air conditioners to estimate individual discount rates, based upon the tradeoffs between capital cost and operating costs.

Estimation of individual discount rates from real-life decisions is often very complicated, as many other factors affect such decisions. See, for an illustration, the discussion of Hausman's work by Kooreman (1995), who showed that accounting for stochastic lifetimes of the appliances substantially affects the estimated discount rates. Consequently, a large literature has arisen that analyzes discount rates in specially designed experimental frameworks, see Frederick, Loewenstein and O'Donoghue (2002). This experimental literature has found a number of systematic deviations from the traditional discounted utility model.

In this paper we analyze two of these findings in particular using a representative household panel rather than small scale experiments: the *sign effect* and the *delay-speedup asymmetry*. The *sign effect* (Thaler, 1981; Loewenstein, 1987, 1988; Shelley, 1993) is the asymmetry in the discounting of gains versus losses, such as postponing receiving a prize versus paying a fine. Experiments show that gains are discounted more than losses. The traditional discounted utility model, however, states that the amount one is willing to pay for postponing a payment of $100 for one month should be about the same as the amount of compensation one demands for receiving $100 one month later. That is to say, gains and losses should be discounted equally. The *delay-speedup asymmetry* is the finding that different discount rates are used depending on whether a gain (or a loss) is delayed or speeded up. For example, Loewenstein (1988) showed that respondents who expected to receive a video cassette recorder one year later were willing to pay an average of $54 to receive it immediately, but those who expected to receive it immediately demanded an average of $126 to delay its receipt by a year.[1]

If present, the *sign effect* and the *delay-speedup asymmetry* should be taken into consideration when designing economic policies. In particular, the framing of the policy measure may determine which discount rate is used, and can, therefore, affect the response to the policy (Brendl and Higgins, 1996). Consider, for example, once again the payment of a lump-sum versus an annuity for the US army separatees as analyzed by Warner and Pleeter (2001). Take an employee who expects to receive an annuity. Suppose this employee is also offered the choice to receive a lump-sum payment instead, giving him a large instantaneous gain. The discount rate applied is, therefore, the discount rate for speeding up gains. Alternatively, if the separatee takes the lump-sum payment as the benchmark, and the separatee is offered as alternative an annuity, then the relevant discount rate will be the discount rate for delay of gain. Due to the *delay-speedup asymmetry* the choice in both cases might be quite different. In other situations, other discount rates will be used. When purchasing energy saving equipment, consumers usually have to pay a higher price. The benefits are lower energy bills in the future.

[1] Other deviations from the traditional discounted model include, in particular, *hyperbolic discounting* (Laibson, 1997), which is the phenomenon that the discount rate over two periods differs from the product of the two corresponding one period rates, and the *magnitude effect*, meaning that small outcomes are discounted more than large ones. In our analysis we will keep the time horizon and the outcome fixed, and, thus, will not address these issues.

Suppose the status quo for this decision is buying the non-energy saving product. The discount rate that will be used for this decision is the discount rate for speeding up losses, which, due to the *sign effect*, might be much higher than the discount rate used for speeding up gains. Given the large differences among the discount rates that are indicated by the literature on time preference, it is important to use information on the appropriate discount rate in analyzing a given policy.

Recently, researchers have started to link discount rates to other individual or household characteristics, see, for example, Harrison, Lau and Williams (2002), and Kirby et al. (2002). An important motivation for investigating this link is that one will be better able to predict the choices made by households with given characteristics. For example, many policy issues that relate to pensions affect only the elderly. The appropriate discount rate for analyzing such a policy will, therefore, be the discount rate used by the elderly, which might differ from the population average. Similarly, programs aimed at getting high-school drop-outs back into the education system in order to let them invest in the short run with the purpose of obtaining long run benefits, mainly affect adolescents.

In this paper we extend this literature by analyzing how individual characteristics affect the different types of discount rates that one can use. We distinguish four different scenarios that (might) lead to different discount rates. These scenarios differ in whether it concerns a gain or a loss (to take account of the potential *sign effect*) and in whether the payment date is postponed or speeded up (to deal with a possible *delay-speedup asymmetry*). The literature suggests that these four scenarios lead to markedly different discount rates, a finding supported by our data. However, to do proper policy analysis on various types of policies, one would need to know the level of the discount rate for the households that are affected. Therefore, we relate the four different discount rates to an extensive set of explanatory variables. This provides information on each of the four discount rates, enabling policy makers to select the appropriate type of discount rate in combination with the relevant household characteristics.

Moreover, using four years of panel data on respondents who answer the questions for all four scenarios, we are able to distinguish between unobserved heterogeneity and idiosyncratic noise. This makes it possible to analyze how much of the variation that is not explained by observed household's or respondent's characteristics reflects genuine heterogeneity in time preferences and how much is noise. By jointly analyzing the four scenarios, we can also identify the extent to which common factors drive the four discount rates.

Our analysis is based on a rich dataset from the Netherlands, the CentER Saving Survey (CSS), which is representative of the Dutch population. The data we use stems from four waves, 1997-2000. The survey contains sixteen questions about time preference, including the four different scenarios, but with varying amounts and time horizons. This wealthy data set gives us a unique opportunity to investigate the relationships of time preference and its determinants in detail. In our study, four ordered probit models are used to simultaneously estimate the time preference for gains and losses, and delay and speedup. Maximum Simulated Likelihood and the GHK simulator[2] are employed to estimate the model.

A remarkable finding is that the mean of the discount rate for speedup of gains has the same size as the rate for delay of losses, but is much smaller than the mean of the discount rate for delay of gains. We find that discount rates vary with individual characteristics, and the four discount rates vary in different ways, implying that policy analysis should take account of different trade-offs in different demographic groups. In particular, females have lower discount rates than males in all four scenarios, but especially in delaying losses and speeding up gains; age has a robust U-shaped relationship with both discount rates of gains and losses. We find that unobserved individual heterogeneity explains a substantial part of the variation in reported discount rates, and that heterogeneity in the four discount rates is positively correlated. Idiosyncratic errors for the four scenarios in a given time period are positively correlated too. The correlation coefficients of the random effects and the error terms in the econometric model are highest between the discount rates for delay of losses and speedup of gains. Thus, we may conclude that the discount rates for speedup of gains is more like that for delaying losses than delaying gains; and, for the same reason, speedup of losses is more similar to delaying gains than losses.

The remainder of this paper is organized as follows. In Section 2, we briefly review the relationships between time preference and a number of socioeconomic variables that have been found in both the theoretical and the empirical literature. We describe the data in Section 3. In Section 4, we discuss the econometric model and the estimation procedure. In Section 5, the results of two models are presented. The first model contains only basic demographics. The second model contains a wide range of variables, including subjective variables that can be seen as alternative indicators of time preference and are helpful in predicting discount rates. Section 6 concludes.

[2]Named after Geweke, Hajivassiliou, Keane who developed the procedure independently in the early 1990s. See Hajivassiliou and Ruud (1994).

2.2 Socioeconomic variables and time preference

In this section we discuss the existing literature on time preference and its relationship with other individual characteristics. Many studies have considered time preference as a determinant of household behavior. Other studies aim at explaining the rate of time preference from other factors. A prominent example of the latter is Becker and Mulligan (1997). They argue that time preference is affected by wealth, mortality, addictions, uncertainty, and many other demographic and socioeconomic variables. Our interest is in predicting who is more and who is less impatient. Therefore, we will not pay attention to causality in the empirical analysis, although we will discuss this issue here, to get insight in the mechanisms through which the empirical relation between time preference and other variables comes about. Given the important role of time preference in all kinds of decisions, only very few variables are unambiguously strictly exogenous to the level of time preference. In the data set we have available, only gender and age would be classified as such. All other variables we discuss below are potentially endogenous, i.e., they might also be determined by the rate of time preference. Still, these variables should help us in identifying who is patient and who is not.

2.2.1 Age and time preference

Becker and Mulligan (1997) give a clear prediction for the age-pattern of time preference: the future would tend to be discounted relatively heavily at both young and old ages, giving a U-shaped relationship between the discount rate and age.

However, empirical studies on the rate of time preference have not been able to confirm this U-shaped pattern. For example, Harrison, Lau, and Williams (2002) use three dummy variables for age: young, middle, old, and found an inverse U-shaped age pattern, which was not significant. Pender (1996) did an experiment with 96 participants in two villages in southern India. He found that the coefficient of age was very small and insignificant. But some studies comparing the elderly with the youngsters, did find that elderly people are more impatient than youngsters. For example, Van Der Pol and Cairns (1999) showed that the average discount rates of respondents younger and older than 64 were 0.118 and 0.166 in case of 5 years delay, respectively, and 0.066 and 0.114 in case of 13 years delay. Kirby et al. (2002) also investigated the age pattern of time preference. They find the lowest discount rates at the age of 20 and increasing discount rates after age

20. Donkers and Van Soest (1999) found that the rate of time preference is negatively correlated with age, using the Dutch CentER Savings Survey waves of 1993 and 1995.

2.2.2 Gender and time preference

A number of empirical studies have included gender as an explanatory variable. Kirby and Marakovic (1996) estimated the discount rate of delay of gains with an experiment that used more than 600 students as subjects. A reliable gender difference was found with males discounting at higher rates than females, on average. Daniel (1994) and Donkers and Van Soest (1999) found the same: females have lower discount rates, on average. Other studies have not found a significant effect of gender, see, for example, Kirby et al. (2002), Harrison, Lau and Williams (2002), and Pender (1996).

2.2.3 Health and time preference

More than 160 years ago, Rae (1834) already realized that the uncertainty of human life has important effects on time preference. He wrote:[3]

"When engaged in safe occupations, and living in healthy countries, men are much more apt to be frugal, than in unhealthy, or hazardous occupations, and in climates pernicious to human life. Sailors and soldiers are prodigals. In the West Indies, New Orleans, the East Indies, the expenditure of the inhabitants is profuse. The same people, coming to reside in the healthy parts of Europe, and not getting into the vortex of extravagant fashion, live economically. War and pestilence have always waste and luxury, among the other evils that follow in their train." (Rae 1834, p. 57)

Clearly, bad health of individuals reduces their life expectancy, and it is easy to understand that people with long life expectancy are more patient. Many studies indicate that there is some relationship between time preference of individuals and their health status. Becker and Mulligan (1997), following Rae's (1834) suggestion, argue that differences in health cause differences in time preference. Better health reduces mortality and, therefore, raises future utility levels, which would make people more patient. On the other hand, Fuchs (1982) and others argue that differences of time preference have big effects on the individual health-related decisions, and, therefore, influence the health of the individual. In all cases, the conclusion is that better health status is associated with lower time preference and more patience. A recent paper of Picone et al. (2004) checked the role of risk and time preference,

[3] Cited from Frederick, Loewenstein and O'Donoghue (2002).

expected longevity, and education on demand for medical tests of women, their results revealed that women with a short time horizon is less likely to do these tests, it means impatience makes people invest less in health.

As explained above, we will not analyze causality between health and time preference, but will model the (partial) correlation between health and time preference so that information on health status can be used to better forecast individual rates of time preference. In this paper we use three variables to present the health status of individuals: First, the Quetelet index or Body Mass Index (BMI), which is a common measure for obesity; second, a self-reported measure on general health; and third, a dummy indicating a serious illness or other health problems in the previous year.

2.2.4 Addiction and time preference

Addiction is an interesting topic in relation to time preference. Many experimental studies illustrate that drug addicts discount the future significantly heavier than those who do not use drugs.[4] Carrillo (1999), O' Donoghue and Rabin (2000), and Gruber and Koszegi (2001) also show that hyperbolic discounting could explain the over-consumption of the harmful addictive products in their models. In general, the causality between addiction and impatience is not clear. Becker and Murphy (1988) assume that people with higher discount rates would consider the future less important, making them more likely to become addicted. But Becker and Mulligan (1997) also stress the reverse causality, arguing that addictions cause persons to discount the future more heavily, and this higher discount rate might lead to an even stronger addiction.

We do not have data about strong addictions like drug-use, but our survey does include information on smoking and drinking behavior. These could be considered as some kind of addiction, though not as strong as drug-use.

2.2.5 Income and time preference

Economic theory provides no reason to expect that people with lower incomes would have higher discount rates, because the relative valuation of consumption now and in the future need not depend on the level of income. In Becker and Mulligan's (1997) model also, there are opposing forces with respect to

[4] Madden et al. (1997) show that Opioid-dependent participants discounted delayed monetary rewards significantly more than non-drug-using participants. Furthermore, opioid-dependent participants discounted delayed heroin significantly more than delayed money. Kirby, Petry and Bickel (1999), and Giordano et al. (2002) have similar results for heroin addicts.

the impact of income on time preference, resulting in an ambiguous overall effect.

The existing empirical literature has some results on the relationship between income and time preference. Hausman (1979) shows that the discount rate is inversely related to income (for example, 39 percent for households with income below $10,000 and 8.9 percent for households earning between $25,000 and $35,000). Harrison, Lau and Williams (2002) also found that rich people have lower discount rates than poor, using Danish data, but their results were not significant. Finally, Houston (1983) presents individuals with a decision of whether to purchase a hypothetical energy-saving device and also concludes that income "played no statistically significant role in explaining the level of discount rate."

2.2.6 Schooling and time preference

Becker and Mulligan (1997) discuss two different causal relationships between schooling and time preference. They argue that schooling focuses students' attention on the future, and at the same time educated people should be more productive at reducing the remoteness of future pleasures. Both effects imply a lower discount rate for people with higher education levels. This is also implied by the notion that inherently more patient people will tend to invest more in education, reducing current consumption in order to reap the benefits later in life.

There are all kinds of empirical results about the relationship between schooling and time preference. Viscusi and Moore (1989) used a multi-period Markov model of the lifetime choice of occupational fatality risks to estimate the discount rate. They used the 1982 wave of the University of Michigan Panel Study of Income Dynamics (PSID) and show that the discount rates decrease with education. Harrison, Lau and Williams (2002) and Pender (1996) found insignificant effects of education.

2.3 Data

The data we use is a panel data set with four waves (1997-2000) taken from the CentER Savings Survey (CSS, formerly known as the VSB Panel). The CSS is a large Dutch household survey, collected every year for a panel of more than 2000 households, starting from 1993. The CSS is a rich data set containing information on employment status, pensions, accommodation, mortgages, income, assets, debts, health, economic and psychological concepts, and personal characteristics. Our data constitutes an unbalanced panel, with a total of 3,938 individuals and 6,962 observations. Table 4.1 shows the structure of this unbalanced panel. The average time that an individual stayed in the panel is 1.8 years.

Table 2.1: Structure of the panel

By wave		By number of waves		
Year	Observations	Number of waves	Obs.	Number of individuals
1997	2,657	1	2,299	2,299
1998	1,363	2	1,378	689
1999	1,366	3	1,545	515
2000	1,576	4	1,740	435
Total	6,962	Total	6,962	3,938

Left panel: Year is the survey year, 1997-2000, in total we have four waves of the survey.

Right Panel: Number of waves that the households stay in the panel.

Starting from the year 1997, a detailed set of questions about time preference is included in the CSS.[5] There are sixteen questions about the way people value opportunities in the future compared to the present. These questions differ on four aspects with each aspect having two levels, resulting in a total of sixteen questions. The first aspect is the amount of money concerned, either Dfl. 1000 or Dfl. 100,000.[6] The second is the time horizon, either three months or one year. The third is whether the amount of money is to be received or to be paid[7]. The last one is whether the transaction (payment or receipt) is speeded up or delayed.

In this paper, we analyze the four questions with the amount of money equal to Dfl. 1000 and a time horizon of one year. Therefore, the four different scenarios we consider are the following (using the (translated) questions asked in the questionnaire):

Delay of gains

> Now imagine that the National Lottery asks if you are prepared to wait A YEAR before you get the prize of Dfl. 1000. There is no risk involved in waiting. How much extra money would you AT LEAST want to receive to compensate for the

[5] In earlier waves time preference was elicited with questions that differ in the answering format (cf. Donkers and van Soest, 1999). Therefore, we do not use these data.

[6] 1 Dfl. is approximately 0.45 Euro.

[7] We consider "need to pay a tax assessment " as a loss and "win a prize of the National Lottery" as a gain.

waiting term of a year? If you agree on the waiting term without the need to receive extra money for that, please type 0 (zero).

Delay of losses

Imagine again that you have to pay a tax assessment of Dfl. 1000 today. Suppose that you could wait A YEAR with settling the tax assessment. How much extra money would you AT MOST be prepared to pay to get the extension of payment of A YEAR? If you are not interested in getting an extension of payment or if you are not prepared to pay more for the extension of payment, please type 0 (zero).

Speedup of gains

Imagine again that you receive notice from the National Lottery that you have won a prize worth Dfl. 1000. The money will be paid out after A YEAR. The money can be paid out at once, but in that case you receive less than Dfl. 1000. How much LESS money would you AT MOST be prepared to receive if you would get the money at once instead of after a year? If you are not interested in receiving the money earlier or if you are not prepared to receive less for getting the money earlier, please type 0 (zero).

Speedup of losses

Imagine again that you receive a tax assessment of Dfl. 1000. The assessment has to be settled within A YEAR. It is, however, possible to settle the assessment now, and in that case you will get a REDUCTION. How much REDUCTION would you AT LEAST want to get for settling the assessment now instead of after a year? If you are not interested in getting a reduction for paying early or if you think there is no need to get a reduction for paying early, please type 0 (zero).

Each of these four questions leads to a different discount rate, providing discount rates for the delay of gains (δ_{DG}), delay of losses (δ_{DL}), speedup of gains (δ_{SG}) and speedup of losses (δ_{SL}). If we use x_{DG}, x_{SG}, x_{DL}, x_{SL} to represent the answer to each question above, then we can compute these four discount rates as follows:

$$\begin{aligned}
\delta_{DG} &= \frac{x_{DG}}{1000} \\
\delta_{DL} &= \frac{x_{DL}}{1000} \\
\delta_{SG} &= \frac{x_{SG}}{1000 - x_{SG}} \\
\delta_{SL} &= \frac{x_{SL}}{1000 - x_{SL}}
\end{aligned}$$

Some descriptive statistics on these discount rates are provided in Table 4.2. Notice here that we use only those observations with a discount rate of at most 120% to compute some of descriptive statistics. In the Appendix we provide a list of definitions of our explanatory variables and some descriptive statistics.

Table 2.2: Descriptive statistics of the discount rates

Variable	Mean	Median	Std. Dev.	Min.	Max.
δ_{DG}	0.208	0.1	0.250	0.0	1.2
δ_{DL}	0.032	0.0	0.076	0.0	1.2
δ_{SG}	0.028	0.0	0.078	0.0	1.2
δ_{SL}	0.109	0.053	0.164	0.0	1.2

Note: the mean, std. dev., min. and max. of δ are computed with the obs.which δ is smaller or equal than 1.2

In case of perfect financial markets without constraints[8], a "rational" individual should have discount rates that are the same for all four scenarios. As expected, our data does not support this. Instead, one can see from Table 4.2 that our data is in line with the findings in the literature. First of all, people discount gains heavier than losses, i.e., the mean of δ_{DG} (the discount rate of delay of gains) is more than five times larger than that of δ_{DL}(the discount rate of delay of losses); this is what is called the *sign effect*. The sign effect is also clearly present when comparing δ_{SG} (the discount rate of speedup of gains) and δ_{SL} (the discount rate of speedup of losses). The second confirmation of existing findings is the *delay-speedup asymmetry*: we find that δ_{DG} is much bigger than δ_{SG}, and we find that δ_{DL} is much smaller than δ_{SL}.

[8]Our data includes information on savings accounts. More than 95% of the individuals had more than Dfl 1,000 in their bank account. This means that these people are not likely to be financially constrained for the amount of Dfl 1,000 we investigate.

The answers to the questions show a number of patterns worth noting. First, respondents have a tendency to provide answers in relatively round numbers, for example, 10, 20, 25, or 50, but not 11 or 37. For example, for the delay of loss question 57% of the observations used one of these four numbers 25, 40, 50, 100, in case the answer is not equal to zero. This is also illustrated in Figure 2.1. We account for this in our econometric model in the next section.

A second feature is that there are a large number of respondents that answer zero. In particular for the delay of loss and the speedup of gain, more than 50% of the answers is zero, indicating that these respondents are not willing to pay more or receive less. Table 2.3 provides some statistics and the average discount rates and the percentage of answers that equal zero for each of the four waves. In general, the average discount rates increase over time, especially in 2000, while the number of answers equal to zero decreases over time.

Table 2.3: Discount rates and zero answers by wave

Variable	Year	Mean	Std.Dev.	% of obs. answered with zero
δ_{DG}	1997	0.190	0.234	18.3%
	1998	0.210	0.251	13.9%
	1999	0.215	0.263	14.1%
	2000	0.232	0.261	9.2%
δ_{DL}	1997	0.029	0.074	67.7%
	1998	0.032	0.076	64.9%
	1999	0.033	0.076	62.3%
	2000	0.037	0.078	57.0%
δ_{SG}	1997	0.026	0.070	70.5%
	1998	0.028	0.082	69.6%
	1999	0.027	0.075	69.1%
	2000	0.033	0.089	63.5%
δ_{SL}	1997	0.105	0.168	38.8%
	1998	0.098	0.149	37.1%
	1999	0.104	0.160	35.6%
	2000	0.131	0.172	20.2%

Note: the mean and Std. Dev. of δ are computed with the obs. which δ is smaller or equal than 1.2.

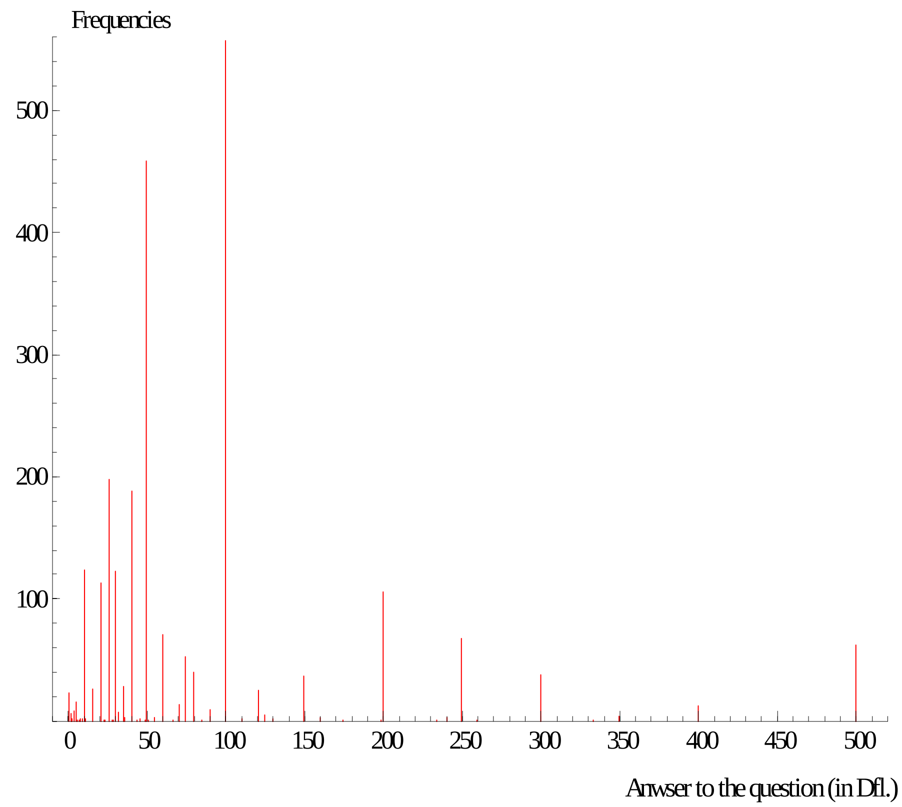


Figure 2.1: Frequencies of the answer to the question on Delay of losses

2.4 Econometric Model

In this section, we present the econometric model for the analysis of the observed discount rates for the four scenarios. To estimate the time preference for the four scenarios jointly, we use a model with four equations and allow for correlated errors and individual effects. As discussed in the previous section, respondents tend to use round numbers to answer the questions. We expect this to be the result of rounding off the "true" answer, an observed discount rate of 10% might indicate that the actual discount rate is, for example, between 5% and 15%. To account for this, we group the reported discount rates and define a categorical outcome y_{it} for each discount rate, which indicates the interval of the reported discount rate δ_{it}. We use 0%, 7.5%, 15%, 30%, and 60% as the cutoff points and classified the data into 6 intervals as follows:

$$\begin{aligned}
y_{it} &= 0 \text{ if } \delta_{it} \leq 0 \\
y_{it} &= 1 \text{ if } 0 < \delta_{it} \leq 7.5 \\
y_{it} &= 2 \text{ if } 7.5 < \delta_{it} \leq 15 \\
y_{it} &= 3 \text{ if } 15 < \delta_{it} \leq 30 \\
y_{it} &= 4 \text{ if } 30 < \delta_{it} \leq 60 \\
y_{it} &= 5 \text{ if } \delta_{it} > 60
\end{aligned}$$

This classification is chosen on the basis of the distribution of the reported discount rates in the data. The categories are located around the focal points with many observations. In particular, there are a large number of observations with 0%, 5%, or 10%, which will dominate the lower three categories of our discretization. Another advantage of this discretized variable is that outliers, i.e., very large observations, will not affect the estimation results too much. When we would use a continuous model, taking the reported values as the actual values, we would use more information on the detailed answers, inferring too much precision from the rounded numbers. This is avoided by using the discretized variable instead of the continuous one.

To explain the ordered discrete dependent variable, we adopt the ordered probit model adapted to a panel data context with multiple equations. Using standard notation for the ordered probit panel data model, the underlying latent variable of individual i at time t for scenario J, $J = DG,\ DL,\ SG,\ SL$, is denoted as y_{it}^{J*}, and we model it as

$$y_{it}^{J*} = X_{it}'\beta^J + \varepsilon_{it}^J \qquad i = 1, \cdots, N;\ t = 1, \cdots, T.$$

For each individual respondent we have four scenarios, i.e., four observed outcomes, per wave and up to four waves. As we expect that the explanatory variables in X_{it} do not capture all individual heterogeneity, we allow for random effects by assuming

$$\varepsilon_{it}^{J} = \eta_i^J + \nu_{it}^J$$

We assume that both the individual effects and the error term are multivariate normal with arbitrary covariance structure, independent of the regressors $X_i' = (X_{i1}', \ldots, X_{iT}')'$:

$$\eta_i \equiv (\eta_i^{DG}, \eta_i^{DL}, \eta_i^{SG}, \eta_i^{SL})' | X_i \sim N(0, \Sigma_\eta)$$

and

$$\nu_{it} \equiv (\nu_{it}^{DG}, \nu_{it}^{DL}, \nu_{it}^{SG}, \nu_{it}^{SL})' | X_i \sim N(0, \Sigma_\nu).$$

Moreover, we assume that ν_{it} is independent of ν_{js}, for all i, j and all $s \neq t$, and that the error terms ν_{it} are independent of the individual effects η_i. The latent variable is transformed into the categorical outcome as in an ordered probit model:

$$\begin{aligned} y_{it}^J &= 0 \text{ if } y_{it}^{J*} \leq cut_1^J \\ y_{it}^J &= l \text{ if } cut_l^J < y_{it}^{J*} \leq cut_{l+1}^J,\ l = 1, .., 4 \\ y_{it}^J &= 5 \text{ if } y_{it}^{J*} > cut_5^J \end{aligned}$$

Here $cut_1^J, \ . \ . \ , cut_5^J$ are cutoff points for equation J. The most common normalization restrictions in ordered probit models are restrictions on the variance of the error term, ε, and the constant term. In our application, however, we are interested in comparing parameter estimates across equations. As the variances of the error terms are not necessarily equal, we adopt a different normalization: we set $cut_1^J = 0$ and $cut_5^J = 60$. These values are the same values that are used in discretizing the observed discount rates. A change from 0 to 60 in the underlying latent variable, y^{J*}, is therefore equal to a change in the observed discount rate from 0% to 60%. This permits us to compare the estimated coefficients across the four equations.

Our aim is to estimate the four equations at the same time. Given the four waves in our panel, this means that we have 4×4 ordered probit equations with correlated error terms. In particular, the 16×16 dimensional covariance matrix is given by

$$\Sigma = \begin{bmatrix} \Sigma_\eta + \Sigma_\nu & \Sigma_\eta & \Sigma_\eta & \Sigma_\eta \\ \Sigma_\eta & \Sigma_\eta + \Sigma_\nu & \Sigma_\eta & \Sigma_\eta \\ \Sigma_\eta & \Sigma_\eta & \Sigma_\eta + \Sigma_\nu & \Sigma_\eta \\ \Sigma_\eta & \Sigma_\eta & \Sigma_\eta & \Sigma_\eta + \Sigma_\nu \end{bmatrix}$$

It is almost impossible to compute the probability of this 16 dimensional multivariate normal distribution using numerical methods directly. Therefore, we rely on simulation techniques to calculate these high dimensional integrals. The GHK simulator and the method of maximum simulated likelihood (MSL) are well known tools to estimate this kind of high dimensional discrete choice models, see Hajivassiliou and Ruud (1994) or Train (2003). We use 100 random draws[9] when we compute the simulated likelihood, and then employ the BHHH[10] algorithm to maximize the simulated likelihood.

2.5 Results

As time preference might be determined simultaneously with a large number of individual characteristics, including, for example, education, income, health status, and home ownership, we estimate two models. The first model includes only strictly exogenous variables, which, in our data, are age and gender. The second model includes many more variables that can improve the predictive performance of the model. This model reveals the correlations between time preference and the explanatory variables, but these relations are not necessarily causal.

The estimation results of the first model with only exogenous variables are presented in Table 2.4. The model includes age, age-squared, a dummy for female, and three dummies to capture changes over time. The eight parameters of age and age-squared are all significant except age-squared in case of δ_{SL}. These estimates indicate that time preference has a U-shaped age pattern, which would fit Becker and Mulligan's (1997) prediction perfectly. The minimum values, however, are attained at age 86, 79, 52, and 168 for δ_{DG}, δ_{DL}, δ_{SG}, and δ_{SL}, respectively. For a reasonable age range like 20 to 85 years old, only the time preference for speedup of gains and to some extent delay of losses have a U-shaped pattern in the relevant age range. The discount rates for the other two scenarios decrease with age. This is in line with the mixed findings on the U-shaped age structure in empirical studies in the literature.

The results in Table 2.4 also show that females are more patient than males. This holds especially for time preference of delay of losses and speedup of gains. Given the scaling of the cut-off levels, the parameters have a direct

[9] Our results are checked by double draws (using 200 draws), and compared to the result using 100 draws, the relative change of estimated parameters is smaller than 5 percent. It seems that it is accurate enough (5%) with 100 draws to estimate our model.

[10] Berndt, Hall, Hall, and Hausman (1974) proposed this procedure of the numerical search for the maximum of the log-likelihood.

interpretation. For example, the estimate of -4.92 for females in the delay-loss-scenario indicates that for this scenario females have a discount rate that is about 4.9% points lower than for males on average. Similar interpretations hold for the other coefficients.

The second part of Table 2.4 are Wald tests for two hypotheses on the parameters. Wald Test 1 is the test of the null hypothesis that the four parameters of the same variable are all zero, i.e., $\beta_i^{DG} = \beta_i^{DL} = \beta_i^{SG} = \beta_i^{SL} = 0$ vs. the alternative that this is not the case. Wald test 2 is the test of the null hypothesis that these four parameters are equal to each other, i.e., $\beta_i^{DG} = \beta_i^{DL} = \beta_i^{SG} = \beta_i^{SL}$, against the alternative that this is not the case. Under the null, both test statistics are asymptotically chi-squared distributed with degrees of freedom equal to four and three, respectively. The critical values at the 5% confidence level are 9.49 and 7.81, respectively. The test results imply that the variables are highly significant. With respect to the age pattern we can conclude that equality of the parameter across the equations cannot be rejected. However, a more appropriate test is whether the joint effects of age and age squared are equal or the same across the four equations. The Wald test 1 and 2 for age and age squared parameters jointly are 379.5 and 82.8 respectively, so the parameters of age structure are jointly significant, and the hypothesis that these four discount rates have the same quadric age structure: $\beta_{age}^{DG} = \beta_{age}^{DL} = \beta_{age}^{SG} = \beta_{age}^{SL}$ and $\beta_{age-sq}^{DG} = \beta_{age-sq}^{DL} = \beta_{age-sq}^{SG} = \beta_{age-sq}^{SL}$ is rejected at the 1% confidence level. This result is important, as it indicates that different discount rates, even as a function of age, should be used for the analysis of different types of economic policies. The same holds for the effect of gender, as the size of the gender effect differs across the four discount rates.

To check whether the quadratic age structure of the rate of time preference is a flexible enough specification, we compared it with two other structures for the age pattern. One is a model with a much more flexible structure for age: a piecewise linear structure with kinks at five year intervals, starting from 16-20 years of age to 86-90 years of age. We calculated 95% confidence bands of the piecewise linear model. The estimated curve of the quadratic model is situated well inside these confidence bands of the piecewise linear model for all ages. This suggests that the quadratic age structure provides a good description. The other age structure we tested is a model with different quadratic age structures for males and females. Based on a likelihood ratio test value of 9.88 with 8 degrees of freedom, we cannot reject the hypothesis that males and females have the same quadratic age structure.

The results of the year dummies imply that there are differences over time, especially for the year 2000. In order to check whether only using a

Table 2.4: Estimation results with only strictly exogenous variables

Variable	δ_{DG}		δ_{DL}		δ_{SG}		δ_{SL}	
	Para.	t-st.	Para.	t-st.	Para.	t-st.	Para.	t-st.
Age	-7.87	-6.53	-5.03	-3.79	-5.82	-3.50	-4.75	-3.59
Age-sq.	4.59	3.59	3.20	2.20	5.62	3.19	1.41	0.99
Female	-0.938	-1.29	-4.92	-6.13	-5.40	-5.24	-1.65	-2.14
Dum. 98	2.49	3.53	2.08	2.59	0.721	0.70	0.163	0.21
Dum. 99	2.85	4.13	3.37	4.04	0.800	0.76	1.27	1.60
Dum. 00	4.39	6.11	5.44	6.51	5.05	4.90	7.23	8.76
Constant	50.0	18.6	7.30	2.51	0.054	0.01	28.7	9.70
Cutoff 1	0.00	-	0.00	-	0.00	-	0.00	-
Cutoff 2	14.5	44.3	16.3	24.7	18.9	30.2	13.3	40.7
Cutoff 3	33.1	89.6	30.6	26.6	39.6	36.4	32.5	68.9
Cutoff 4	45.8	123	42.4	28.7	49.0	42.7	44.4	85.2
Cutoff 5	60.0	-	60.0	-	60.0	-	60.0	-
Loglikelihood				-34198.2				

Variable	Wald Test 1	Wald Test 2
Age	53.5*	5.47
Age-sq.	20.9*	6.98
Female	46.4*	21.8*
Dum. 98	17.9*	7.55
Dum. 99	30.3*	7.61
Dum. 00	113*	9.73*
Constant	375*	172*
Cutoff 2	-	64.1*
Cutoff 3	-	40.4*
Cutoff 4	-	16.9*

* Significant at 5% level for Wald test.

Wald Test 1 is the test of the null hypothesis that four parameters in one row are all zero, Wald Test 2 is the test of the null hypothesis that the four parameters in one row are equal.

time dummy is sufficient to pick up the changing trend or not, we estimated a model allowing for different coefficients in each period and tested equality of parameters over time by using Wald tests. None of the alternative specifications resulted in a significantly better model, so there is no evidence that age and gender parameters change over time. What remains unexplained is the large coefficient for the dummy for the year 2000.

Note that we discretized our data with non-equidistant cutoff levels at 0.0, 7.5, 15, 30, and 60. Looking at the parameter estimates, we find that the distances between the cutoff levels are about equal. In terms of the underlying latent variable, an increase in the discount rate from 7.5% to 15% seems to be about as same large as a change from 15% to 30%, or from 30% to 60%. This indicates that there is some non-linearity in this model, that is easily accounted for in our ordered probit model. We did a likelihood ratio test, and the results show that the three cut-off points we estimated for each equation are significantly different from 7.5, 15, and 30, at the 1% level.

Table 2.5: Standard deviations and correlation coefficients of random effects, η.

	Standard deviations	Correlation coefficients			
		δ_{DG}	δ_{DL}	δ_{SG}	δ_{SL}
δ_{DG}	15.9 (40.6)	1.0			
δ_{DL}	15.1 (23.4)	0.19 (4.65)	1.0		
δ_{SG}	19.3 (21.5)	0.24 (5.82)	0.67 (20.4)	1.0	
δ_{SL}	15.8 (32.1)	0.47 (15.8)	0.36 (9.13)	0.36 (9.04)	1.0

t-statistics in parentheses.

Table 2.6: Standard deviations and correlation coefficients of error terms, ν.

	Standard deviations	Correlation coefficients			
		δ_{DG}	δ_{DL}	δ_{SG}	δ_{SL}
δ_{DG}	17.9 (85.0)	1.0			
δ_{DL}	17.4 (25.2)	0.17 (7.69)	1.0		
δ_{SG}	21.4 (35.7)	0.18 (7.45)	0.33 (13.8)	1.0	
δ_{SL}	20.0 (60.5)	0.24 (13.0)	0.17 (8.00)	0.27 (12.2)	1.0

t-statistics in parentheses.

In the four equations model, the correlations between the error terms and individual effects are also of interest, as they provide insights into the dependencies among the observations for a given respondent, conditional on the explanatory variables. Tables 2.5 and 2.6 present the variances and correlation coefficients of the random effects and error terms. The amount of unexplained variation is of comparable order of magnitude in three of the four scenarios, but substantially larger in the speedup of gains scenario. In all four scenarios, the variance of the individual specific effects is somewhat smaller than the variance of the error terms, implying that there is a substantial amount of unobserved heterogeneity, explaining between 38 and 45% of the nonsystematic variation in the four discount rates.

All correlation coefficients are significant at the 1% level, with values ranging from 0.17 to 0.67. In general, there is a substantial dependency among the observations. Correlations between random effects are always larger than corresponding correlations between error terms, in line with the notion that the errors also capture purely idiosyncratic noise. One would expect that discount rates of gains (δ_{DG} and δ_{SG}) are more similar, as well as the two discount rates of losses (δ_{DL} and δ_{SL}), so we expect relatively high correlation coefficients between these variables. However, the correlation coefficients of random effects and error terms are highest between δ_{DL} and δ_{SG}. This suggests that time preference of speedup of gains (δ_{SG}) has a closer relationship with δ_{DL} than δ_{DG}; in this sense δ_{SG} behaves more like delayed losses (δ_{DL}) than delayed gains (δ_{DG}). Similarly, speedup of losses is more similar to a delayed gain, as δ_{SL} has a higher correlation coefficient with δ_{DG} than with δ_{DL}.

To gain more predictive power, we also included a much larger set of socioeconomic and demographic variables. The estimation results are shown in Table 2.7. Given the large number of explanatory variables, we focus our discussion mainly on those variables that are jointly significant for the four scenarios, i.e., the variables with a significant value for Wald test 1. The interpretation of the size of the coefficients is similar to that in case of the model in Table 4, as the scaling of the problem is the same. The coefficient of -0.79 for females in the delay of gains scenarios, therefore, means that females have a discount rate that is on average about 0.8% point lower than that of males with the same other characteristics.

We get a similar result for the age pattern as in the earlier model, with a significant quadratic structure; the coefficients of age and age-squared are all significant except the coefficient of age-squared in δ_{SL}, just like before. Also the result that females are more patient than males is not affected by including the other variables. The gender difference, however, varies significantly, but also substantially across the four scenarios. For δ_{DG} and δ_{SL} the difference is less than 1% point and insignificant, while for the other two scenarios the difference is more than 4% points and highly significant. Notice that similar to the correlation coefficients in the earlier model, the effects are more similar for the delay of gains and speedup of losses and not for the two

Table 2.7: Estimation results for the full model.

Variable	δ_{DG}		δ_{DL}		δ_{SG}		δ_{SL}		Wald Test	
	Para.	t-st.	Para.	t-st.	Para.	t-st.	Para.	t-st.	1	2
Age	-9.10	-7.14	-5.96	-4.18	-6.71	-3.75	-5.24	-3.70	63.8*	6.70
Age-sq.	5.53	4.01	4.10	2.57	6.94	3.57	1.91	1.23	25.7*	7.43
Female	-0.79	-1.02	-4.76	-5.63	-4.54	-4.25	-0.99	-1.22	36.5*	19.2*
Child	-1.60	-1.98	0.48	0.53	1.37	1.18	-1.15	-1.29	7.64	6.43
School2	-0.93	-1.21	0.91	1.05	1.74	1.52	2.27	2.63	12.5*	11.5*
School3	-0.01	-0.01	2.36	2.42	6.01	4.74	3.68	3.84	33.4*	22.5*
Owner	0.44	0.47	1.03	1.05	3.02	2.28	1.94	2.02	7.45	4.35
Urban.	-0.01	-0.02	-0.19	-0.65	0.02	0.06	-0.06	-0.21	0.57	0.42
Job	0.80	0.91	-1.04	-1.02	0.40	0.30	0.07	0.07	2.55	2.53
Income	-0.42	-1.13	0.38	0.88	-0.56	-1.02	-0.67	-1.60	5.59	4.47
Manag.	-0.87	-1.77	-1.84	-3.16	-1.31	-1.79	-0.82	-1.45	12.9*	2.28
Fin. sit.	-0.76	-1.00	-0.91	-1.06	1.91	1.80	2.25	2.68	15.4*	15.2*
Pri. exp.	2.67	3.00	2.92	2.54	2.52	1.73	5.00	4.75	30.8*	4.45
Time-h.	-0.09	-0.33	-0.24	-0.80	-0.36	-0.89	-0.20	-0.68	1.30	0.40
Smoke	0.39	1.09	-0.08	-0.20	-0.36	-0.71	0.09	0.25	2.16	1.94
Drink	0.78	0.56	0.89	0.59	2.22	1.24	0.26	0.18	1.70	0.95
Que. Ind.	0.34	4.16	0.02	0.20	-0.20	-1.56	0.07	0.73	22.5*	15.6*
Health	-0.33	-0.65	0.46	0.82	-0.19	-0.26	1.47	2.63	10.3*	9.30*
Illness	1.45	1.75	0.64	0.68	0.18	0.15	1.57	1.67	4.57	1.28
Dum. 98	2.30	3.22	2.10	2.59	0.89	0.85	0.13	0.16	15.9*	6.34
Dum. 99	2.67	3.82	3.44	4.03	0.99	0.92	1.24	1.54	27.8*	6.65
Dum. 00	4.21	5.56	5.67	6.39	6.36	5.78	7.86	8.85	116*	13.3*
Constant	48.7	11.4	9.48	1.91	5.13	0.82	19.3	3.92	130*	57.9*
Loglikeli.	-34104.7									

* Significant at 5% level for Wald test.

gains scenarios or the two loss scenarios. However, this pattern is not found for all variables.

Households that can manage financially are more patient than households that cannot, keeping total household income (and other variables) constant. This effect is not different across the four scenarios. Households that spend less than their income tend to have higher discount rates for the speedup scenarios and lower rates for the delay scenarios, although the latter is not significant.

Price expectation has quite a strong positive effect on all four discount rates. It is easy to understand that people with higher inflation expectations also will have higher nominal discount rates. As the effect of inflation is the same for all scenarios, we expect the effect of this variable to be the same across all scenarios. This hypothesis is not rejected.

For three of the four scenarios, we find that higher levels of education imply higher discount rates. This effect is particularly strong for the speedup scenarios, where the highest education level corresponds to an increase in the discount rates with 3.7 to 6% points. As there are two variables related to education level, we also tested the joint significance and equality of all eight parameters of school2 and school3. The Wald test statistics for these hypotheses are 38.2 and 27.4, respectively, implying that both hypotheses are rejected. Note that this result is not in accordance with the theoretical predictions in the literature that higher educated people are more patient.

Although a large number of theoretical and empirical papers suggest a positive relationship between addiction and discount rates (impatience), we do not find such a relationship for smoking and drinking. Obviously, these might not be very severe addictions, which might explain this result.

Two of the three health related variables are significant: the Quetelet index and the measure of general health. More obese people have higher discount rates of δ_{DG}, while better general health positively affects δ_{SL}. This last finding is surprising, as good health makes people live longer and, therefore, they might be expected to be more patient.

The estimates of the cutoff levels and the covariance structure of the error terms and the random effects are similar to the simple model. In general, the variances of the error terms and the random effects are a bit smaller than for the simple model. This is expected as some of the variances are now explained by the additional variables into the model. Given the similarity of the results we do not present the details.

The differences in estimation between the four discount rates should be very useful for policy making. For example, the discount rates of delay of gains and speedup of gains are quite often used. When policy makers want to take into account these two discount rates, they can understand from our model not only that the intercepts are different, but also that a number of variables, like age pattern, gender, education level, and the Quetelet index all have different effects on different discount rates.

2.6 Conclusions

Time preference has a substantial impact on households' responses to all kinds of policy measures. Information on individual rates of time preference will, therefore, help in predicting the effectiveness of such policy measures. As individuals tend to have different discount rates for different types of intertemporal tradeoffs, one needs to investigate the different discount rates simultaneously. For example, a typical finding in the existing literature is the *sign effect*, meaning that gains are discounted at higher rates than losses; not only do we find the *sign effect* for the delaying, like mostly mentioned in the literature, but we also get a *sign effect* for the two discount rates of speeding up, but in the opposite direction: the mean of δ_{SG} is 2.8%, much lower than δ_{SL} (11%). In addition, discount rates for speeding up and delaying gains or losses behave differently, the *delay-speedup asymmetry*: the mean of δ_{SG} is 2.8%, which is much lower than δ_{DG} (21%). A similar finding applies to the two discount rates of losses, δ_{SL} and δ_{DL}, which are also quite different, and again the effect is in the opposite direction, δ_{SL} is much higher than δ_{DL}.

To predict discount rates for each scenario, we estimated a multivariate ordered probit model. Our estimation results indicate that females have lower rates of time preference than males, especially for delay of losses and speedup of gains. For females, δ_{DL} and δ_{SG} are, on average, more than 4% points lower than for males; for δ_{DG} and δ_{SL} this difference is less than 1% points. Income has no significant effect on time preference in all four situations, while education increases discount rates for three of the four scenarios. We find a U-shaped age structure for all four scenarios, in line with the predictions of Becker and Mulligan (1997). For delay of losses δ_{DL} and speedup of gains δ_{SG}, the lowest discount rates are at age 73 and 48, respectively. For these scenarios, we observe that people are really discounting heavier both at young and old ages. However, the lowest points of δ_{DG} and δ_{SL} are found above the age of 80 years. This implies that, in general, discount rates are decreasing with age.

We find significant (all at the 1%) correlations between the random effects and error terms of the four scenarios. An unexpected finding is that the correlation coefficients of the random effects and the error terms are highest between δ_{DL} and δ_{SG}. This suggests that time preference of speedup of gains (δ_{SG}) has a closer relationship with δ_{DL} than δ_{DG}; in this sense δ_{SG} behaves more like delayed losses (δ_{DL}) than delayed gains (δ_{DG}). Similarly, speedup of losses is more like a delayed gain, as δ_{SL} has a higher correlation coefficient with δ_{DG} than with δ_{DL}.

Further research can focus on finding a model to explain why people have asymmetries between time preference of gains and losses, delay and speedup. Models with interaction between members of a household, and causality of time preference and many socioeconomic variables will also be interesting topics for further study. Future research can also focus on the role of time preference in particular policy applications, for example, over-education and the dropout from high school; decisions that are likely to be related to an individuals degree of impatience.

2.7 Appendix to Chapter 2

Explanatory variables

Table 2.8: Variable definitions

Variable	Definition
Age	Age of the individual / 10
Age-sq.	Age-squared of the individual / 1000
Female	Dummy variable for female, 0 no, 1 yes.
Child	Dummy for the presence of children in the household, 1 present, 0 not
School2	Dummy variable for education level 2 (middle), 1 yes, 0 no.
School3	Dummy variable for education level 3 (high), 1 yes, 0 no.
Owner	Dummy variable for homeowner, 1 yes, 0 no.
Urban.	Degree of urbanization of the residence place, 5 levels, 1 very high, 5 very low.
Job	Dummy variable for having a paid job, 1 yes, 0 no.
Income	Total net income of household, 6 levels, 1 is lowest.
Manage	Can you manage the total income of your household? 5 levels, 1 is very hard and 5 is very easy.
Fin. sit.	Dummy variable for financial situation, 1 for those "expenditure were lower than the income".
Pri. exp.	Dummy variable for price expectation, 1 is for prices increasing.
Time-h.	When deciding about what part of the income to spend, and what part to save. Which time-horizon is in your household MOST important with regard to planning expenditures and savings? 1: the next couple of months, 2: the next year, 3: the next couple of years, 4: the next 5 to 10 years, 5: more than 10 years from now

Smoke	Dummy variable for smoking, 1 yes, 0 no.
Drink	Dummy variable for dinking, 1 yes, 0 no.
Que. ind.	Quetelet index; a measure for fatness.
Health	Self-measured general health, 5 levels, 1 poor and 5 excellent.
Illness	Suffer from a long illness, disorder, or handicap, or the consequences of an accident, 1 yes, 0 no.
Dum. 98	Dummy variable for observations in year 1998.
Dum. 99	Dummy variable for observations in year 1999.
Dum. 00	Dummy variable for observations in year 2000.

Table 2.9: Descriptive statistics

Variable	Mean	Std.Dev.	Min	Max
Age	4.77	1.54	1.6	8.7
Age-sq.	2.52	1.50	0.26	7.57
Female	0.433	0.495	0	1
Child	0.437	0.496	0	1
School2	0.307	0.461	0	1
School3	0.315	0.464	0	1
Owner	0.697	0.430	0	1
Urban.	3.02	1.33	1	5
Job	0.592	0.488	0	1
Income	3.74	1.03	1	6
Manag.	3.51	0.799	1	5
Fin. sit.	0.376	0.460	0	1
Pri. exp.	0.888	0.299	0	1
Time-h.	2.20	1.15	1	5
Smoke	0.246	0.419	0	1
Drink	0.078	0.261	0	1
Que. ind.	24.9	3.92	15.6	63.2
Health	3.93	0.715	1	5
Illness	0.260	0.426	0	1
Dum. 98	0.196	0.397	0	1
Dum. 99	0.196	0.397	0	1
Dum. 00	0.226	0.418	0	1

Chapter 3

Reference Points and Loss Aversion in Intertemporal Choice

3.1 Introduction

Reference points and loss aversion are two essential concepts in behavioral economics[1]. After the seminal paper of Kahneman and Tversky (1979) about the reference dependent model for decision-making under uncertainty, prospect theory, Tversky and Kahneman (1991) also proposed a reference-dependent model in riskless choice. The most important idea of reference-dependent theory is a value function with three basic features: (1) reference dependence, gains and losses are defined on deviations from a reference point, rather than on the final level of wealth, (2) loss aversion, the value function is steeper for losses than for gains, losses loom larger than corresponding gains, (3) diminishing sensitivity, the decreasing marginal value for both domains of gains and losses. This kind of value function is quite different from the utility function in traditional expected utility theory. The concepts of reference points and loss aversion came from psychology. As Kahneman and Tversky mentioned in their paper, the reference point is fragile, might depend on expectations, consumption level in previous periods, comparison with others, status quo, etc.; and with the changing of the reference point, a gain could be reframed as a loss, and vice versa (see also Rabin (1998)). The complexity of the reference point makes it quite difficult to use in empirical studies.

As far as we know, no paper estimated the reference point of individuals before, although it is prevalent in economic theories nowadays. In this paper, we propose a structural reference-dependent model for intertemporal choice based on the insight of Loewenstein's (1988) paper. We estimate the reference points of individuals and other important parameters for individual decision-making, in our case the coefficient of loss aversion and the discount rate. Using a Dutch representative household survey, we investigate the distributions of the reference point, loss aversion and discount rate in a population.

Reference points and loss aversion are widely used in models of decision-making under uncertainty since a long time ago (Fishburn 1977, Kahneman and Tversky 1979, Bell 1982, Loomes and Sugden 1982), and are becoming dominant in behavioral economics. For example, a reference-dependent model would give a different prediction of the consumption growth pattern when future income changes, than the traditional Permanent Income Hypothesis (Shea 1995a, 1995b, Bowman, Minehart, and Rabin 1999). A recent paper of Rizzo and Zeckhauser (2003) examined the effects of reference income on the behavior of young male physicians. Their result shows that

[1] See for a comprehensive overview of behavioral economics, Camerer and Loewenstein (2003).

the reference point and loss aversion can explain the data well, but other traditional theories without a reference point and loss aversion have difficulty in doing so. An extensive literature illustrates that there is a big divergence between valuations of willingness-to-pay (WTP) and willingness-to-accept (WTA). The analysis of Kahneman, Knetsch and Thaler (1990) illustrated that reference points and loss aversion might be a possible explanation of this divergency. Bateman et al. (1997) attempted to test a set of predictions derived from the loss aversion hypothesis of reference-dependent theory, by using the data of WTP and WTA from experiments. They showed that divergences between WTP and WTA could be predicted by reference-dependent theory.

Loewenstein introduced the reference point and loss aversion to the analysis of intertemporal choice in 1988. He found that the delay premium is larger than the speed-up cost, and the speed-up cost is larger than the difference between the immediate and delayed consumption prices from several experiments. For example, in one experiment, he found that respondents who expected to receive a video cassette recorder (VCR) one year later would be willing to pay on average \$54 to receive it immediately, but those who expected to receive it immediately demanded on average \$126 to delay its receipt by one year. As explained in his paper, the traditional discounted utility (DU) model could not explain this phenomenon, but a descriptive model with the reference point and loss aversion can solve the problem easily. Furthermore, with the help of a reference point model, Loewenstein and Prelec (1992) could explain two often observed anomalies of time preference: the gain-loss asymmetry and the delay-speedup asymmetry. The gain-loss asymmetry is also called the "sign effect", gains are discounted more than losses. The delay-speedup asymmetry is an asymmetric preference between speeding up and delaying gains or losses. For gains, the amount required to compensate for delaying is much bigger than the amount willing to sacrifice to speed them up. Shelley (1993) employed an experiment to test the reference point model of Loewenstein (1988), and the result is quite consistent with the model.

The main problem to make full use of the reference dependent model is that it is very difficult to observe and estimate the reference point of individual. The reference point might change with expectations, consumption level in previous periods, and comparison with others. From previous studies, we still have no idea about how big the reference point is. Donkers (2000) estimated an intertemporal choice model based on the model proposed by Loewenstein and Prelec (1992). To fix the reference point, he assumed people totally adjusted their reference points to the hypothesized situation, e.g. winning a lottery prize. Four scenarios, delay of gains, delay of losses, speedup of gains, and speed-up of losses, were considered in his paper, but with setting the reference point to one, Donkers (2000) got a negative average discount rate and relatively small coefficient of loss aversion between 1.01 and 1.05. If we employ the ratio of the amount of money people used as the reference divided by the total amount of money involved as the reference point, then generally, it should be a number between zero and one. The result is hard to understand and inconsistent with the literature. The assumption that the reference point equals one might be too strong, as Loewenstein (1988) proposed, the reference point probably could be some number between zero and one.

There are also some papers that estimated the coefficient of loss aversion, but all cases were based on experiments with selective samples. With the shortage of the data used in these studies, we still have no idea about the distribution of loss aversion in the total population or how loss aversion might change with demographics, like age, gender and education. After Kahneman and Tversky's (1979) seminal paper about loss aversion, several papers estimated the coefficient of loss aversion, measured by the ratio of the slope of the utility function in the loss domain and the slope in the gain domain. Kahneman, Knetsch and Thaler (1990) and Tversky and Kahneman (1991) employed an experiment to estimate the coefficient of loss aversion. The subjects were randomly divided into two groups: sellers who were given a mug and had the option to sell it, and choosers who were given the option of receiving the mug or a sum of money. According to their analysis, the different evaluations of the mug were because of loss aversion, the sellers considered it as a loss, while the choosers considered it as a gain. The coefficient of loss aversion they found was slightly greater than two. Pennings and Smidts (2003) based on an investigation of the global shape of the utility function among 332 owner-managers, with 46% of them having an S-shaped utility function, found that the average loss aversion coefficient of those with an S-shaped utility function was 1.8. Schmidt and Traub (2002) used an experiment to test loss aversion in the framework of cumulative prospect theory. They found mixed evidence for loss aversion, with women exhibiting both a more frequent occurrence and a higher extent of loss aversion. Actually loss aversion and reference point should be consider together, it is difficult to get an accurate estimation of the coefficient of loss aversion without consider the reference point carefully. We propose a new method, intertemporal choice, to estimate the coefficient of loss aversion and the reference point at the same time.

In this paper, we consider four often used scenarios in intertemporal choice for eliciting the rates of time preference: delay of gains, delay of losses, speed-up of gains, and speed-up of losses; and we extend Loewenstein's (1988) idea, by reframing the questions in intertemporal choice, we develop a structural model with the reference point and loss aversion for intertemporal choice. We will discuss it in detail in Section 2. With four equations in four scenarios, we can estimate the reference points of delay and speedup, the coefficient of loss aversion, and the discount rate at the same time. The main contribution of our paper is that we propose a new method to estimate the coefficient of loss aversion in intertemporal choice without risk[2], and our paper might be the first time in literature attempting to estimate the reference points of individuals.

After constructing a structural model with reference points and loss aversion in an intertemporal choice setting, we employ a rich panel dataset from a Dutch representative household survey over the years 1997-2002, to estimate the coefficients involved in decision making in intertemporal choices simultaneously, namely, the reference points of delay and speedup, the coefficient of loss aversion and the discount rate at the same time, and we can also investigate their relationships with some demographic variables. Our findings show that the average coefficient of loss aversion is around two, in line with other experimental findings. We also find that the reference point of delay is larger than that of speedup, and reference points are between zero and one, all consistent with Loewenstein's (1988) hypothesis. We find that females are more loss averse than males, and high-education and aging make people less loss-averse. For the discount rate, we find that high-educated and older people are more patient. The observed relationships of these parameters could be used to better understand and predict the behavior of households for policy purposes, such as households' decisions of investment, saving and pension. Loss aversion, reference points, and the discount rate are all main factors that affect these decisions of households.

The remainder of this paper is organized as follows. In Section 2, we introduce a structural model of time preference with the reference points and loss aversion. Then in Section 3 we use the model to explain some anomalies of time preference prominent in the literature. We describe the data in Section 4. In Section 5, we discuss the econometric model and the estimation procedure. In Section 6, the results of the models are presented. Section 7 concludes.

---

[2] With risk situation, loss aversion may be different. This paper only discusses the situation without risk.

3.2 Economic Model

In this paper, we consider the four different scenarios of time preference: delay of gains, speedup of gains, delay of losses, and speedup of losses, which are commonly used in the literature. According to the traditional DU model, individuals should have discount rates that are the same in all four scenarios. In this section, we first describe the four questions for the four scenarios, which we use to elicit four discount rates. Then we discuss the reference point model for intertemporal choice as proposed by Loewenstein (1988), and present a structural model with reference points and loss aversion to explain the discount rates in these four scenarios.

We consider "win a prize in the National Lottery" as a gain, and "pay a tax" as a loss. We select four questions with the amount of money equal to Dfl. 1000[3] and a time horizon of one year[4]. Therefore, the questions about the four different scenarios we consider are the following:

Delay of gains

> Imagine you win a prize of Dfl. 1000 in the National Lottery. The prize is to paid out today. Imagine, however, that the National Lottery asks if you are prepared to wait A YEAR before you get the prize of Dfl. 1000. There is no risk involved in this wait. How much extra money would you ask to receive AT LEAST to compensate for the waiting term of a year? If you agree on the waiting term without the need to receive extra money for that, please type 0 (zero).

Speedup of gains

> Imagine again that you receive a notice from the National Lottery that you have won a prize worth Dfl. 1000. The money will be paid out after A YEAR. The money can be paid out at once, but in that case you receive less than Dfl. 1000. How much LESS money would you be prepared to receive AT MOST if you would get the money at once instead of after a year? If you are not interested in receiving the money earlier or if you are not prepared to receive less for getting the money earlier, please type 0 (zero).

Delay of losses

[3] 1 Dfl. $\approx$ 0.45 Euro.

[4] We will discuss the data in detail in the next section.

> Imagine again that you have to pay a tax assessment of Dfl. 1000 today. Suppose that you could wait A YEAR with settling the tax assessment. How much extra money would you be prepared to pay AT MOST to get the extension of payment of A YEAR? If you are not interested in getting an extension of payment or if you are not prepared to pay more for the extension of payment, please type 0 (zero).

Speedup of losses

> Imagine again that you receive a tax assessment of Dfl. 1000. The assessment has to be settled within A YEAR. It is, however, possible to settle the assessment now, and in that case you will get a REDUCTION. How much REDUCTION would you like to get AT LEAST for settling the assessment now instead of after a year? If you are not interested in getting a reduction for paying early or if you think there is no need to get a reduction for paying early, please type 0 (zero).

Each of these four questions leads to a different discount rate, providing discount rates for the delay of gains (δ_{DG}), speedup of gains (δ_{SG}), delay of losses (δ_{DL}), and speedup of losses (δ_{SL}) respectively. We use x_{DG}, x_{SG}, x_{DL}, x_{SL} to represent the answer of each question above, and compute these four discount rates as follows:

$$\begin{aligned}\delta_{DG} &= \frac{x_{DG}}{1000}\\ \delta_{SG} &= \frac{x_{SG}}{1000 - x_{SG}}\\ \delta_{DL} &= \frac{x_{DL}}{1000}\\ \delta_{SL} &= \frac{x_{SL}}{1000 - x_{SL}}\end{aligned}$$

With its simple and elegant structure, the traditional DU model is the basic way to analyze intertemporal choice. If the individual's utility function is $v(\cdot)$, and $\delta(t)$ is her discount factor, then the present value of utility at time t of two-period consumption is $V = v(c_t) + \delta(1)v(c_{t+1})$, where c_t and c_{t+1} are consumption levels in periods t and $t+1$ respectively. Normally we assume that people have positive time preference that is $\delta(t) < 1$, and people will discount more for a longer period, that is $\delta(t) < \delta(t')$ if $t > t'$.

By introducing the reference point, Loewenstein (1988) wrote the present value of utility of the above problem at time t in a different way, $V =$

$v(c_t - R_t) + \delta(1)v(c_{t+1} - R_{t+1})$, where $v(0) = 0$ and R_t and R_{t+1} are the reference points of the individual in period t and $t+1$ respectively. The DU model becomes a special case when the reference points are equal to zero, $R_t = R_{t+1} = 0$.

In this paper, we follow the insight of Loewenstein's (1988) reference point model, utility of outcome C depends on the reference point R, $V(C) = v(C - R)$, $v(\cdot)$ is the value function. We also normalize the utility of the reference points to zero, that is $v(0) = 0$. In order to make the model analytically tractable, we use a piecewise linear value function like many other studies: for gains ($x \geq 0$), the utility of $x + R$ is $V(x+R) = v(x) = x$. We also use a simple structure the utility of a negative outcome $-x + R$, (compared to the reference point R), $V(-x+R) = v(-x) = -\lambda v(x) = -\lambda x$, with $-x < 0$. Note that the outcome $-x + R$ we assume is considered as a loss compared to the reference point R, even if $-x+R$ is a positive outcome. Here, λ is the coefficient of loss aversion, and we have $\lambda > 0$. Loss aversion means $\lambda > 1$, that is the disutilities of losses are larger than utilities of corresponding gains, $-v(-x) > v(x)$. We denote with d the discount rate for utility in the next period, so the present value of x in the next period becomes $v_1(x) = \frac{1}{1+d}v(x) = \frac{1}{1+d}x$, when $x \geqslant 0$. For a loss, similarly, we have $v_1(-x) = \frac{1}{1+d}v(-x) = -\frac{\lambda}{1+d}x$, where $-x < 0$ is a loss.

We can retrieve the traditional DU model as a special case of our model when $\lambda = 1$ and $R = 0$; this means that without loss aversion ($\lambda = 1$) or with reference point equal to zero ($R = 0$, like the model without considering the reference point), our model is the same as the traditional DU model, predicting that people should have the same discount rates d for all four scenarios. We now consider the predictions of this model for the four scenarios separately.

3.2.1 Delay of Gains

For the question of "delay of gains", the hypothesized situation is winning a prize of Dfl. 1000 today. It seems reasonable to assume that the prize today might affect the reference point today but not for next year. Then the reference points of today and next year could be set as "to get Dfl. R_D today and zero next year". In previous studies like Loewenstein (1988), Shelly (1992) and Donkers (2000), the reference points were always equal to the amount of money involved, that is $R_D = 1000$ for our situation. The answer of the question should be interpreted as that people make a choice between the following two: original choice a) is "to receive Dfl. 1000 today"; alternative choice b) is to postpone the gain, wait and receive more later, Dfl. $1000 + x_{DG}$, in next year. We introduce a notation (x_1, x_2) to present it, where x_1 is the outcome today and x_2 is the outcome next year. A positive number is a gain, and a negative number is a loss; x_1 and x_2 can both be either gains or losses.

Taking "to get Dfl. R_D today and zero next year" as the reference point, it should be $(R_D, 0)$ with the new notation, the choice a) becomes to "gain Dfl. $1000 - R_D$ today and zero next year", $(1000 - R_D, 0)$, compared to the reference point; and the choice b) becomes to "lose Dfl. R_D today and gain Dfl. $1000 + x_{DG}$ next year", $(-R_D, 1000 + x_{DG})$, compared to the reference point. Generally, we can expect that R_D is a number between 0 and 1000[5]. From now on, we use the ratio $r_D = R_D/1000$, which should be a number between 0 and 1, as the reference point. Reframing of the outcome for the questions in other three scenarios is done similarly.

Choosing of x_{DG} means that these two alternatives have the same utility levels. We can write this as the following equation:

$$\begin{aligned}
V(1000) + \frac{1}{1+d}V(0) &= V(0) + \frac{1}{1+d}V(1000 + x_{DG}) \\
v(1000 - R_D) + \frac{1}{1+d}v(0-0) &= v(0 - R_D) + \frac{1}{1+d}v(1000 + x_{DG} - 0) \\
1000 - R_D &= -\lambda R_D + \frac{1000 + x_{DG}}{1+d} \\
1000 + x_{DG} &= (1000 - R_D + \lambda R_D)(1+d) \\
1 + \delta_{DG} &= \frac{1000 + x_{DG}}{1000} = (1 + (\lambda - 1)r_D)(1+d)
\end{aligned}$$

where δ_{DG} is the discount rate of delay of gains, λ is the coefficient of loss aversion, r_D is the reference point (the ratio $R_D/1000$) of delay, d is the discount rate for future utility.

Here, $r_D = 1$ and $r_D = 0$ are two special cases which are often used in the literature. For $r_D = 1$, we have $\delta_{DG} = \lambda(1+d) - 1$. With loss aversion, $\lambda > 1$ and non-zero reference point, we have $\delta_{DG} > d$ for all $r_D > 0$, an individual will discount delay of gains quite heavily, with δ_{DG} increasing in r_D and λ. When $r_D = 0$ or $\lambda = 1$, we have $\delta_{DG} = d$, as in the DU model.

3.2.2 Speed-up of Gains

For the question of "speed-up of gains", the hypothesized situation is winning a prize of Dfl. 1000 next year. We assume that the reference point is "to

[5] If R_D is equal to zero, receiving nothing, it means that the respondant has not adjusted to the new situation at all; and equal to 1000, it means that the respondant has fully adjusted to the hypothesized new situation.

get zero today and Dfl. R_S next year", $(0, R_S)$. People choose x_{SG} such that utilities of the following two choices are equal: a) wait for payment Dfl. 1000 next year, $(0, 1000)$; b) now receive less: Dfl. $1000 - x_{SG}$, $(1000 - x_{SG}, 0)$. Taking $(0, R_S)$ as the reference point, the choice a) becomes to "gain 0 today and Dfl. $1000 - R_S$ next year", $(0, 1000 - R_S)$, compared to the reference point; and the choice b) becomes to "gain Dfl. $1000 - x_{SG}$ today, and lose Dfl. R_S next year", $(1000 - x_{SG}, -R_S)$, compared to the reference point. We can write this as the following equation:

$$V(0) + \frac{1}{1+d}V(1000) = V(1000 - x_{SG}) + \frac{1}{1+d}V(0)$$

resulting in

$$1 + \delta_{SG} = \frac{1000}{1000 - x_{SG}} = \frac{1}{1 + (\lambda - 1)r_S}(1 + d)$$

where δ_{SG} is the discount rate of speed-up of gains, r_S is the reference point (the ratio $R_S/1000$) of speedup.

For the special case $r_S = 1$, we have $\delta_{SG} = \frac{1}{\lambda}(1 + d) - 1$. It is quite possible that we have $\lambda > 1 + d$, then δ_{SG} could be a negative number. With loss aversion, $\lambda > 1$, an individual will discount speedup of gains slightly, we have $\delta_{SG} < d$ for all $r_S > 0$, and δ_{SG} is decreasing in r_S if $r_S > 0$. For the case of $r_S = 0$ or $\lambda = 1$, we have $\delta_{SG} = d$, as in the DU model.

3.2.3 Delay of Losses

For the question of "delay of losses", the hypothesized situation is a requirement to pay for taxes Dfl. 1000 today. We assume paying a taxes is a loss for individuals, and the reference point is "to lose Dfl. R_D today and zero next year", $(-R_D, 0)$. People choose x_{DL} such that utilities of the following two choices are equal: a) to lose Dfl. 1000 today, $(-1000, 0)$; b) to lose money next year: Dfl. $1000 + x_{DL}$, $(0, -1000 - x_{DL})$. Taking $(-R_D, 0)$ as the reference point, the choice a) becomes to "loss Dfl. $1000 - R_D$ today and zero next year", $(-1000 + R_D, 0)$, compared to the reference point; and b) becomes to "gain Dfl. R_D today, and lose Dfl. $1000 + x_{DL}$ next year", $(R_D, -1000 - x_{DL})$, compared to the reference point. We can write this as the following equation:

$$V(-1000) + \frac{1}{1+d}V(0) = V(0) + \frac{1}{1+d}V(-(1000 + x_{DL}))$$

resulting in

$$1 + \delta_{DL} = \frac{1000 + x_{DL}}{1000} = (1 - \frac{\lambda - 1}{\lambda} r_D)(1 + d)$$

where δ_{DL} is the discount rate of delay of losses.

For the special case $r_D = 1$, we have $\delta_{DL} = \frac{1}{\lambda}(1 + d) - 1$, and it is quite possible that $\lambda > 1 + d$, then people even have a negative discount rate of δ_{DL}. With loss aversion, $\lambda > 1$, an individual will discount for delay of losses slightly, we have $\delta_{DL} < d$ for all $r_D > 0$, and δ_{DL} is decreasing in r_D if $r_D > 0$. When $r_D = 0$ or $\lambda = 1$, we have $\delta_{DL} = d$, as in the DU model.

3.2.4 Speed-up of Losses

For the question of "speed-up of losses", the hypothesized situation is a requirement to pay for taxes Dfl. 1000 next year. We assume that the reference point is "to lose zero today and Dfl. R_S next year", $(0, -R_S)$. People choose x_{SL} such that utilities of the following two choices are equal: a) lose Dfl. 1000 next year, $(0, -1000)$; b) lose less today: Dfl. $1000 - x_{SL}$, $(-1000 + x_{SL}, 0)$. Taking $(0, -R_S)$ as the reference point, the choice a) becomes to "loss zero today and Dfl. $1000 - R_S$ next year", $(0, -1000 + R_S)$, compared to the reference point; and b) becomes to "lose Dfl. $1000 - x_{SL}$ today and gain Dfl. R_S next year", $(-1000 + x_{SL}, R_S)$, compared to the reference point. We can write this as the following equation:

$$V(0) + \frac{1}{1 + d} V(-1000) = V(-(1000 - x_{SL})) + \frac{1}{1 + d} V(0)$$

resulting in

$$1 + \delta_{SL} = \frac{1000}{1000 - x_{SL}} = \frac{1}{1 - \frac{\lambda - 1}{\lambda} r_S}(1 + d)$$

where δ_{SL} is the discount rate of speed-up of losses.

For the special case $r_S = 1$, $\delta_{SL} = \lambda(1+d) - 1$. With loss aversion, $\lambda > 1$, an individual will discount delay of losses heavily, we have $\delta_{SL} > d$ for all $0 < r_S < 1$, and δ_{SL} is increasing in r_S. When $r_S = 0$ or $\lambda = 1$, we have $\delta_{SL} = d$, as in the DU model.

3.2.5 Summary of the Model

We now summarize the way the outcomes of delay of gains, speedup of gains, delay of losses, and speedup of losses are framed in Table 3.1.

Table 3.1: Summary of outcomes with and without considering the reference point

Discount rate	Delay of gains	Speed-up of gains
Reference point	$(R_D,\ 0)$	$(0,\ R_S)$
Without considering the reference point		
Choice a	$(1000,\ 0)$	$(0,\ 1000)$
Choice b	$(0,\ 1000 + x_{DG})$	$(1000 - x_{SG},\ 0)$
With considering the reference point		
Choice a	$(1000 - R_D,\ 0)$	$(0,\ 1000 - R_S)$
Choice b	$(-R_D,\ 1000 + x_{DG})$	$(1000 - x_{SG},\ -R_S)$
Discount rate	**Delay of losses**	**Speed-up of losses**
Reference point	$(-R_D, 0)$	$(0, -R_S)$
Without considering the reference point		
Choice a	$(-1000, 0)$	$(0, -1000)$
Choice b	$(0, -1000 - x_{DL})$	$(-1000 + x_{SL}, 0)$
With considering the reference point		
Choice a	$(-1000 + R_D, 0)$	$(0, -1000 + R_S)$
Choice b	$(R_D, -1000 - x_{DL})$	$(-1000 + x_{SL}, R_S)$

Note: Choice a) is the hypothesized situation, choice b) is the alternative.
Positive number is a gain, and negative one is a loss.

For this structural model, we have four equations for four scenarios. From our data, we can compute the four discount rates for each person from the reported values x_{DG}, x_{SG}, x_{DL} and x_{SL}, so we can solve and identify the model with four unknown coefficients. Considering the equations with four unknowns, actually we have two choices: one is with the same discount rate d for gains and losses, but with different reference points r_D for delay and r_S for speed-up, that is the model we just presented. The second one is with different discount rates for gains d_G and losses d_L, but with the same reference point r for all four scenarios[6]. We want to test the hypothesis of Loewenstein (1988): whether the reference point of delay is larger than that of speedup, we will only discuss and estimate the first model, which has different reference points for delay and speedup but same discount rate for gains and losses, hereafter.

Loewenstein (1988) proposed that the reference points might be smaller than one and larger in the delay conditions than in the speedup conditions, that is, $1 > r_D > r_S > 0$. This might be the reason that people are more familiar with the delay than with the speedup situation, and in the speedup scenario r_S is more distant in the future (next year) and less vivid than the delay scenario r_D (today). We can test this hypothesis in our model with four coefficients, which has different reference points r_D for delay and r_S for speedup. This model can be written as follows:

$$
\begin{aligned}
1+\delta_{DG} &= (1+(\lambda-1)r_D)(1+d) \\
1+\delta_{SG} &= \frac{1}{1+(\lambda-1)r_S}(1+d) \\
1+\delta_{DL} &= (1-\frac{\lambda-1}{\lambda}r_D)(1+d) \\
1+\delta_{SL} &= \frac{1}{1-\frac{\lambda-1}{\lambda}r_S}(1+d)
\end{aligned}
$$

As a starting point, we will estimate a model with only three coefficients: loss aversion λ, the same discount rate d for gains and losses, and the same reference point r for all these four scenarios. This model with three coefficients is a special case when $r_D = r_S$.

[6] Actually, we could solve these two models analytically, four equations with four unknown parameters, and calculate the four parameters for each individual directly. But the problem is that for about 40% of the observations we get a negative value of loss aversion, and reference points less than zero or larger than one. That is difficult to understand. With the help of panel data, we can use a statistic model with random effects instead of an analytic one in our paper. We will discuss it in more detail in the next section.

3.3 Explanation of Some Anomalies with the Model

In this section, first we will show some results of our model, and then use these results to explain some of the anomalies prevalent in the literature. In the literature on time preference, most of the studies employed experiments to collect data on discount rates, mainly focusing on differences of some factors, which should be expected to have no effect on the discount rates according to the traditional economic theories. For example, many empirical studies of time preference find anomalies that contradict the assumption underlying the DU model, which the discount rate should be constant for all goods and all time periods. It is obvious that our model can explain some anomalies of time preference prominent in the literature, see Loewenstein and Prelec (1991, 1992), Shelley (1993), and Frederick, Loewenstein and O'Donoghue (2002) for more detailed discussion.

With the reference points equal to zero, $r_D = r_S = 0$, or without loss aversion, that is with $\lambda = 1$, we have $\delta_{DG} = \delta_{SG} = \delta_{DL} = \delta_{SL} = d$, as in the DU model. Thus, it is clear that the DU model is a special case of our model.

Figure 3.1 shows how δ_{DG} and δ_{DL} change with r_D, and how δ_{SG} and δ_{SL} change with r_S. Here, we assume $\lambda = 2$ and $d = 10\%$, which are quite reasonable compared to findings in literature. From the figure we can see, that δ_{DG} and δ_{SL} will increase, and δ_{DL} and δ_{SG} will decrease with r_D or r_S monotonically when r_D and r_S are numbers between zero and one. We have $\delta_{DG} > \delta_{SL} > d > \delta_{DL} > \delta_{SG}$ for any $0 < r_D = r_S < 1$. This is in line with our data and our estimation result presented in Section 6.

One might expect the two discount rates for gains, δ_{DG} and δ_{SG}, or two discount rates for delay, δ_{DG} and δ_{DL} might be more close linked. However our model implies that δ_{DG} and δ_{SL}, δ_{DL} and δ_{SG} have the same trends increasing and decreasing with reference points respectively. Thus, we expect δ_{DG} to be more close linked with δ_{SL} other than δ_{DL} and δ_{SG}. Same for δ_{DL}, we expect it is more close linked with δ_{SG} other than δ_{DG} and δ_{SL}.

With the help of this structural model, we can explain some anomalies of time preference prominent in the literature.

3.3.1 The sign effect

One anomaly of time preference dominant in the literature is the sign effect: gains are discounted at a higher rate than losses, in studies with the discount rates of gains and losses estimated at the same time. According to traditional

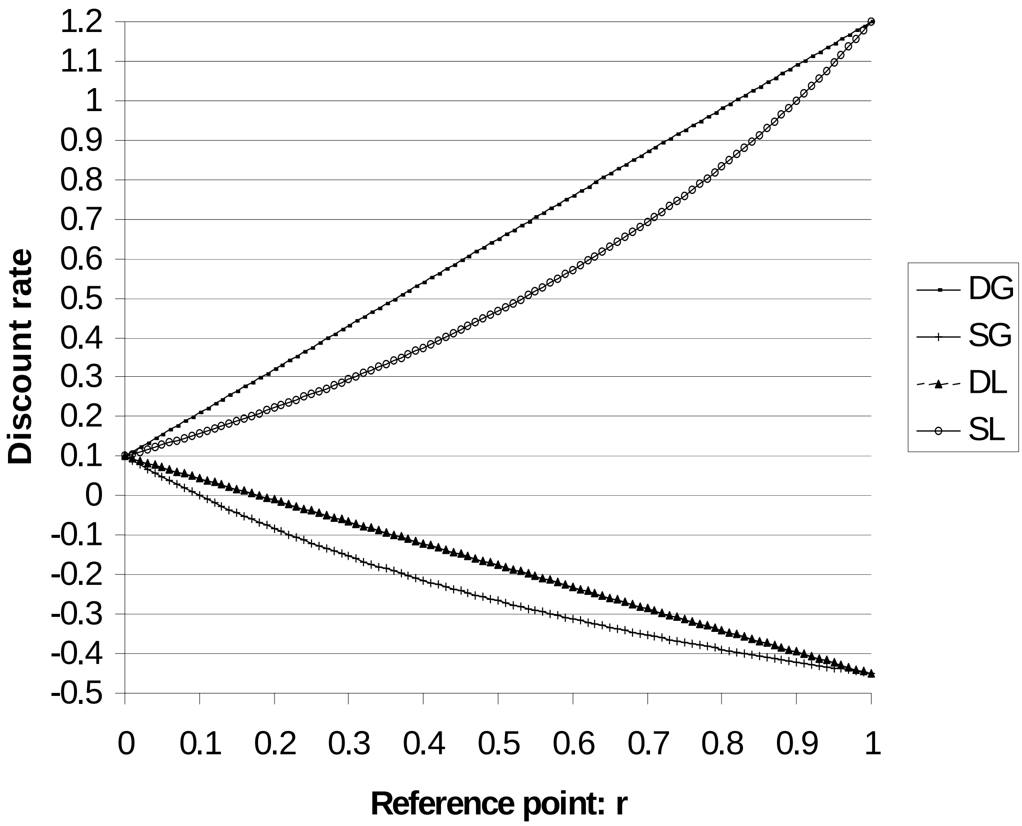


Figure 3.1: δ_{DG}, δ_{SG}, δ_{DL}, δ_{SL} change with r: $\lambda = 2$ and $d = 0.1$

economic theory[7], one should be willing to pay a similar amount to receive \$100 a month later or to postpone paying \$100 for a month. That is, gains and losses should be discounted equally.

In the literature, the discount rate of gains is the delay of gains δ_{DG}; and the discount rate of losses is the delay of losses δ_{DL}. The sign effect means $\delta_{DG} > \delta_{DL}$. Using the simple model with loss aversion and reference points we presented here, we can easily explain the sign effect. From the model above, we always have $1+\delta_{DG} = (1+(\lambda-1)r_D)(1+d) > 1+d > 1+\delta_{DL} = (1-\frac{\lambda-1}{\lambda}r_D)(1+d)$, that implies $\delta_{DG} > d > \delta_{DL}$, if we assume that $\lambda > 1$, and $0 < r_D \leq 1$. For example, if we assume $\lambda = 2$ and $d = 10\%$, then $\delta_{DG} = 0.1 + 1.1r_D > d = 0.1 > \delta_{DL} = 0.1 - 0.55r_D$; so even r_D is a small number like 0.1, then $\delta_{DG} = 0.21 > d = 0.1 > \delta_{DL} = 0.045$. We thus have the result that δ_{DG} is much bigger than δ_{DL}: the sign effect.

3.3.2 The "Delay-Speedup" Asymmetry

In the literature, the "delay-speedup" asymmetry means people demand more to delay a gain than they would like to pay for speeding it up, that is $\delta_{DG} > \delta_{SG}$. It is also obvious that $1 + \delta_{DG} = (1 + (\lambda - 1)r_D)(1 + d) > 1 + d > 1 + \delta_{SG} = \frac{1}{1+(\lambda-1)r_S}(1 + d)$, if $\lambda > 1$, r_D and r_S are between zero and one. That indicates $\delta_{DG} > d > \delta_{SG}$ in our model with reference points and loss aversion, without special assumption for the relationship of r_D and r_S.

Consider the situation with $\lambda = 2$ and $d = 10\%$, then we have $\delta_{DG} = 0.1 + 1.1r_D > d = 0.1 > \delta_{SG} = \frac{0.1-r_S}{1+r_S}$, if $0 < r_i < 1$, i=delay, speedup. Thus, our model can explain the "delay-speedup" asymmetry well.

3.3.3 Negative Discount Rate

The negative discount rates, which are sometimes observed in literature, are very difficult to understand in traditional theory, see Loewenstein and Prelec (1992) for more discussion. In our model, δ_{DG} and δ_{SL} should always be positive, but δ_{SG} and δ_{DL} can be negative, it can be used to explain why the negative discount rates sometimes can appear.

From the Figure , we can find that δ_{SG} and δ_{DL} are easily become negative numbers if the reference points are large than 0.3. For example, if $r_D = r_S = 1$, as many studies in the literature assumed, then $\delta_{DG} = 120\%$, $\delta_{SG} = -45\%$, $\delta_{DL} = -45\%$, and $\delta_{SL} = 120\%$. From the model, we can get negative discount rates of δ_{SG} and δ_{DL} when reference points are relatively high.

[7] Without loss aversion.

3.3.4 Hyperbolic Discounting

In the literature, the first dominant anomaly is the finding that discount rate declines with the time horizon, i.e., the discount rate over longer time horizon is lower than that over a shorter time horizon. This anomaly is called hyperbolic discounting, which makes a hyperbolic functional form fit data better, as observed in many studies, Benzion, Rapoport, and Yagil (1989), Chapman (1996), Chapman and Elstein (1995), Pender (1996), Redelmeier and Heller (1993), Thaler (1981). This might lead to preference reversal, an example of dynamically inconsistent behavior. A famous example is from Thaler (1981), who found that a person might prefer one apple today to two apples tomorrow, but at the same time would prefer two apples in 51 days to one apple in 50 days.

Our model with reference points and loss aversion might be an alternative explanation for this finding, not requiring hyperbolic discounting. When a gain is very close, for example one apple today, a person can image it easily; if she wants it eagerly, she will feel a real loss if she can not get it today. So the reference point when she makes decision for one apple today or two apples tomorrow could be quite high, close to one. For example, we can assume that $\lambda = 2$ and $d = 1\%$, if $r_D > 0.98$, then $\delta_{DG} > 100\%$, she will choose one apple today. But for choices of 51 days and 50 days, the time is quite far away, it is reasonable to assume that she has a smaller reference point compared to choose between today and tomorrow, and if $r_D < 0.98$, she will choose two apples in 51 days. If the coefficient of loss aversion is a little bit large, δ_{DG} can be larger than 100% with small r_D even with $d = 0$. For a person with $\lambda = 3$ and $d = 0$, if $r_D > 0.5$, then $\delta_{DG} > 100\%$, she might prefer one apple today over than two apples tomorrow; and if she uses $r_D < 0.5$ when she makes the second decision, she will prefer two apples in 51 days than one apple in 50 days. We can see it clear, not because of the changing of discount rate makes people have a dynamically inconsistent behavior, but the changing of feeling of loss and loss aversion make her reverse her preference.

3.4 Data

The data we use is a panel data set with six waves (1997-2002) taken from the CentER Savings Survey (CSS). The CSS is a large Dutch household survey starting from 1993, collected every year of more than 1500 households. The CSS is a rich data set containing information on employment status, pensions, accommodation, mortgages, income, assets, debts, health, economic and psychological concepts, and personal characteristics. Our data constitute an unbalanced panel, with a total of 5,480 individuals and 11,847 observations. Table 4.1 shows the structure of this unbalanced panel. The average time an individual stayed in the panel is 2.2 years.

Table 3.2: Structure of the panel

By wave		By number of waves		
Year	Observations	Number of waves	Observations	Number of individuals
1997	2,659	1	2,498	2,498
1998	1,365	2	2,716	1,358
1999	1,368	3	2,247	749
2000	1,783	4	1,160	290
2001	2,467	5	1,420	284
2002	2,205	6	1,806	301
Total	11,847	Total	11,847	5,480

Left panel: Year is the survey year, in total we have four waves of the survey, 1997-2002.
Right Panel: Number of waves the households stay in the panel.

Starting from the year 1997, a detailed set of questions about time preference is included in the CSS.[8] In total, there are sixteen questions about the way people value opportunities in the future compared to the present. These questions differ in four aspects with each aspect having two levels, resulting in a total of sixteen questions. The first aspect is the amount of money concerned, either Dfl. 1000 or Dfl. 100,000[9]. The second is the time horizon, either three months or one year. The third is whether the amount of money is to be received or to be paid[10]. The last one is speedup or delay of the receipt or payment of the money.

Some descriptive statistics on these discount rates and demographic variables are provided in Table 4.2. We only use observations with a discount rate of at most 100% to compute the descriptive statistics of discount rates. We use a Tobit model with right censoring at $\delta \leq 1.0$ for each equation to estimate our model.

According to the traditional DU model, in case of perfect financial markets without constraints[11], a "rational" individual should have the same dis-

[8] In earlier waves time preference was elicited with questions that differ in the answering format. We therefore do not use these data. In the wave of 2003, these questions were deleted from the survey.

[9] 1 Dfl. $\approx$ 0.45 Euro.

[10] We consider "a tax assessment need to pay" as a loss and "win a prize of the National Lottery" as a gain.

[11] Our data includes information on savings accounts. More than 96% of the individuals

Table 3.3: Descriptive statistics for the discount rates and demographic variables.

Variable	Mean	Std. Dev.	Min	Max	Median	Number of obs. with $\delta > 1$
δ_{DG}	0.212	0.246	0	1	0.100	138
δ_{SG}	0.030	0.080	0	1	0.0	5
δ_{DL}	0.036	0.085	0	1	0.0	5
δ_{SL}	0.121	0.176	0	1	0.081	11
Age/10	4.67	1.53	1.6	9.5	4.6	
Age-squared/1000	2.42	1.49	0.26	9.03	2.12	
Female	0.445	0.497	0	1		
School2	0.329	0.461	0	1		
School3	0.378	0.476	0	1		

Note: the mean and std. dev. of δ are computed with the obs. which δ is smaller or equal than 1

count rate for all four scenarios, implying that δ_{DG}, δ_{SG}, δ_{DL} and δ_{SL} have similar values on average. As expected from the existing literature, our data do not support this prediction of traditional DU model. Instead, one can see from the table 4.2 that our data are in line with the findings in the literature. First of all, people discount gains heavier than losses, i.e., the mean of δ_{DG} (the discount rate of delay of gains) is more than five times larger than that of δ_{DL} (the discount rate of delay of losses); this is what is called the "sign effect". The second confirmation of existing findings is the asymmetry of delay and speedup, which states that δ_{DG} is much bigger than δ_{SG} (the discount rate of speedup of gains). However, we also find that δ_{DL} is much smaller than δ_{SL} (the discount rate of speedup of losses).

From table 4.2, we find for the means of the four discount rates that $\delta_{DG} > \delta_{SL} > \delta_{DL} > \delta_{SG}$. We use t-tests to check the inequality of the means, the differences between the four discount rates are all highly significant. This is exactly the prediction from our economic model. In order to verify the inequality of the means carefully, we also compute the percentage of observations which are consistent with $\delta_{DG} \geq \delta_{SL} \geq \delta_{DL} \geq \delta_{SG}$ by individual data. We also compute the correlation coefficients between different discount rates. The results are presented in table 3.4. The upper right part of the table 3.4 is about the relations, which are percentages of observations who have $\delta_i \geq \delta_j$, $i, j = DG, SL, DL, SG$. For example, 65% observations

had more than Dfl 1,000 in their bank account. This means that these people are not likely to be financially constrained for the amount of Dfl 1,000 we investigate.

have $\delta_{DG} \geq \delta_{SL}$, and 93% observations have $\delta_{SL} \geq \delta_{SG}$. All these numbers are much bigger than 50%, means that this inequality also stands for individual data. The lower left part of the table 3.4 are correlation coefficients of the four discount rates. Our result indicates that the discount rate of delay of gains δ_{DG} is more close linked with the discount rate of δ_{SL} rather than with δ_{DL} or δ_{SG}, δ_{DL} is more close linked with δ_{SG} rather than with δ_{DG} or δ_{SL}. This result is consistent with our expectation.

Table 3.4: Relations and correlation coefficients of the four discount rates

δ_i	δ_j			
	δ_{DG}	δ_{SL}	δ_{DL}	δ_{SG}
δ_{DG}	-	65%	97%	91%
δ_{SL}	0.26	-	90%	93%
δ_{DL}	0.16	0.13	-	73%
δ_{SG}	0.09	0.14	0.29	-

Note: the upper right part is the percentage of obs. who have $\delta_i \geq \delta_j$
the lower left part is the correlation coefficients between δ_i and δ_j

3.5 Econometric Model

In this section, we present an econometric model, a nonlinear random coefficients model with panel data, to estimate the coefficients in the structural model. The main advantage of a random coefficients model is that we can examine the variation of the coefficients we are interested in, like the coefficient of loss aversion λ, the reference points of delay r_D and speedup r_S, and the discount rate d across individuals, while still allowing for estimation of the overall mean effects. We can write the nonlinear structural model as follows:

$$\begin{aligned}
y_{it}^{DG} &= 1+\delta_{DGit} = (1+(\lambda_{it}-1)r_{Dit})(1+d_{it}) + e_{it}^{DG} \\
y_{it}^{SG} &= 1+\delta_{SGit} = \frac{1}{1+(\lambda_{it}-1)r_{\mathrm{Si}\,t}}(1+d_{it}) + e_{it}^{SG} \\
y_{it}^{DL} &= 1+\delta_{DLit} = (1-\frac{\lambda_{it}-1}{\lambda_{it}}r_{Dit})(1+d_{it}) + e_{it}^{DL} \\
y_{it}^{SL} &= 1+\delta_{SLit} = \frac{1}{1-\frac{\lambda_{it}-1}{\lambda_{it}}r_{\mathrm{Si}\,t}}(1+d_{it}) + e_{it}^{SL} \\
i &= 1,\cdots,N,\ t=1,...,6
\end{aligned}$$

We assume that all four coefficients in the structural model are random and that they are linear functions of other demographic variables, that is

$\lambda_{it} = x_{it}\beta_\lambda + \varepsilon_i^\lambda$, $r_{Dit} = x_{it}\beta_{r_D} + \varepsilon_i^{r_D}$, $r_{Sit} = x_{it}\beta_{r_S} + \varepsilon_i^{r_S}$, and $d_{it} = x_{it}\beta_d + \varepsilon_i^d$. Here x_{it} is the vector of demographic variables of individual i at time t, and β_K, $K = \lambda, r_D, r_S, d$, are vectors of the parameters. We assume that the random effects ε_i^K are normally distributed with mean zero and variance matrix Ω_ε, given by

$$\Omega_\varepsilon = \begin{bmatrix} \sigma_\lambda^2 & cov(\lambda, r_D) & cov(\lambda, r_S) & cov(\lambda, d) \\ cov(\lambda, r_D) & \sigma_{r_D}^2 & cov(r_D, r_S) & cov(r_D, d) \\ cov(\lambda, r_S) & cov(r_D, r_S) & \sigma_{r_S}^2 & cov(r_S, d) \\ cov(\lambda, d) & cov(r_D, d) & cov(r_S, d) & \sigma_d^2 \end{bmatrix}$$

Then, the four coefficients are also normally distributed, we have $\begin{pmatrix} \lambda_{it} \\ r_{Dit} \\ r_{\mathrm{Si}\,t} \\ d_{it} \end{pmatrix} \sim N\left(\begin{pmatrix} x_{it}\beta_\lambda \\ x_{it}\beta_{r_D} \\ x_{it}\beta_{r_S} \\ x_{it}\beta_d \end{pmatrix}, \Omega_\varepsilon\right)$. Here e_{it}^J, $J = DG, SG, DL, SL$, is the error term of equation J of individual i at time t. We assume that they are all independent and normally distributed with mean zero and variance $\sigma_{e^J}^2$, $e_{it}^J \sim N(0,\, \sigma_{e^J}^2)$.

We will estimate two versions of this model in this paper. For a simple version, we assume that every coefficient of each individual i is equal to a constant plus the random effects term without the demographic variables. That is $\lambda_i = \lambda + \varepsilon_i^\lambda$, $r_{Di} = r_D + \varepsilon_i^{r_D}$, $r_{Si} = r_S + \varepsilon_i^{r_S}$, $d_i = d + \varepsilon_i^d$, $i = 1, \cdots, N$. The four constants λ, r_D, r_S and d do not change over i, only the individual specific effects ε_i^λ, $\varepsilon_i^{r_D}$, $\varepsilon_i^{r_S}$ and ε_i^d vary across different individuals i, but not over time t. We also estimate a more complicated version of the model using some demographic variables.

Because there are some outliers in our data, we use a Tobit model with right censoring at $\delta_J \leq 1.0$ (J = DG, SG, DL, SL) for each equation, that is the discount rate is smaller than or equal to one, and $y_{it}^J \leq 2.0$. The econometric model becomes:

$$\begin{aligned} y_{it}^{J*} &= z_{it}^J + e_{it}^J \\ y_{it}^J &= \min\{y_{it}^J,\, 2\} \\ i &= 1, \cdots, N,\ t = 1, ..., 6,\ J = DG, SG, DL, SL \\ & \quad \text{and where} \\ z_{it}^{DG} &= (1 + (\lambda_{it} - 1)r_{Dit})(1 + d_{it}) \\ z_{it}^{SG} &= \frac{1}{1 + (\lambda_{it} - 1)r_{\mathrm{Si}\,t}}(1 + d_{it}) \end{aligned}$$

$$\begin{aligned}
z_{it}^{DL} &= (1 - \frac{(\lambda_{it} - 1)r_{Dit}}{\lambda_{it}})(1 + d_{it}) \\
z_{it}^{SL} &= \frac{1}{1 - \frac{(\lambda_{it} - 1)r_{\mathrm{Si}\,t}}{\lambda_{it}}}(1 + d_{it}) \\
\lambda_{it} &= x_{it}\beta_\lambda + \varepsilon_i^\lambda \\
r_{Dit} &= x_{it}\beta_{r_D} + \varepsilon_i^{r_D} \\
r_{Sit} &= x_{it}\beta_{r_S} + \varepsilon_i^{r_S} \\
d_{it} &= x_{it}\beta_d + \varepsilon_i^d
\end{aligned}$$

Here, we observe y_{it}^J. The likelihood function is given by:

$$\begin{aligned}
L(\beta, \sigma) \;=\; & \int_{-\infty}^{+\infty} .. \int_{-\infty}^{+\infty} \prod_{i,\ t,\ J:\ y_{it}^J > 2} \Phi(\frac{z_i^J - 2}{\sigma_{e^J}}) \prod_{i,\ t,\ J:\ y_{it}^J \leq 2} \frac{1}{\sigma_{e^J}} \phi(\frac{y_{it}^J - z_{it}^J}{\sigma_{e^J}}) \\
& f(\varepsilon^\lambda, \varepsilon^{r_D}, \varepsilon^{r_S}, \varepsilon^d) d\varepsilon^\lambda d\varepsilon^{r_D} d\varepsilon^{r_S} d\varepsilon^d
\end{aligned}$$

where $f(\varepsilon^\lambda, \varepsilon^{r_D}, \varepsilon^{r_S}, \varepsilon^d)$ is the joint density function of random effects of four coefficients, so we need four dimensional integral to compute the likelihood. Together with the nonlinearity of the model, it is very time consuming to compute the likelihood directly with a numerical method. A feasible approach is to use a simulation technique, then we can compute the simulated likelihood of the model, and use maximum simulated likelihood to estimate all the parameters of this model. First, we draw four random variables for four coefficients λ, r_D, r_S and d independently. In order to increase the efficiency, we use the Halton sequences, which can be considered as well-placed draws from a standard uniform distribution. Then we should transform it into standard normal distribution η, $\eta \sim N(0, I)$, where I is the identity matrix. See Train (2003) and Hajivassiliou and Ruud (1994) for a detailed discussion of Halton draws and maximum simulated likelihood. We can calculate a Choleski factor C of Ω_ε, which is a lower-triangular matrix such that $CC' = \Omega_\varepsilon$. Then we can obtain a draw of random effects $\varepsilon_i = (\varepsilon_i^\lambda, \varepsilon_i^{r_1}, \varepsilon_i^{r_2}, \varepsilon_i^d)'$, since for $\varepsilon_i = C\eta$, we have $\varepsilon_i \sim N(0, \Omega_\varepsilon)$. Secondly, after getting a draw of the random effects ε_i of individual i, we can compute the four coefficients λ_i, r_{i1}, r_{i2} and d_i with all other parameters, and then repeat this for all four equations J, time periods t and all individuals i, we can get the value of the total likelihood $L(\beta, \sigma)$. After that, we repeat the first and second steps for many times (M) and compute the average of the likelihood, $\hat{L} = \frac{1}{M} \sum_{m=1}^{M} L_m$ and then employ the BHHH[12] algorithm to maximize the simulated likeli-

[12] Berndt, Hall, Hall and Hausman (1974) proposed this procedure of the numerical search for the maximum of the log-likelihood.

hood. We use 200 Halton draws when we simulate the random effects terms and estimate the model.

3.6 Results

In this section, we present the results of our model. In order to check the procedure of estimation we presented in previous section, first we estimate a simple model with only three coefficients: loss aversion λ, the same reference point r for delay and speedup, the same discount rate d for gain and loss, and without correlation between the random effects terms ε_i^λ, ε_i^r and ε_i^d. We will allow correlations between the random effects terms in a more complex model later on.

Table 4.3 shows the result of this model. We get a reasonable coefficient of loss aversion of 1.34, smaller than previous studies. The reference point is around 0.2, meaning that on average people put 20% of the money involved as the reference point. It is significantly different from zero and from one, implying that it might be too strong an assumption if we put the reference point equal to one, like previous studies did. The result $\lambda > 1$ and $r > 0$ also give us some evidence that it might be a good way to model the intertemporal choice by using a model with the reference point and loss aversion. The estimated discount rate of 9.6% is also a quite reasonable result. All the parameters estimated are highly significant at any level, it suggests that the result is robust.

Table 3.5: Results of the model: with only three coefficients

Variable	Parameter	t-statistic
Loss aversion λ	1.34	70.0
σ_λ	0.070	9.64
Reference point $r : r_D = r_S$	0.198	19.1
σ_r	0.065	18.1
Discount rate d	0.096	74.6
σ_d	0.039	66.1
$\sigma_{e^{DG}}$	0.254	240.6
$\sigma_{e^{SG}}$	0.045	306.0
$\sigma_{e^{DL}}$	0.063	331.7
$\sigma_{e^{SL}}$	0.167	238.6
Loglikelihood	32178.3	

Note: without correlation between random effects terms

The model we are more interested in is the model with the four coefficients: loss aversion λ, different reference points for delay and speedup, r_D and r_S respectively, and the same discount rate for gains and losses d. Table 4.4 presents the estimation results of the simple version of this model. All the parameters we estimated for this model are highly significant at any level. The coefficient of loss aversion we got is exactly 2.0 on average, consistent with previous empirical studies. The values of the reference point for delay and speedup are 0.13 and 0.07 on average respectively. The reference point of delay is 1.8 times as big as that of speedup, consistent with Loewenstein's (1988) hypothesis. The discount rate is around 10%. The standard deviation of each random coefficient is around half of the average value of the coefficient, this makes sure that in most cases we have $\lambda > 0$, $0 < r_D < 1$, $0 < r_S < 1$, and $d > 0$. We will look at the distributions of the four coefficients across all individuals later on in this section.

Table 3.6: Results of the simple version of the model with four coefficients

Variable	Parameter	t-statistic
Loss aversion λ	2.00	131.5
σ_λ	0.646	113.4
Reference points of delay r_D	0.129	55.8
σ_{r_D}	0.076	74.3
Reference points of speedup r_S	0.071	45.8
σ_{r_S}	0.033	54.8
Discount rate d	0.104	97.0
σ_d	0.068	98.2
$\sigma_{e^{DG}}$	0.173	243.9
$\sigma_{e^{SG}}$	0.042	289.3
$\sigma_{e^{DL}}$	0.054	348.7
$\sigma_{e^{SL}}$	0.164	323.2
Loglikelihood	36023.6	

Note: different reference points for delay and speedup.
with correlation between random effects terms.

The correlation coefficients between the random effects of coefficients λ, r_D, r_S, and d for this model are significant at any level. This result is presented in table 4.5. The random effect of the discount rate is positively correlated with those of other three coefficients, especially highly correlated with loss aversion. The correlation coefficient of 0.96 indicates that loss averse people are impatient. The correlation coefficient between the random effects of the two reference points, delay r_D and speedup r_S, is also quite high, 0.55, consistent with our expectation. It means that people with a high reference point of delay are also more likely to have a high reference point for speedup.

Table 3.7: Correlation Coefficients of random coefficients

	λ	r_1	r_2	d
r_1	0.11 (11.4)	1.0		
r_2	0.19 (12.4)	0.55 (45.8)	1.0	
d	0.96 (380.7)	0.36 (26.7)	0.29 (15.6)	1.0

Note: t-statistic in parentheses. This is the result from the simple version of the model with four coefficients.

In table 4.6, the result of a more complicated version of the model with four coefficients is presented; the four coefficients are linear functions of some demographic variables and a random component independent of these demographic variables. It seems that females are slightly more loss aversion than males, with a 0.03 higher coefficient of loss aversion on average; higher education level, especially the middle education level (school2) makes people less loss averse; and age also makes people less loss averse. For example, a 20-year-old person on average has a loss aversion coefficient at 2.22, 0.36 higher than a 60-year-old person, on average. Females have higher reference points r_D and r_S than males; people with the highest education level (school3) have lower r_D and r_S; school2 has different effects on reference points than school3, it has no significant effect on r_D, but it makes r_S bigger. r_D and r_S also have different age-patterns, r_D has a U-shaped age-pattern, but r_S has an inverted U-shape.

For the discount rate d, we have the result that education makes people more patient, the effects are not very strong but significant. Higher education makes people slightly more patient, on average, the discount rate is 0.3-0.4 percentage point lower than those low educated. This is in line with the theoretical prediction of Becker and Mulligan (1997) that higher educated people are more patient. With the same data set, but different waves, Tu et al. (2004) got the different result that education had no significant effects on discount rates, Tu et al. (2004) did not use a structural model with loss aversion and reference points like this paper, so that the effects on discount rates might be mixed with effects on loss aversion and reference points. It seems that there is no significant difference between discount rate of females and males, so that females and males have the same patience level on average. Age has a quite big positive effect on patience. For example, on average, a person of 20-years-old has a discount rate of 0.129, 0.05 higher than a 60-year-old.

The estimated correlation coefficients of the random effects are almost the same as in the simple version model in table 4.5. We do not present them here again. The log likelihood of this model is much bigger than that

Table 3.8: Result of the model with four coefficients

Variable	λ		r_D		r_S		d	
	Para.	t-stat.	Para.	t-stat.	Para.	t-stat.	Para.	t-stat.
Age/10	0.062	6.4	-0.012	-5.1	0.005	3.3	-0.010	-4.4
Age-sq./1000	-0.189	-14.1	0.016	4.0	-0.008	-4.2	-0.003	-1.0
Female	0.031	5.9	0.006	3.2	0.004	3.6	0.001	0.6
School2	-0.082	-13.9	-0.001	-0.53	0.003	3.0	-0.003	-2.3
School3	-0.067	-4.7	-0.009	-3.7	-0.003	-2.6	-0.004	-2.2
Constant	2.211	107.1	0.160	24.6	0.056	16.3	0.152	32.5
σ of REs	0.716	109.7	0.077	72.9	0.037	56.6	0.071	106.2
	$\sigma_{e^{DG}}$		$\sigma_{e^{SG}}$		$\sigma_{e^{DL}}$		$\sigma_{e^{SL}}$	
	Para.	t-stat.	Para.	t-stat.	Para.	t-stat.	Para.	t-stat.
	0.171	242.3	0.040	318.0	0.052	359.0	0.165	317.1
Loglikelihood				36599.3				

Note: four coefficients are two reference points, for delay and speedup respectively, the coefficient of loss aversion, and the discount rate.

of the simple version in table 4.4 without demographic variables, indicating that the model with these demographic variables fits our data better.

We use the model and parameters in table 4.6 to compute the estimated values of the four coefficients we are interested in, λ, r_D, r_S, and d; the random effects are simulated once according to the variances and correlation coefficients of the random terms. In total we have 11,847 observations. Table 3.9 gives the descriptive statistics of the simulated results. Figure 3.2 shows the distributions. The mean values of the four coefficients are quite similar to the results we get from the simple version of the model; the mean coefficient of loss aversion λ is 2.00, also consistent with previous empirical studies. The reference point of delay is almost twice as big as of that of speedup on average, 0.120 and 0.063 respectively. The last column of Table 3.9 is the percentages of estimated coefficients which are smaller than zero. For most cases, we have $\lambda > 0$, $0 < r_i < 1$, $i = D, S$, and $d > 0$. There are 91.5% of the simulated parameter values have $\lambda > 1$, showing people are loss averse. 83% estimated observations satisfy the conditions $\lambda > 1$, $0 < r_i < 1$, $i = D, S$, and $d > 0$ jointly.

3.7 Conclusions

Our model shows that the discount rates individuals implicitly used to make intertemporal trade offs depends enormously on the level of loss aversion

Table 3.9: Descriptive statistics for the simulated individual coefficients

Variable	Mean	Median	Std. Dev.	Min	Max	%<0
Loss aversion λ	1.99	2.00	0.741	-0.343	4.59	0.5%
Ref. point of delay r_D	0.119	0.120	0.077	-0.153	0.379	5.7%
Ref. point of speedup r_S	0.063	0.063	0.037	-0.074	0.193	4.3%
Discount rate d	0.096	0.097	0.073	-0.143	0.338	9.0%

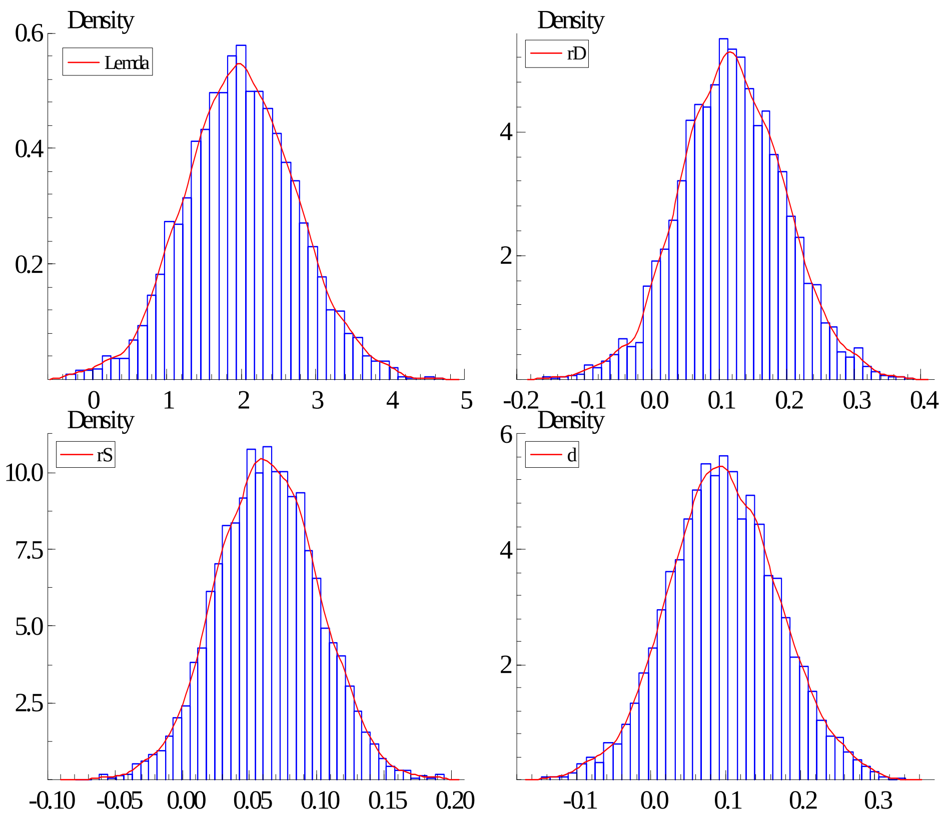


Figure 3.2: Distributions of estimated coefficients

and the reference point being used. Reference points, loss aversion, and of course the discount rate itself have a substantial impact on households' responses to all kinds of policy measures that relate to intertemporal choices. Understanding the joint distributions of these parameters in the population and their effects on time preference in different scenarios will help us to predict the behaviors of individuals and households much better.

In this paper, we present a structural model with reference points and loss aversion for intertemporal choice for four scenarios: delay of gains, delay of losses, speed-up of gains, and speed-up of losses. In order to estimate this structural model with four equations, we use a simulation technique to compute the likelihood. With help of the rich data set representative for the Dutch population, we are the first to get the population distribution of reference points and loss aversion. We get clear evidence to support the hypothesis of Loewenstein (1988) that the reference point of delay is larger than that of speedup. We have the result that the mean coefficient of loss aversion is around two, which is in line with previous findings in literature. Our results also show that females are more loss averse than males, and high-education and age make people less loss averse; high educated or older people are also more patient. We find significant (all at the 1%) correlations between the random effects of the four coefficients. There is a positive correlation between random effects of loss aversion and the discount rate, so impatient persons are more loss averse. These results, distribution of coefficients and observed relationships of those parameters, may help us predicting and understanding the behavior of households for policy purposes.

Reference points might be more fragile than loss aversion and discount rate, as the characteristics of an individual, the coefficient of loss aversion and the discount rate might not like reference points, change so often. The data we use here came from four scenarios with fixed time period, one year, and fixed amount of money, Dfl. 1000. A further research project can focus on testing whether the coefficient of loss aversion and the discount rate keep as stable when we only change the amount of money or the time horizon concerned. Models with interaction between loss aversion and risk aversion will also be interesting topics for further study. Future research can also focus on the role of loss aversion and reference points in particular policy applications.

Chapter 4

Joint Estimation of Loss Aversion and Probability Weighting Function

4.1 Introduction

In traditional economic theory, risk aversion is a crucial determinant of individual decision-making under uncertainty. Nowadays, risk aversion is widely used in all economic theories, especially in financial decision-making, like portfolio choice. Risk aversion, intuitively, implies that when facing choices with comparable expected returns, individual tends to choose the less risky alternative. Traditional theory for decision-making under uncertainty is expected utility (EU) theory, in which risk aversion is equivalent to concavity of the utility function. This means that the utility function exhibits diminishing marginal utility of wealth. The probabilities used to compute EU are the actual probabilities perceived by the decision-maker, the risk attitude is solely captured by the shape of utility function.

The EU model cannot explain Allais' paradox (Allais 1953, Kahneman and Tversky 1979), suggesting that non-linear probability weights are needed to transform both very small and very large probabilities. Matthew and Thaler (2001) criticized the EU model, described it as a "dead parrot", even though it has "beautiful plumage" - a simple and elegant structure - but it is dead, and economists should develop better descriptive models for choice under uncertainty.

The EU model cannot explain the common empirical finding that the same individual exhibits risk-aversion and risk-seeking behavior at the same time. For example, we often observe a person who is risk-averse in most cases and buys lottery tickets at the same time; it means that she prefers risk aversion for some prospects and at the same time prefers risk-seeking for other prospects. This phenomenon cannot be explained by EU with a concave (or convex) utility function only, but theories with probability weighting function could be good explanations. In all these theories, cumulative prospect theory (CPT) introduced by Tversky and Kahneman (1992) is the dominant one. Compared to the EU model, not only the shape (concavity or convex) of the utility function affects the risk attitude of individual, but the probability weighting function also has an impact on the risk attitude. As Tversky and Kahneman show in their paper, people overweight low probabilities, and underweight moderate and high probabilities both for gains and losses, and this can make people risk-seeking in small probabilities for gains and in high probabilities for losses, and risk averse in high probabilities for gains and small probabilities for losses. By introducing probability weighting, it is thus possible to explain why risk averse people would also buy lottery tickets. Therefore, an individual's risk attitude is determined in two domains in CPT instead of the one domain in the EU model: one is the domain of value - concavity of the utility function; the other is domain of probability -

the nonlinearity of the probability weighting function. Tversky and Kahneman's experimental results indeed confirmed that people overweighted low probabilities, and underweighted moderate and high probabilities both for gains and losses, making people risk-seeking in small probabilities for gains and in high probabilities for losses, and risk averse in high probabilities for gains and in small probabilities for losses.

Because the risk attitude is determined by the shapes of both the utility function and the probability weighting functions, and since these two functions are closely linked when people make decisions under uncertainty, it is important to estimate these two functions together if we want to understand the risk attitude and the behavior of individuals under uncertainty. Because of lack of adequate data, no existing study jointly estimates these two functions with a population representative data set. In this paper, we will use a survey which is a representative of the Dutch population to estimate the shapes of the value and probability weighting functions for gains and losses together with the coefficient of loss aversion, and investigate how these coefficients vary with observed demographics and unobserved characteristics of the individuals.

Nowadays, risk aversion and loss aversion are commonly used and have become two fundamental concepts in behavioral economics. From a psychological view, loss aversion is more fundamental; people do not like a risk not because of the risk itself, but because of the potential loss, which is the possible result of risk. A recent paper Duxbury and Summers (2004) investigated whether individuals' perceptions of risk are linked to variance aversion[1] or loss aversion by an experiment, and found that a link to loss aversion is supported. After the seminal paper of Kahneman and Tversky (1979), the importance of loss aversion in decision under risk has become common knowledge, with more and more studies attempting to combine risk aversion and loss aversion to analyze decisions under uncertainty. Preferences incorporating loss aversion can reconcile significant small-scale risk aversion with reasonable degrees of large-scale risk aversion (Rabin 2000). Kahneman and Tversky (1979) interpreted loss aversion using a symmetric 50-50 bet. Formally, loss aversion holds if $(0.5, x; 0.5, -x) \prec (0.5, y; 0.5, -y)$ for all $x > y \geq 0$.[2] Actually this preference can be also explained by risk aversion, because the variance of the bet $(0.5, x; 0.5, -x)$ is larger than that of $(0.5, y; 0.5, -y)$ if $x > y \geq 0$, so we can say that the bet $(0.5, x; 0.5, -x)$ is more risky. We could use this kind of simple bets to estimate the coefficient

[1] Variance aversion here refers to the traditional concept of risk aversion. Normally we call outcomes with higher variance more risky.

[2] $(p, x, 1-p, y)$ refers to a bet that pays x with probability p and y with probability $1-p$.

of loss aversion.

In the literature, many experimental studies estimated probability weighting functions from individual choices, see Tversky and Kahneman (1992), Tversky and Fox (1995), Wu and Gonzalez (1996), Gonzalez and Wu (1999), Abdellaoui (2000), Bleichrodt and Pinto (2000), Kilka and Weber (2001), Brandstatter, Kuehberger, and Schneider (2002). Their most remarkable finding is that the weighting functions are inverse S-shaped both for gains and losses, which means that people overweight low probabilities and underweight moderate and high probabilities both for gains and losses. This result implies that people are risk-seeking in small probabilities for gains and in high probabilities for losses, and risk-averse in high probabilities for gains and in small probabilities for losses.

There are also many papers in the literature estimating individual risk attitude under uncertainty. Based on CPT, Donkers, Melenberg, and Van Soest (2001) used questions about lotteries in a large household survey to estimate an index for risk aversion. Their result showed that both the value function and the probability weighting function vary significantly with age, income, and wealth of the individual. Neilson and Stowe (2002) examined many experimental studies of probability weighting functions based on CPT; they suggested that the functional forms proposed in the literature were not suitable for generalization to applied settings. CPT, and especially the weighting function's parameterization, still needs more research.

In this paper, based on CPT and the structure of Tversky and Kahneman's (1992) model, we use data from a large representative survey to estimate the shapes of value functions and weighting functions jointly with loss aversion. To account for heterogeneity in the population, we allow all the coefficients we are interested in to vary across individuals using a random coefficients model.

The remainder of this paper is organized as follows. In Section 2, we introduce an economic model we will use in this paper. Then in Section 3, we describe the data. We discuss the econometric model and the estimation procedure in Section 4. In Section 5, the results are presented, and Section 6 concludes.

4.2 Economic Model

We follow CPT of Tversky and Kahneman (1992) to construct an economic model. The utility of a prospect $p = (x_1, p_1; x_2, p_2; \cdots, x_n, p_n)$, where p_i is the probability of outcome x_i, $i = 1, \ldots, n$, and $\sum_{i=1}^{n} p_i = 1$, is

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

in which payoffs x_i are increasing order, $x_1 \leq \cdots \leq x_m \leq 0 \leq x_{m+1} \leq \cdots \leq x_n$, π_i^+ and π_i^- are the decision weighting functions for gains and losses respectively, where for gains, $\pi_n^+ = w^+(p_n)$ and $\pi_i^+ = w^+(\sum_{j=i}^{n} p_j) - w^+(\sum_{j=i+1}^{n} p_j)$ for $m+1 \leq i \leq n-1$, and for losses, $\pi_1^- = w^-(p_1)$ and $\pi_i^- = w^-(\sum_{j=1}^{i} p_j) - w^-(\sum_{j=1}^{i-1} p_j)$ for $2 \leq i \leq m$. For CPT, these probability weights do not necessarily sum to one.

Again following Tversky and Kahneman (1992), we use a power value function, and with loss aversion, the value function is defined differently in domain of gains and losses.

$$v(x) = \left\{ \begin{array}{ll} x^a & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\lambda(-x)^b & \text{if } x < 0 \end{array} \right\}$$

Here λ is the coefficient of loss aversion. Tversky and Kahneman expect that $0 < a \leq b \leq 1$ and $\lambda \geqslant 1$, implying that the value function is steeper for losses than for gains: "losses loom larger than corresponding gains." In this case, the value function is S-shaped: concave for gains and convex for losses. Without probability weighting, the shape of the value function will capture the risk attitude of individual. With an S-shaped value function, the individual is then risk-averse for gains and risk-seeking for losses. On the other hand, if $a > 1$, the individual is risk-seeking for gains and if $b > 1$, the individual is risk-averse for losses. Thus with a power function specification, the values of the parameters a and b determine completely the shape of value function and the individual's risk attitude in the value domain.

We use the Arrow-Pratt measure of relative risk-aversion to represent the degree of risk aversion of an individual in the value domain, defined as $R_v(x) = -xv''/v'$. For a power value function x^a, $R_v = 1 - a$ for gains, that is we have constant relative risk-aversion (CRRA). If $a < 1$, $R_v = 1 - a > 0$, the individual is risk averse for gains; if $a > 1$, $R_v = 1 - a < 0$, the individual is risk-seeking for gains. It is opposite for losses, if $b < 1$, the individual is risk-seeking for losses; if $b > 1$, the individual is risk averse for losses.

The probability weighting functions also have impact on individual risk attitude. It is believed that people overweight low probabilities and underweight moderate and high probabilities both for gains and losses, so that people are risk-seeking in small probabilities for gains and in high probabilities for losses, and risk-averse in high probabilities for gains and small probabilities for losses.

Tversky and Kahneman (1992) proposed to use a two-part power function as the weighting function for probabilities. They use the following functional form for the weighting functions:

$$p_+ = w^+(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}$$

for gains, and

$$p_- = w^-(p) = \frac{p^\delta}{(p^\delta + (1-p)^\delta)^{\frac{1}{\delta}}}$$

for losses. γ and δ are parameters of weighting functions for gains and losses respectively, and they should be positive numbers between zero and one. Tversky and Kahneman used the median value of subjects in their experiment to estimate these parameters. The parameter values they obtained are $a = b = 0.88$, $\lambda = 2.25$, $\gamma = 0.61$ and $\delta = 0.69$.

In our paper, we use a different weighting function. The weighting function for gains is:

$$p_+ = \exp(-(-\ln p)^m)$$

and the weighting function for losses is given by:

$$p_- = \exp(-(-\ln p)^n)$$

We have three reasons to choose this functional form. First, it has better properties than the two-part power function, Prelec (1998) shows that this weighting function satisfies all four target properties he mentioned: regressive, asymmetric, s-shaped, reflective, and has an invariant fixed point at $p = 1/e = 0.37$. Second, it fits our data better, the log likelihood is much bigger than for the two-part power function. The last reason is most important. We use a random coefficients model, so the parameters of the weighting functions will vary over individuals; the two-part power weighting function is not a monotonically increasing function when γ and δ is small, for example, if $\gamma = 0.2$, $w^+(0.5) = 0.0544 < w^+(0.1) = 0.0583$, this is quite difficult to understand and will cause some problem when we interpret our model.

With the setting above, the shape of the value functions, a and b, together with the shapes of the weighting functions, m and n, will determine the risk attitude of an individual jointly. The value function will affect the individual risk attitude in the value domain, as in traditional EU theory, but in addition, the probability weighting function will determine the risk attitude in the probability domain.

In this paper, we use data of questions about lotteries to estimate the model introduced above. For a lottery $L(-x_1, p_1; x_2, 1-p_1)$, with probability p_1 to lose x_1 and probability $1-p_1$ to gain x_2, the utility is $V_L = p_- v(-x_1) + p_+ v(x_2) = -p_- \lambda x_1^b + p_+ x_2^a$, where λ is the coefficient of loss aversion, a and b are parameters of power value function, and $p_- = \exp(-(-\ln p_1)^n)$ and $p_+ = \exp(-(-\ln(1-p_1)^m)$ are the weights of probabilities for losses and gains respectively.

4.3 Data

The data we use is taken from the DNB Household Survey (DNBHS) collected by CentERdata in August 2004. This is a large representative Dutch survey with panel data for more than 1500 households, starting in 1993. The DNBHS is a rich data set containing information on employment status, pensions, accommodation, mortgages, income, assets, debts, health, economic and psychological concepts, and personal characteristics. It is administered over the Internet and people without access to Internet or without a personal computer have been provided with the necessary equipment to participate. Researchers can use the system to conduct their own module of survey questions. We constructed a questionnaire with seven questions on simple bets[3]

Questions 1 to 4 have one bet and questions 5 to 7 have two bets. Question 1 presents several bets with a 50% chance of losing 100 euros and a 50% chance of winning an amount varying from 100 to 900 euros, depending on randomization of the question. In each case, the question is asked whether the respondent would be willing to take part in such a bet. Question 2 asks for the minimum prize the individual should be able to win with 50% chance if the lottery also has 50% chance of losing either 500 or 1000 euros (randomized). Question 3 asks for the minimum prize if the chances of losing 100 euros are 20%, 30% or 40% (randomized) and the chances of winning are, accordingly, 80%, 70% or 60%. Question 4 is similar but reverses the probabilities of winning and losing. Question 5 asks the respondent to choose between two lotteries, one with a 30% chance of winning 100 euros and a 70% chance of neither winning nor losing, and the other with a 50% chance of winning a larger (randomized) amount but also with a 50% chance of losing 100 euros. Questions 6 and 7 ask about the maximum loss and minimum prize required for specific choices in a similar trade offs between two lotteries.

[3] See appendix for the exact questions.

There are in total 2,062 individuals who take part in the survey, and we deleted 230 individuals who did not answer any of the questions in our module. Normally, data of individual choices are relatively noisy, which might be the main reason why most of the existing experimental studies only used some kind of aggregate data in their analysis, like median or mean. Not every respondent answered all seven questions. The response rates are between 70%, the lowest for question 4, and 97%, the highest for question 1. The average answer rate is 78.6% for all seven questions.

To reduce the amount of noise in our data, we decided to check consistency of the answers and retained only the observations who gave consistent answers.[4] For a lottery $L(-x_1, p_1; x_2, 1-p_1)$, the utility is $V_L = w^-(p_1)v(-x_1) + w^+(1-p_1)v(x_2)$, and the value and weighting functions should be monotonically increasing functions. According to this basic rule, we can construct ten checks for the consistency of answers for six of our seven questions.[5]

1. The lower bound of question 1 should be smaller than the answer of question 2.
2. The higher bound of question 1 should be larger than the answer of question 3.
3. The lower bound of question 1 should be smaller than the answer of question 4.
4. The lower bound of question 1 should be smaller than the higher bound of question 5.
5. The lower bound of question 1 should be smaller than the answer of question 7.
6. The answer of question 2 should be larger than that of question 3.
7. The answer of question 3 should be smaller than that of question 4.
8. The answer of question 3 should be smaller than the higher bound of question 5.
9. The answer of question 3 should be smaller than that of question 7.
10. The the lower bound of question 5 should be smaller than that of question 7 when the randomization of question 7 is equal to 2.

We deleted all the answers with a consistency problem, (i.e., if two answers were inconsistent, we deleted both of them) and also deleted around twenty outlier answers for each question, leaving 89 observations without any answer. This gives a final data set of 1,743 observations for our estimation.

---

[4] If we use all the observations of our data without checking for consistency, the algorithm for simulated maximum likelihood does not converge.

[5] Question 6 is not involved in the consistency checks, because it needs an additional assumption other than monotonicity of the value and weighting functions if we want to compare its answer with other questions.

This kind of inconsistency might be caused by two reasons. It seems that low educated people are more likely to answer questions in an inconsistent way, the reason might be that the questions are more difficult for them to understand. The other reason is that people do not answer the questions carefully and randomly pick a number. In order to make full use of our data, we do not delete those individuals who only answered part of total seven questions. Table 4.1 shows the structure of our data. After cleaning all the data, 3.7 questions are answered consistently by each individual on average.

Table 4.1: Structure of the panel

Number of observations for each question			Number of consistent answers	
Question	All answers	Consistent	Number of ques.	Individuals
Question 1	1,768	872	1	315
Question 2	1,338	1,110	2	323
Question 3	1,398	625	3	293
Question 4	1,282	907	4	266
Question 5	1,566	835	5	128
Question 6	1,354	1,321	6	79
Question 7	1,371	721	7	339
Total answers	10,077	6,391	Total	1,743

Exploiting the advantages of an internet survey, in order to increase the design variance and to help to identify our model, we used a randomization technique for the survey design. From Figure 4.1 (in **Results** section), we can see the distribution of the actual answers of question 2 when randomization=1 (50% chance of losing 500), and it seems that the distribution of the data is more like the log-normal distribution than like a normal distribution. This will affect the econometric model we want to estimate, which we will discuss in the next section. Some descriptive statistics on these questions and demographic variables are provided in Table 4.2. We take the logarithm of the answers for all questions.

For question 2, if the randomization variable ("randomization") is equal to 1, the bet is $L(-500, 0.5; x_2, 0.5)$, and the bet is $L(-1000, 0.5; x_2, 0.5)$ for randomization=2. It is easy to understand that the average (or median) answer of the former should be smaller than that of the latter, and this is exactly the result we get from the data. We can have similar expectations for other questions with different randomization, and Table 4.2 shows that all our expectations can be confirmed by our data.

Table 4.2: Descriptive statistics for the answers of questions and demographic variables

Variable	Obs.	Median	Mean	Std. Dev.	Min	Max
Question 2, random.=1	557	7.60	8.04	1.97	1.61	13.8
Question 2, random.=2	553	8.52	8.68	2.13	2.30	14.5
Question 3, random.=1	241	5.01	5.08	1.45	0.0	10.8
Question 3, random.=2	206	5.30	5.22	1.61	0.0	10.8
Question 3, random.=3	178	5.41	5.39	1.97	0.0	11.5
Question 4, random.=1	288	7.82	8.57	2.44	2.30	14.5
Question 4, random.=2	293	6.91	7.86	2.51	1.61	14.5
Question 4, random.=3	326	6.91	7.69	2.25	2.30	14.2
Question 6, random.=1	643	4.61	4.25	1.23	0.693	8.82
Question 6, random.=2	678	4.61	4.51	1.28	1.39	9.90
Question 7, random.=1	336	6.21	6.59	1.44	4.50	12.2
Question 7, random.=2	385	6.91	7.30	1.46	5.30	12.4
Female	1743	0	0.447	0.497	0	1
High-education	1743	0	0.439	0.496	0	1
Age/10	1743	4.7	4.80	1.56	1.6	9.0
Age-squared/1000	1743	2.21	2.55	1.54	0.256	8.1

4.4 Econometric Model

In this section, we develop an econometric model and use our survey data to estimate this. In order to introduce heterogeneity, we use a random coefficients model. The first four questions, questions 1 to 4, concern one lottery, and questions 5 to 7 involve two lotteries. For questions with only one lottery, we assume that the utility of the lottery $L(-x_1, p_1; x_2, 1-p_1)$ should be equal to zero if the subject is indifferent between accepting and not accepting the lottery. Then the utility V_{ij} of question i, $i = 1, \cdots, 4$, for individual j, $j = 1, \cdots, N$, can be written as:

$$V_{ij} = p_{-1ij}v(-x_{1ij}) + p_{+1ij}v(x_{2ij}) = -p_{-1ij}\lambda_j x_{1ij}^{b_j} + p_{+1ij}x_{2i}^{a_j} = 0$$

For these four questions with one lottery, we always know $-x_1$, p_1, and ask for the relevant x_2. If we use the weighting function we presented in section two, then $p_{-1ij} = \exp(-(-\ln p_{1ij})^{n_j})$ and $p_{+1ij} = \exp(-(-\ln(1-p_{1ij}))^{m_j})$. We can solve the above equation and get:

$$x_{2ij} = (\frac{p_{-ij}}{p_{+ij}}\lambda_j x_{1ij}^{b_j})^{\frac{1}{a_j}},\ i = 1, \cdots, 4,\ j = 1, \cdots, N$$

The five coefficients we are interested in are the powers of value functions for gains and losses a_j, and b_j, the coefficient of loss aversion λ_j, and the coefficients of the weight functions for gains and losses m_j and n_j. These coefficients vary across individuals j, but have nothing to do with a specific question i. Because the distribution of data is more like log-normal distribution (see Figure 3.1) and it seems that a log-normal distribution fit our data better than a normal distribution, we use a logarithm form:

$$\begin{aligned} \ln(y_{ij}) &= \ln(x_{2ij}) + e_{ij} \\ &= \frac{1}{a_j}\ln(\frac{p_{-ij}}{p_{+ij}}\lambda_j x_{1ij}^{b_j}) + e_{ij} \\ i &= 1,\cdots,4,\ j = 1,\cdots,N \end{aligned}$$

Here y_{ij} is the observed answer to the question on x_{2ij}, the answers of individual j for questions i, $i = 1, \cdots, 4$. Compared to a question with only one lottery, a question with two lotteries is more complicated. We assume that the utilities of two lotteries $L_A(0, p_3; x_3, 1-p_3)$ and $L_B(-x_4, p_4; x_5, 1-p_4)$ should be equal, $V_A = V_B$, if the subject is indifferent between these two lotteries. Then we have:

$$\begin{aligned} 0 + p_{+3ij}v(x_{3ij}) &= p_{-4ij}v(-x_{4ij}) + p_{+4ij}v(x_{5ij}) \\ p_{+3ij}x_{3ij}^{a_j} &= -p_{-4ij}\lambda_j x_{4ij}^{b_j} + p_{+4ij}x_{5ij}^{a_j} \\ i &= 5,6,7,\ j = 1,\cdots,N \end{aligned}$$

The weighting functions are the same as in questions 1 to 4, so $p_{+3ij} = \exp(-(-\ln p_{3ij})^{m_j})$, $p_{+4ij} = \exp(-(-\ln(1-p_{4ij}))^{m_j})$, and $p_{-4ij} = \exp(-(-\ln(1-p_{4ij}))^{n_j})$.

For equations 5 and 7, only x_5 is unknown, and we have

$$x_{5ij} = (\frac{p_{+3ij}x_{3ij}^{a_j} + p_{-4ij}\lambda x_{4ij}^{b_j}}{p_{+4ij}})^{\frac{1}{a_j}},\ i = 5,7,\ j = 1,\cdots,N$$

For equation 6, x_4 is asked, and given by

$$x_{4ij} = (\frac{p_{+4ij}x_{5ij}^{a_j} - p_{+3ij}x_{3ij}^{a_j}}{p_{-4ij}\lambda_j})^{\frac{1}{b_j}},\ i = 6,\ j = 1,\cdots,N$$

Like for the one lottery questions, we think our data on x_{4ij} and x_{5ij} are log-normally distributed, and we still use a logarithm form. Then we have

$$\ln(y_{ij}) = \frac{1}{a_j}\ln(\frac{p_{+3ij}x_{3ij}^{a_j} + p_{-4ij}\lambda x_{4ij}^{b_j}}{p_{+4ij}}) + e_{ij},\ i = 5,7,\ j = 1,\cdots,N$$

and

$$\ln(y_{ij}) = \frac{1}{b_j}\ln(\frac{p_{+4ij}x_{5ij}^{a_j} - p_{+3ij}x_{3ij}^{a_j}}{p_{-4ij}\lambda_j}) + e_{ij},\ i = 6,\ j = 1,\cdots,N$$

Here y_{ij} is the observed answer of x_{5ij} for equations 5 and 7, and x_{4ij} for equation 6.

We assume that the error terms e_{ij} of all seven equations are iid with a normal distribution $N(0, \sigma_j^2)$, $j = 1, \cdots, 7$, implying that y_{ij} has a log-normal distribution. We employ a random coefficients model here, so all five coefficients we are interested in are random variables, specified as a linear function of demographics plus an additive random effect term. We can write them as $a_j = x_j' \beta_a + \varepsilon_{aj}$, $b_j = x_j' \beta_b + \varepsilon_{bj}$, $\lambda_j = x_j' \beta_\lambda + \varepsilon_{\lambda j}$, $m_j = x_j' \beta_m + \varepsilon_{mj}$, $n_j = x_j' \beta_n + \varepsilon_{nj}$. Here x_j is the vector of demographic variables of individual j, and the β_k, $k = a, b, \lambda, m, n$, are vectors of the parameters. We assume that the random effects terms ε_{kj} of the five coefficients are normally distributed with mean zero and covariance matrix Ω_ε given by

$$\Omega_\varepsilon = \begin{bmatrix} \sigma_a^2 & \rho_{21}\sigma_a\sigma_b & \rho_{31}\sigma_a\sigma_\lambda & \rho_{41}\sigma_a\sigma_m & \rho_{51}\sigma_a\sigma_n \\ \rho_{21}\sigma_a\sigma_b & \sigma_b^2 & \rho_{32}\sigma_b\sigma_\lambda & \rho_{42}\sigma_b\sigma_m & \rho_{52}\sigma_b\sigma_n \\ \rho_{31}\sigma_a\sigma_\lambda & \rho_{32}\sigma_b\sigma_\lambda & \sigma_\lambda^2 & \rho_{43}\sigma_\lambda\sigma_m & \rho_{53}\sigma_\lambda\sigma_n \\ \rho_{41}\sigma_a\sigma_m & \rho_{42}\sigma_b\sigma_m & \rho_{43}\sigma_\lambda\sigma_m & \sigma_m^2 & \rho_{54}\sigma_m\sigma_n \\ \rho_{51}\sigma_a\sigma_n & \rho_{52}\sigma_b\sigma_n & \rho_{53}\sigma_\lambda\sigma_n & \rho_{54}\sigma_m\sigma_n & \sigma_n^2 \end{bmatrix}.$$

We also assume that the five coefficients are jointly normally distributed,

$$\begin{pmatrix} a_j \\ b_j \\ \lambda_j \\ m_j \\ n_j \end{pmatrix} \sim N \begin{pmatrix} x_j'\beta_a & \\ x_j'\beta_b & \\ x_j'\beta_\lambda & ,\Omega_\varepsilon \\ x_j'\beta_m & \\ x_j'\beta_n & \end{pmatrix}.$$

For questions 1 and 5, the data is discrete. We use ordered probit models for these two questions. Since the ordered probit model is quite standard nowadays, we will not discuss it in detail here, see many econometric textbooks. Answers to the other five equations are continuous, with normally distributed error terms for which we can compute the density easily. The conditional likelihood conditional on random coefficients can now be straightforwardly written as:

$$L(\beta, \sigma | a_j, b_j, \lambda_j, m_j, n_j) = \prod_{j=1,N} \Big(\prod_{i=2,3,4,6,7} \frac{1}{\sigma_{e_i}} \phi\Big(\frac{\ln(y_{ij}) - \ln(x_{kij})}{\sigma_{e_i}}\Big) \prod_{i=1,5} P_{ij} \Big)$$

The unconditional likelihood function is then given by

$$\begin{aligned} L(\beta, \sigma) \;=\; & \int \cdots \int \prod_{j=1,N} \Big(\prod_{i=2,3,4,6,7} \frac{1}{\sigma_{e_i}} \phi\Big(\frac{\ln(y_{ij}) - \ln(x_{kij})}{\sigma_{e_i}}\Big) \prod_{i=1,5} P_{ij} \Big) \\ & dF(\varepsilon_{aj}, \varepsilon_{bj}, \varepsilon_{\lambda j}, \varepsilon_{mj}, \varepsilon_{nj}) \end{aligned}$$

Here $\phi(\cdot)$ is the density function of the standard normal distribution, k will change according to different questions, x_{2ij} for questions 2, 3, and 4, x_{4ij} for question 6, and x_{5ij} for question 7. P_{ij} is the probability calculated from ordered probit models, for questions 1 and 5. $F(\varepsilon_{aj}, \varepsilon_{bj}, \varepsilon_{\lambda j}, \varepsilon_{mj}, \varepsilon_{nj})$ is the joint distribution of the random parts of the random coefficients. Due to the nonlinear way in which the random effects enter the model, we cannot compute the likelihood function with four dimensional integral analytically. With the help of a simulation technique, we can approximate the exact likelihood with a simulated likelihood, and use maximum simulated likelihood to estimate all the parameters of this model. First, we draw five random variables for five coefficients a_j, b_j, λ_j, m_j and n_j of individual j independently from the standard normal distribution; η is the vector of these draws, $\eta \sim N(0, I)$, where I is the identity matrix. With the variance matrix Ω_ε, we can compute a Choleski factor C of Ω_ε, which is a lower-triangular matrix such that $CC' = \Omega_\varepsilon$. Then we can obtain a draw for the random effects $\varepsilon_j = (\varepsilon_{aj}, \varepsilon_{bj}, \varepsilon_{\lambda j}, \varepsilon_{mj}, \varepsilon_{nj})'$, since we know for $\hat{\varepsilon}_j = C\eta$ we have $\hat{\varepsilon}_j \sim N(0, \Omega_\varepsilon)$. Secondly, after getting a draw $\hat{\varepsilon}_j$ for the random effects of individual j, we can compute the five coefficients of individual j: $\hat{a}_j$, $\hat{b}_j$, $\hat{\lambda}_j$, $\hat{m}_j$, and $\hat{n}_j$ together with all other parameters, and then we can compute the simulated likelihood of individual j $\hat{L}_{j1}(\beta, \sigma)$ straightforwardly for the first draw. After that, we repeat the first and second steps K times (for a large value of K) and calculate the average of the likelihood, that is $\hat{L}_j = \frac{1}{K}\sum\limits_{k=1}^{K} \hat{L}_{jk}$, as the simulated likelihood of individual j; at last, we can get the simulated likelihood of all data $\hat{L}(\beta, \sigma)$, and employ the BHHH[6] algorithm to maximize $\hat{L}(\beta, \sigma)$. In this paper, we use 400 draws when we simulate the random effects terms and estimate the model. See Train (2003) for a detailed discussion of maximum simulated likelihood.

4.5 Results

In this section, we present the estimation results of the econometric model presented in the previous section. Table 4.3 shows some results of this model. There are five coefficients we are interested in: powers of value function for gains and losses, a and b, the coefficient of loss aversion, λ, parameters for weighting functions of gains and losses, m and n, are all modelled as linear functions of some background variables: dummy of female, dummy of high-educated, age, age-squared, and logarithm of individual total income, and an

[6] Berndt, Hall, Hall and Hausman (1974) proposed this procedure of the numerical search for the maximum of the log-likelihood.

unobserved random effect. In general, our results are well in line with those of Tversky and Kahneman (1992); the model with power value functions, weighting functions together with loss aversion can explain our data well. Because our data represents the Dutch population, it seems that the model could be used for modelling decision making under risk at the individual level.

Table 4.3: Estimation results of model

Variable	a		b		λ	
	Para.	t-st.	Para.	t-st.	Para.	t-st.
Female	-0.061	-2.33	-0.027	-0.68	0.065	0.38
High-edu.	-0.023	-0.91	-0.007	-0.18	0.075	0.42
Age/10	0.020	0.48	0.032	0.53	0.180	0.48
Age-sq./1000	-0.030	-0.74	-0.051	-0.84	-0.177	-0.48
Loginc	-0.002	-0.36	0.002	0.19	-0.066	-1.65
Constant	0.681	30.4	0.730	31.2	3.14	13.5
σ	0.151	25.3	0.000	0.00	1.14	11.9
Variable	m		n			
	Para.	t-st.	Para.	t-st.		
Female	0.083	2.04	-0.093	-0.79		
High-edu.	0.078	1.76	-0.359	-3.10		
Age/10	-0.026	-0.34	0.145	0.87		
Age-sq./1000	-0.007	-0.10	-0.122	-0.78		
Loginc	-0.001	-0.06	-0.037	-1.64		
Constant	0.999	30.3	0.593	8.41		
σ	0.447	19.7	0.000	0.00		
Log-like.			-10779.1			

Note: all regressors centered around their sample means.

Tversky and Kahneman (1992) expected that $0 < a \leq b \leq 1$, and their experimental result is $a = b = 0.88$ - the median of the individual estimates, supports their statement. Our result does fully support their expectation, the average of a is 0.68, less than that of b, 0.73, and they are all between 0 and 1. The result $a < b$ means that the marginal utility of gains is diminishing faster than that of losses; and the averages of a and b are both smaller than 0.88, meaning that on average people are more risk-averse for gains and more risk-seeking for losses in our data than in those of Tversky and Kahneman. The average relative risk-aversion coefficient $(1 - a)$ is 0.32 for gains. We also get the result that females are significantly more risk averse than males, because the estimated value of a for females is 0.061 smaller than that of males with the same other characteristics. Other variables have no significant effect on a, and all variables have an insignificant effect on b.

People with high income are more loss averse, but the income effect is only significant at the10% level. Other variables have no significant effects on λ. Our result imply that people are more loss averse on average than the sample of Tversky and Kahneman (1992). We find an average value of λ equal to 3.1, larger than their result, 2.25. Because we get the result that $a < b$, the marginal value functions of gains and losses are diminishing at different speeds. If we define the "real" loss aversion as $\tilde{\lambda} = -v(-x)/v(x) = \lambda x^{b-a}$, then $\tilde{\lambda}$ is not a constant anymore, because a is a little bit smaller than b, $b - a > 0$, so $\tilde{\lambda}$ will increase slowly when x increases. For example, if $x = 10$, $\tilde{\lambda} = 3.5$; if $x = 100$, $\tilde{\lambda} = 3.9$; if x=$1,000$, $\tilde{\lambda} = 4.4$; if x=$10,000$, $\tilde{\lambda} = 4.9$; if x=$1,000,000$, $\tilde{\lambda} = 6.2$. Compared to our pervious study,[7] the coefficient of loss aversion that we find here in a setting with risky choices is much bigger than in the riskless situation of the previous chapter.

For the specification of the probability weighting function we use, parameters m and n less than one mean that people are overweighting small probabilities and underweighting for moderate and high probabilities. The average parameters for the weighting functions are 1.0 and 0.59 for gains and losses respectively. The fact that the average value of m is very close to one means that the weights of probabilities of gains should are almost equal to the objective probabilities, on average. We only use probabilities between 0.1 and 0.9 in the design used to estimate m in our data, and we do not use very small probabilities like 0.001 or very high probabilities like 0.99, so it might be that people just use objective probabilities for most cases, except for very small or a very big probabilities, like in real life lotteries, where people may then still overweight the very small probability of winning. The standard deviation of m is substantial, and approximately half the people have m larger than one, and half less than one. Those with m smaller than one are risk-seeking in small probabilities and risk-averse in high probabilities; on the contrary, the 50% with m larger than one are risk-averse in small probabilities and risk-seeking in high probabilities. Only those people with m small enough, making them heavily overweight small probabilities, will buy lotteries, so it seems more than half people will not buy any lotteries in our sample.

The average of n is much smaller than one, meaning that people are really overweighting low probabilities and underweighting moderate and high probabilities for losses. This difference between gains and losses is different from Tversky and Kahneman (1992), who find that people overweight low

[7] See the previous chapter in this thesis.

probabilities and underweight moderate and high probabilities both for gains and losses.[8]

Our result shows that females have a higher m than otherwise similar males. Also, high educated people have a higher m, though the education effect is only significant at 10% level. With high m, people are less likely to buy lotteries, so our results also indicate that females and high educated people are less likely to buy lotteries. In contrast, high education has a different effect on n: high educated people have a significantly lower n, and people with high income also have lower n, significant at 10% level.

Of the standard deviations of the random effects terms of the five coefficients we estimated, two are close to zero: those for b, the power of the value function for losses, and n, the parameter of the weighting function of losses. The other three standard deviations are significantly different from zero. They are all smaller than half of the means of corresponding coefficients, implying that for the large majority of the observations, we get a positive a, λ and m (and b and n). Why only the two random effects that are connected with losses are close to zero is not clear. Standard deviations of error terms of all questions are presented in Table 4.4. They are highly significant, indicating noise in the answers.

Table 4.4: Standard deviations of error terms for all questions

Equation 1		Equation 2		Equation 3		Equation 4	
S.D.	t-stat.	S.D.	t-stat.	S.D.	t-stat.	S.D.	t-stat.
1.10	41.7	1.10	44.3	1.39	50.6	1.00	63.7
Equation 5		Equation 6		Equation 7			
S.D.	t-stat.	S.D.	t-stat.	S.D.	t-stat.		
0.704	31.3	0.431	17.9	0.581	25.3		

We present the estimated correlation coefficients of the random effects in Table 4.5. Because two of the five variances of the random effects go to zero, we only have estimates for the correlations between the other three, a, λ, and m. a and m are highly negatively correlated, the correlation coefficient is -0.85. People with a small m will thus have a big a, perhaps because a and m both capture risk attitudes for gains, in two different domains. A smaller m means people are more risk-seeking in small probabilities and risk-averse in moderate and high probabilities. We don't discuss very small and big probabilities in this paper, and for moderate probabilities, a smaller m means that people are more underweighting the probabilities for gains, and are more risk-averse in the probability domain. A bigger a means that

[8] We use a different weighting function.

people have a less concave value function for gains, and are less loss-averse in the value domain. A highly negatively correlation between a and m might indicate that the level of risk aversion in the value and probability domains are complementary; if a person has a high risk aversion in one domain she should have a low risk aversion in another domain. This property might be useful to understand and predict the behavior of individual decision-making under risk. The correlation coefficient of the random effects of λ and m is 0.58, meaning that people with higher loss aversion might be less risk-averse in the probability domain.

Table 4.5: Correlation coefficients of random effects

	a	b	λ	m
b	-	-		
λ	-0.05 (-0.35)	-	1.0	
m	-0.85 (-21.3)	-	0.58 (8.99)	1.0
n	-	-	-	-

Using the model we estimated in this paper, we computed all five coefficients for each individual, using only one draw for the random effects. The descriptive statistics for the five estimated coefficients are presented in Table 4.6. The means of all these coefficients are almost the same as in Table 4.3, as expected. Figures 2 to 6 are kernel density estimates of the five random coefficients (estimated using Stata).

Table 4.6: Descriptive statistics for the five estimated coefficients

Variable	Median	Mean	Std. Dev.	Min	Max
a	0.679	0.683	0.153	0.118	1.26
b	0.737	0.730	0.032	0.565	0.775
λ	3.18	3.19	1.15	0.01	6.88
m	1.00	0.995	0.454	0.01	2.57
n	0.771	0.593	0.219	0.191	1.08

Figure 4.1 also shows the distribution of the estimated result of question 2 when randomization=1 (50% chance of losing 500), using the model we presented before and coefficients we estimated.

4.6 Conclusion

Cumulative Prospect Theory is becoming more and more prevalent in decision making under uncertainty, and our results show that a simple structure

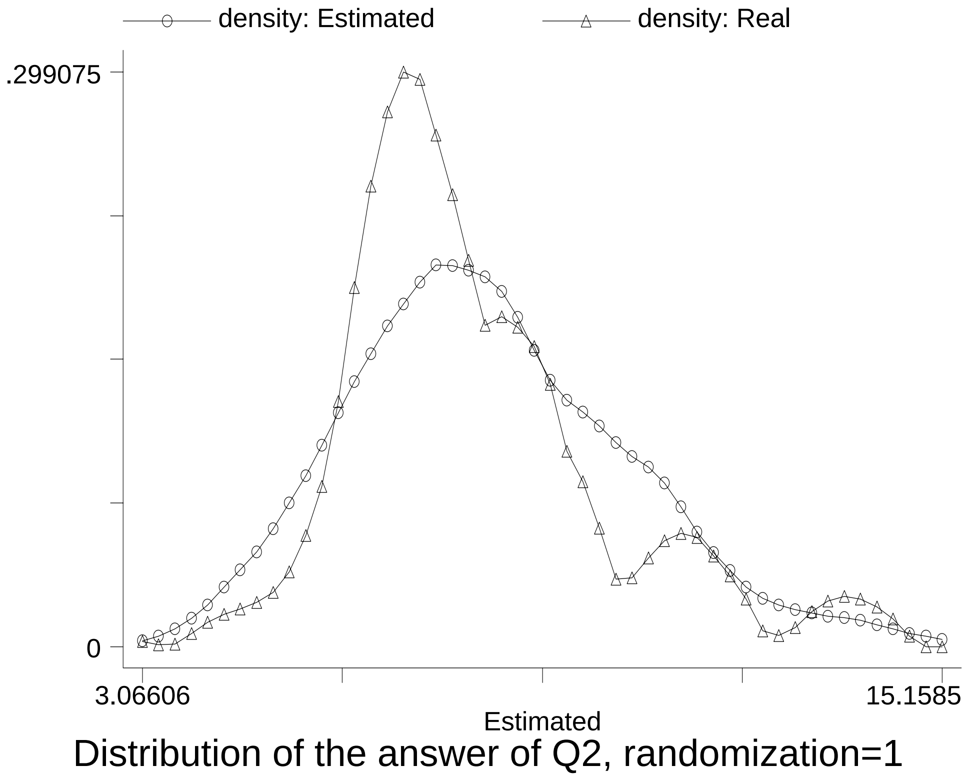

density: Estimated
density: Real
.299075
0
3.06606
15.1585
Estimated
Distribution of the answer of Q2, randomization=1

Figure 4.1:

of a CPT model with power value function, loss aversion, and probability weighting can explain our data well. A random coefficients model is employed in this paper to capture the heterogeneity of individuals in the population of interest.

With the power value function, we get on average powers for gains and losses are all between zero and one, and the power of gains is smaller than for losses, implying that the marginal utility of gains is diminishing faster than that of losses. The standard deviation of the power of losses is close to zero, so that its distribution is more concentrated around the median, while the power of gains is more dispersed. Females have a significantly smaller power of gains than males, indicating that females are more risk averse than males in the value domain. No demographic variable has a significant effect on the power of losses.

The mean of the coefficient of loss aversion we get is 3.1, a little bit bigger than the results in literature. But the results in the literature mostly come from decisions without uncertainty. Our result might indicate that uncertainty makes people more loss averse, or the coefficient of loss aversion grabs some of the risk aversion under uncertainty.

The parameter of the probability weighting function for gains, m, is close to one on average. Because its standard deviation is about half of its mean, it seems that m is varying heavily across individuals. Half of the people in our data have an m smaller than one, and half larger than one. We could conclude that half of Dutch population will overweight small probabilities and is risk-seeking in small probabilities.

Further research can focus on models with more complicated weighting functions both for gains and losses. With only one parameter for weighting functions, there are only two kinds of people considering their risk attitude to gains, either risk-seeking in small probabilities and risk-aversion in big probabilities ($m < 1$), or risk-aversion in small probabilities and risk-seeking in big probabilities for gains ($m > 1$). At least we should consider a model with four kinds of risk attitude to gains: the two kinds of people we mention above, plus people always risk averse and people always risk-seeking in all probabilities.

4.7 Appendix to Chapter 4

Exact Wordings of the Questions

The following questionnaire concerns your attitude towards events with uncertain outcomes. For this purpose we would like to ask you a number of questions concerning lotteries. In case you do not participate in lotteries out of principle, please indicate so here.

Out of principle, I never participate in lotteries: Yes/No

The lotteries that these questions refer to are not standard lotteries. In these lotteries you can win money, but you can also lose money, while participation to the lottery is for free. We would like to know under what conditions you are willing to participate in such lotteries. Most people will like to participate (for free) in a lottery where you can only win prizes, but where you cannot lose. In contrast, almost nobody would like to participate in a lottery where you always lose money. We are interested in the situation where you are indifferent between participating or not. It is important to realize that this will be different for everybody. The right answer will therefore also be different for different people.

Question 1: Imagine a lottery where you lose a certain amount of money with 50% chance and with 50% chance you win a certain amount of money. The probability of winning is the same as the probability of losing, but the amounts are not necessarily the same. In the first question, the amount you can lose is 100 euro (with 50% chance). This amount is in the table below in the left column. In the right column is the amount you can win (with 50% chance). For every combination of a loss and a gain, please indicate whether you would like to participate in this lottery or not.

You lose with 50% chance	You win with 50% chance	Will you participate?
100 euro	100 euro	Yes/No?
100 euro	150 euro	Yes/No?
100 euro	200 euro	Yes/No?
100 euro	250 euro	Yes/No?
100 euro	300 euro	Yes/No?
100 euro	350 euro	Yes/No?
100 euro	400 euro	Yes/No?
100 euro	500 euro	Yes/No?
100 euro	750 euro	Yes/No?

The answer “You win with 50% chance” is randomized as four choices:

1	2	3	4
100 euro	750 euro	100 euro	900 euro
150 euro	500 euro	200 euro	800 euro
200 euro	400 euro	300 euro	700 euro
250 euro	350 euro	400 euro	600 euro
300 euro	300 euro	500 euro	500 euro
350 euro	250 euro	600 euro	400 euro
400 euro	200 euro	700 euro	300 euro
500 euro	150 euro	800 euro	200 euro
750 euro	100 euro	900 euro	100 euro

In case a respondent answers yes to a question and later on no, ask whether she is certain about this.

In the previous question, you indicated that you would not want to participate when you could lose 100 euro and could win xx euro, but that you do want to participate when one could win yy euro instead of xx. Probably, there is a certain prize for which you would just like to participate in the lottery, but when the prize is 1 euro lower, you would not. This prize we call the lowest required prize for you to participate in this lottery.

Question 2: Imagine you could participate in a lottery where you lose X2 (randomized, 500 or 1000) with 50% chance. With the same probability you can win a certain amount of money. What is the lowest acceptable prize for you to participate in this lottery? .... Euro

Question 3: Imagine you could participate in a lottery where you lose 100 euro with P% (randomized, 20, 30 or 40) chance. With (100-P)% chance you can win a certain amount of money. The probability of loosing in this lottery is therefore smaller. What is in this case the lowest acceptable prize for you to participate in this lottery? .... Euro

Question 4: Imagine you could participate in a lottery where you lose 100 euro with P% (randomized, 80, 70 or 60) chance. With (100-P)% chance you can win a certain amount of money. The probability of loosing in this lottery is therefore larger. What is the lowest acceptable prize for you to participate in this lottery? .... Euro

Question 5: In the next question you are asked to choose between two lotteries, A and B. In lottery A you win 100 euro with 30% chance and nothing otherwise. In lottery B you win with 50% chance a larger amount of money, but at the same time you lose 100 euro with 50% chance. Please indicate in each row which lottery you prefer, A or B.

Lottery A		Lottery B		Your
You win with 30% chance	You win with 70% chance	You lose with 50% chance	You win with 50% chance	choice
100 euro	0 euro	100 euro	150 euro	A/B?
100 euro	0 euro	100 euro	200 euro	A/B?
100 euro	0 euro	100 euro	250 euro	A/B?
100 euro	0 euro	100 euro	300 euro	A/B?
100 euro	0 euro	100 euro	350 euro	A/B?
100 euro	0 euro	100 euro	400 euro	A/B?
100 euro	0 euro	100 euro	500 euro	A/B?
100 euro	0 euro	100 euro	700 euro	A/B?
100 euro	0 euro	100 euro	900 euro	A/B?

Same “soft” consistency check here.

The answer “Lottery B: with 50% chance you win” is randomized as four choices:

1	2	3	4
150 euro	900 euro	100 euro	900 euro
200 euro	700 euro	200 euro	800 euro
250 euro	500 euro	300 euro	700 euro
300 euro	400 euro	400 euro	600 euro
350 euro	350 euro	500 euro	500 euro
400 euro	300 euro	600 euro	400 euro
500 euro	250 euro	700 euro	300 euro
700 euro	200 euro	800 euro	200 euro
900 euro	150 euro	900 euro	100 euro

Question 6: In lottery A you can win X6 euro (randomized, 500 or 1000) with 10% chance. In case you do not win, you win or lose nothing. In lottery B you win with 50% chance X6 euro, but in case you do not win, you lose a certain amount of money. What is the highest acceptable loss for you to participate in this lottery? .... Euro

Question 7: In lottery A you can win X7 euro (randomized, 50 or 100) with 80% chance. In case you do not win, you win or lose nothing. In lottery B you win a certain amount of money with 50% chance, but in case you do not win, you lose 2*X7 euro. What is the lowest acceptable prize for you to participate in this lottery? .... Euro

Chapter 5

Modelling Mobility and Housing Tenure Choice

Housing tenure choice (renting or owning a dwelling) is an important decision for all households and closely linked to housing consumption. In most industrialized countries, housing expenditure makes up a substantial share of the household budget on average. For example, as a whole, households in the United States spent almost one third of their total budget on housing during the years 1998-2002[1]. Moreover, for home owners, the investment in the own home is typically much larger than the amount held in financial assets, such as saving accounts, stocks, or bonds, and a mortgage on the house is the most important form of debt holding.

Modelling the household's choice between renting and owning is not straightforward. A seminal article is Henderson and Ioannides (1983)[2]. In their theoretical framework, if there were no transaction cost, tax distortion or borrowing constraint, the household's tenure choice would be entirely determined by the difference between two kinds of demands for housing: investment demand and consumption demand. Ioannides and Rosental (1994) empirically analyzed the relationship of housing tenure choice, consumption, and investment demand for housing under this stylized theoretical structure in a static model.

Housing tenure choice is closely linked to residential mobility. We often observe that a household moves and changes its housing tenure choice at the same time. It is very difficult to distinguish the causality of moving and tenure-choice because of lack of data. Some households may decide to move first, and then choose between renting and owning. Changing the tenure choice first may be the motive for others; then, in most cases, in order to change the status from renter to owner or vice versa, they need to move. For a long time, the housing tenure choice and residential mobility have been studied separately in the literature. There are only few exceptions. Boehm (1981) is one of the first to recognize that the household decides on the housing tenure choice and residential mobility jointly, and employed a multinomial logit model to capture this feature. Ioannides and Kan (1996) used a dynamic multinomial probit model with random effects to estimate the probabilities of the choices *stay*, *move & rent*, and *move & own*. Their probit model allows for a more flexible error structure than the corresponding multinomial logit model. Kan (2000) used a dynamic random effects simultaneous equations model for the household's housing tenure choice, the residential mobility decision, and the expectation of future mobility.

In these dynamic models, a lagged dependent variable is used as one of

[1] From expenditure shares tables, 1998-2002. Bureau of Labor Statistics, U.S. Department of Labor. http://www.bls.gov/cex/csxshare.htm.

[2] See also Fu (1991).

the explanatory variables to model state dependence. The household's lagged tenure status and lagged moving decision in the first sample period cannot be observed, resulting in an initial conditions problem. In short panels, this needs to be accounted for in order to get consistent estimates of dynamic discrete choice models with random effects, as already discussed by Heckman (1981). Ioannides and Kan (1996), and Kan (2000) assume that the initial conditions are exogenous, leading to consistent estimates if the number of panel waves tends to infinity. Heckman (1981) shows that this problem may lead to biased estimates in short panels and suggest solutions for this.

This paper has a similar structure for the econometric model as Ioannides and Kan (1996): a dynamic binary probit model with random effects is used to model the household moving decision and a dynamic multinomial probit model with random effects is used to model the household tenure choice and residential mobility decisions jointly. The main methodological novelty of the present study compared to the existing panel data models of Ioannides and Kan (1996) and Kan (2000) is to account for the initial conditions in an appropriate way for a short panel: Following Heckman (1981), the initial conditions are modelled by using flexible reduced form equations. We use the method of maximum simulated likelihood and the GHK (Geweke - Hajivassiliou - Keane) simulator to estimate the model, allowing for a flexible structure of error terms.

The model is applied to the Dutch housing market, using an unbalanced panel containing ten annual waves. In line with finding in the literature, the probability to move is significantly lower for home owners than for renters, keeping other variables constant. We find a robust and interesting age pattern of the households' moving and housing tenure choice decisions. When the head of the household is young, the moving rate is high, and the household is more likely to rent. Around retirement age of the head of household, the probability of moving attains its minimum and the probability of ownership attains a maximum. After the usual retirement age, the moving rate increases and the household becomes more likely to change to renting.

The remainder of this paper is organized as follows. We give a brief description of the Dutch housing market and household moving rates in Section 2. We describe the data we use, the CentER Panel 1994-2003, in Section 3. We present the econometric models and discuss how these can be estimated in Section 4. In Section 5, we discuss the estimation results, and the conclusions are given in Section 6. In the appendix we introduce a behavioral model, which constitutes a theoretical framework for the econometric models.

5.1 The Dutch Housing Market and Moving Rates

It is well known that the social rented dwelling sector plays an important role in the Dutch housing market. After World War II, the Netherlands has experienced a housing shortage until recently; large numbers of social rented dwellings were built in order to solve this problem. The market share of the social rented dwelling sector had been growing continuously for more than 40 years since World War II, and reached its peak in the early 1990s, when 41% of the total housing stock belonged to this sector.

Social housing provides dwellings of reasonable quality at relatively low prices for families whose incomes are below some threshold. Other than in the United States, it does not only give access to the poor families: many Dutch households with median or even higher income live in social rented housing. Consequently, the home-ownership rate of Dutch households is still relatively low compared to the United States, although it has been rising remarkably in the last 40 years, see Figure 5.1[3]. The Dutch home-ownership rate was 54% in the year 2000 compared to 66.2% in the United States. See Priemus (1998, 2000), and van Kempen and Priemus (2002) for a more detailed discussion of the Dutch housing market[4].

Compared to the United States, the household moving rate, defined as the percentage of households that moved in a year, is very low in the Netherlands. It is only around 5% in the late 1990s, see Figure 5.2[5], less than one third of the United States' level[6]. Furthermore, there are relatively more renters in the Netherlands than in the United States. In addition, in the United States renters move much more frequently than owners. For example, in the year 1999, 34.7% of the renters and 8.2% of the owners moved in the previous

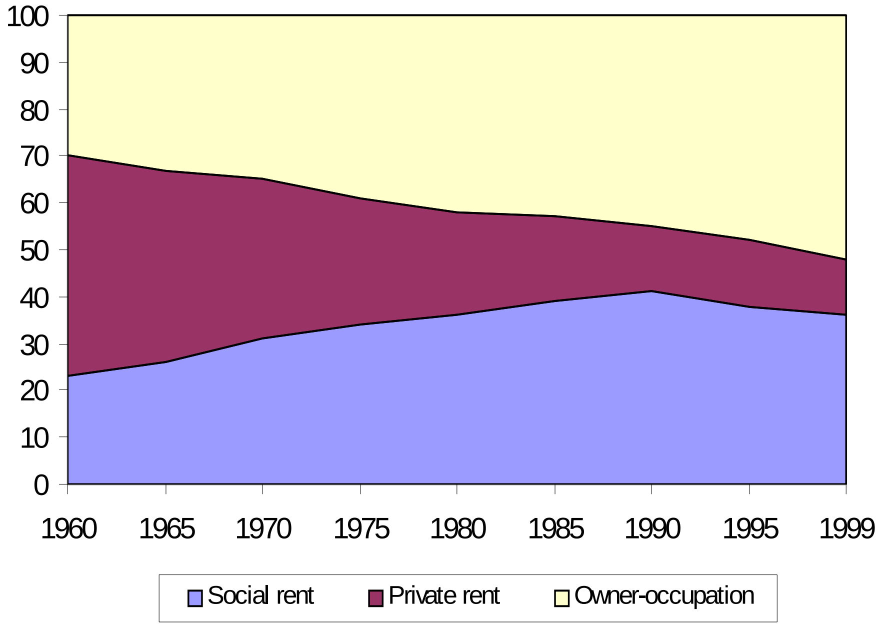
[3] The data in Table 1 are taken from van Kempen and Priemus (2002) (years 1960–85), from the Ministry of Housing (VROM; year 1989), and from Housing Demand Surveys (years 1990-1999) by Statistics Netherlands.

[4] The situation is similar in Germany, where owner-occupied housing has also increased over the past years, but remains at a rather low level, even compared to other EU countries. In 1998, the home-ownership rate in Germany was 40.9% (43.1% in former West-Germany). The data comes from the Federal Statistical Office Germany, 2002.

[5] We use the household moving rate for CentER Panel; for the CBS data the percentage of the total population moved (except those do not live within a family) in the previous year is used.

[6] In the United States, the residential housing market is extremely dynamic, the household moving rates are 17.6% and 17.0% in 1997 and 1999, respectively. See "The American Housing Survey", from the website of U.S. Census Bureau, in particular, for the year 1997: http://www.census.gov/hhes/www/housing/ahs/ahs97/ahs97.html and for the the year 1999: http://www.census.gov/prod/2000pubs/h150-99.pdf.

Figure 5.1: The changing structure of the Dutch housing market



year so that the moving rate of renters in 1999 was more than three times higher than that of owners.

The huge difference in the household moving rate between these two countries may be due to different structures of the housing market, geographic, cultural, and economic or social differences. The huge difference could be partly explained by the big market share of the social rented housing sector in the Netherlands. Quite a lot of evidence indicates that renters move less frequently and consume less on housing in rent-controlled housing markets than their optimal status without rent-control, see, for example, Olsen (1990), Clark and Heskin (1982) for the U.S., and Turner (1988) for Sweden. A detailed review of rent control can be found in Turner and Malpezzi (2003). The other reason might be that, compared to the United States, the Netherlands is a very small country with a convenient and concentrated public transportation system; as a result, even when people change their jobs, commuting is potentially a good alternative for moving.[7] Ekamper and van Wissen (2000) found some evidence of substituting commuting for domestic migration, pointing at the remarkable increase of the total number of commuting persons in the Netherlands after 1987, especially in the period 1989-1992. Though the overall increase of commuting diminished in the period 1992-1997, the increase of commuting was even higher for those regions with positive employment growth.

5.2 Data

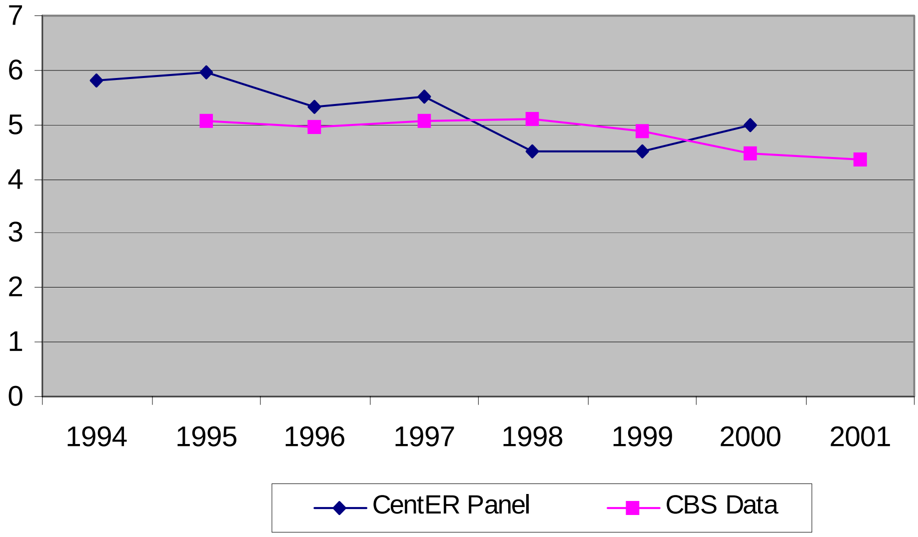
The data we use is an unbalanced panel with ten waves (1994-2003)[8] taken from the DNB Household Survey (formerly known as the VSB Panel, then the CentER Savings Survey). Data is collected every year for a panel of more than 1,500 households. The data contains rich information about employment, pensions, accommodation, mortgages, income, assets, debts, health, economic and psychological concepts, and personal characteristics.

The variables we use to estimate the model stem from several parts of the questionnaire. We employ a dynamic model and, therefore, only use the

[7] Job mobility in the Netherlands is of the same order of magnitude as in the US: In our data (1994-2000), at least one member of the household has changed job since the previous year in 17.2% households, on average; according to Ioannides and Kan (1996), 15.1% of the households in the U.S. (1970-1987) have job-changes every year.

[8] Data from the first two years (1993 and 1994) are merged into one wave. The survey of the 1994-wave was conducted over the period May through December 1994, while the 1993-wave had only finished by the end of April 1994. As a consequence, households who answered the questions about their accommodations in the 1993-wave did not answer the same questions again in the 1994-wave.

Figure 5.2: Dutch household moving rate (percent) in 1994-2001



households that participate in at least two consecutive waves. There is a very small number of households that change tenure status without moving, about 0.6% of all observations. These observations are deleted, and our models assume that households can only change tenure status by moving.[9] A total of 4,189 households and 15,320 observations are used for estimation. Table 5.1 shows the structure of the unbalanced panel that is retained for the estimations. The balanced subpanel of households that are included in the panel during all ten waves consists of only 141 households. Therefore, we will use the complete unbalanced panel in all the estimations.[10] The average time that one household stays in the panel is 3.7 years for the data we use.

As mentioned before, the moving rate is quite low in the Netherlands, around 5% per year on average over the years 1994-2003. In figure 5.2, we

[9]Dutch households rarely change their housing tenure choice without moving: renters can buy the dwellings they rent before and become owners; and owners can sell their dwellings and still live there as renters; owners can also change their housing consumption without moving by reconstructing their house, but we ignore them because we cannot observe these in our data. In recent years, the Dutch government started to encourage households to buy the dwellings they were renting, but in practice the sale of rented dwellings is negligible even in recent years. See van Kempen and Priemus (2002) for a more detailed discussion.

[10]The small number of observations in the 2000 wave is probably due to a change in interviewing technique; since 2000, the panel is internet based.

Table 5.1: Structure of the unbalanced panel

By wave		By number of waves		
Year	Number of Obs.	Number of waves	Obs.	Number of households
1994	2,113	2	2,992	1,496
1995	2,464	3	3,207	1,069
1996	2,203	4	2,604	651
1997	1,928	5	1,420	284
1998	1,550	6	1,890	315
1999	1,202	7	742	106
2000	791	8	704	88
2001	966	9	351	39
2002	1,111	10	1,410	141
2003	992			
Total	15,320	Total	15,320	4,189

Left panel: Year is the survey year, in total we have ten waves, 1994-2003.
Right Panel: Number of waves that the households stay in the panel.

compare our data with the aggregate moving rates published by Statistics Netherlands (CBS). CBS data comes from its online service, StatLine, the central database of Statistics Netherlands. The two rates are quite similar, suggesting that our data is representative for the residential mobility in the Netherlands.

Table 5.2 contains descriptive statistics on the dependent and independent variables used in the empirical models, the definitions of all variables can be found in the appendix. The total moving rate (m) can be decomposed in the moving rate of renters (7.3%) and that of owners (2.9%), showing that the group of renters is a more mobile group than the group of owners. If owners move, more than 86% of them keep the same tenure mode: homeowner as before; but for renters, less than half, only 48% retain their status as a renter. Table 5.3 contains the transition matrices of the household mobility and tenure choices. The moving rate is 4.4% for those households that did not move in the previous year, higher than the rate 2.7% of those that already moved in the previous year. Compared to the 3.7% of the households that change from renter to owner, only 0.4% of the households change from owner to renter.

Table 5.2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
m	0.0513	0.221	0	1
m_{t-1}	0.0535	0.225	0	1
tc	0.677	0.468	0	1
tc_{t-1}	0.674	0.468	0	1
age	4.83	1.39	1.9	10.2
age-sq.	2.53	1.42	0.361	10.4
school2	0.279	0.448	0	1
school3	0.417	0.493	0	1
$morc_{t-1}$	0.0874	0.282	0	1
nchild	0.783	1.12	0	7
nadult	1.76	0.467	1	6
div	0.0104	0.100	0	1
jobch	0.166	0.367	0	1
$retire_{t-1}$	0.0372	0.187	0	1
$logr_{t-1}$	0.00	0.210	-4.53	4.75
$ratio_{t-1}$	0.00	0.796	-0.994	4.41
$loghv_{t-1}$	0.00	.631	-5.22	4.42
$logw_{t-1}$	6.30	7.13	-13.86	15.01

Note: the lagged variable labelled $t-1$ has only 11,131 obs. while total obs. is 15,320.

Table 5.3: Transition matrixs of the househould mobility and tenure choice

Household mobility			Household tenure choice		
In the year t	In the year $t-1$		In the year t	In the year $t-1$	
	Stay	Move		Renter	Owner
Stay	95.6%	97.3%	Renter	96.3	0.4%
Move	4.4%	2.7%	Owner	3.7%	99.6%
Total	100%	100%	Total	100%	100%

5.3 Econometric Model

In this section we present the econometric models that we employ in our estimation. The appendix contains a behavioural model, which serves as a motivation for the econometric models.

Discrete choice models are often used to model the discrete alternatives chosen by decision-makers. Among cross-section data discrete choice models, the multinomial (and binomial) logit and probit model are the most popular ones. The multinomial logit model exhibits restrictive substitution patterns between choice options due to the Independence from Irrelevant Alternatives (IIA) property; its generalization to panel data has the drawback that it cannot be used when unobserved factors are correlated over time and/or alternatives for each decision-maker. The multinomial probit model can solve the IIA problem. The probit model is based on the assumption that the unobserved factors are jointly normally distributed, allowing for complicated patterns of correlation. Compared to the logit model, the flexibility in handling correlations over alternatives and time is the main advantage of the probit model. Therefore, we use a dynamic binary choice model to describe the moving decision and a dynamic multinomial probit model for the joint decision of moving and tenure choice.

5.3.1 A Binary Choice Model for Moving Decisions

In time t, the household can choose from two possibilities: *stay*, or *move*. The decision will be based upon the difference in utility between moving and staying. This difference in utility of the household can be written as a reduced form

$$U_{it} = U_{it}^{Move} - U_{it}^{Stay} = V_{it} + \varepsilon_{it} \qquad i = 1, \cdots, N;\ t = 1, \cdots, T$$

The household chooses to move if $U_{it} > 0$, resulting in $y_{it} = 1$, otherwise the household chooses to stay, i.e., $y_{it} = 0$. Here U_{it} is the difference in utility of household i in time period t consisting of a systematic part (V_{it}), representing the effects of observed factors, and an error term (ε_{it}), representing the unobserved factors. We postulate that the systematic part V_{it} includes family characteristics and the lagged dependent variable:

$$V_{it} = X_{it}'\beta + \gamma y_{it-1}.$$

where X_{it} is a vector of observed socioeconomic variables that may influence the household's moving decision, the lagged variable y_{it-1} captures the state dependence, and β and γ are parameters. Thus we can write

$$y_{it} = 1(X_{it}'\beta + \gamma y_{it-1} + \varepsilon_{it} > 0)$$

The parameter γ is expected to be negative: a household that just moved has adjusted its housing consumption recently towards an optimal value, and, unless a large shock takes place, it is unlikely to move again because moving is costly.[11]

We use a random effects structure for the unobserved factors, so $\varepsilon_{it} = \eta_i + \delta_{it}$, where η_i is a household-specific component which reflects a time invariant unobserved component of the cost of moving; δ_{it} is the time variant part, for which we allow for an AR(1) structure: $\delta_{it} = \rho_\delta \delta_{it-1} + \zeta_{it}$.

We also need to specify the initial conditions (ICs) for the dynamic process, see Heckman (1981), and Heckman and Singer (1986) for a detailed discussion on this. We follow the approach suggested by Heckman (1981), which consists of using a flexible reduced form to approximate the initial conditions. We implement this approach by using the initial period regressors (X_{i0}) for y_{i0} and allowing the initial period error term (ε_{i0}) to be correlated with other period errors (ε_{it}) in an arbitrary way. Hyslop (1999) uses the same approach and compares it with several other ways to deal with the initial conditions, estimating an intertemporal labor force participation model of married women.

With these inputs we can present the model. For households $i = 1, ..., N$, we model

$$\begin{aligned} y_{i0} &= 1(X_{i0}'\beta_0 + \varepsilon_{i0} > 0) \\ y_{it} &= 1(X_{it}'\beta + \gamma y_{it-1} + \varepsilon_{it} > 0) \text{ with } \varepsilon_{it} = \eta_i + \delta_{it},\ t = 1, \cdots, T \end{aligned}$$

where

$$\delta_{it} = \rho \delta_{it-1} + \zeta_{it}.$$

By assumption, the error term subvector

$$(\eta_i, \zeta_{i1}, \cdots, \zeta_{iT}, \delta_{i0})'$$

follows a $(T+2)$ −variate normal distribution, with zero mean vector, and covariance matrix the diagonal matrix with diagonal

$$\left(\sigma_\eta^2, \sigma_\zeta^2, \cdots, \sigma_\zeta^2, \sigma_\delta^2\right),$$

with

$$\sigma_\delta^2 = \frac{\sigma_\zeta^2}{1 - \rho^2}.$$

The resulting distribution of the subvector

$$\epsilon_i = (\varepsilon_{i1}, \cdots, \varepsilon_{iT})'$$

[11] Alternatively, habit formation could induce a positive effect, although we expect this is less relevant for moving than for, for example, consumption.

is then $T-$variate normal with zero mean vector, and covariance matrix having as components

$$cov(\varepsilon_{it}, \varepsilon_{is}) = \sigma_\eta^2 + \rho^{|t-s|}\sigma_\delta^2 = \sigma_\eta^2 + \rho^{|t-s|}\frac{\sigma_\zeta^2}{1-\rho^2}, \qquad t, s = 1, \cdots, T.$$

Combining with the ICs we postulate that

$$\begin{pmatrix} \varepsilon_{i0} \\ \varepsilon_i \end{pmatrix}$$

follows a $(T+1)-$variate normal distribution, with zero mean vector, and covariance matrix having as additional components

$$\begin{aligned} var(\varepsilon_{i0}) &= \sigma_0^2 \\ cov(\varepsilon_{i0}, \varepsilon_{it}) &= \rho_t\sqrt{\sigma_0^2\left(\sigma_\eta^2 + \sigma_\delta^2\right)} = \rho_t\sqrt{\sigma_0^2\left(\sigma_\eta^2 + \frac{\sigma_\zeta^2}{1-\rho^2}\right)}, \qquad t = 1, \cdots, T. \end{aligned}$$

For identification, we normalize both the variance of ε_{i0} and δ_{it} equal to 1, i.e., $\sigma_0^2 = 1$,and $\sigma_\delta^2 = 1$. This latter condition means $\sigma_\zeta^2 = 1-\rho^2$. In addition, in order to reduce the number of parameters we have to estimate, we assume that the absolute value of the correlations ρ_t between the errors in the initial period and the other periods, ε_{i0} and ε_{it}, $t = 1, ..., T$, is decreasing over time, i.e., $\rho_t = \rho_1^t$, where ρ_1 is the correlation between ε_{i0} and ε_{i1}.

Assuming a random sample (over the observations $i = 1, ..., N$) this model (and restricted versions of it) can be estimated by maximum simulated likelihood (MSL).

5.3.2 A Multinomial Probit Model for the Moving and Tenure Choice Decisions

We use a dynamic multinomial probit model for the joint decisions of moving and housing tenure choice. Households choose among three alternatives: *stay* (S), *move & rent* (MR), and *move & own* (MO). The utility household i gets from alternative j in time period t is

$$U_{it}^j = V_{it}^j + \varepsilon_{it}^j \qquad j = \{S, MR, MO\}\,;\ t = 1, \cdots, T;\ i = 1, \cdots, N.$$

Without loss of generality, the normalization $U_{it}^S = 0$ can be imposed, or, equivalently, U_{it}^{MR} and U_{it}^{MO} can be seen as the differences between the utility of alternatives MR and MO with the utility of the benchmark alternative S.

The choice of household i at time t is represented by two dummy variables, representing the moving decision (M) and the tenure choice decision (TC):

$$\begin{aligned} \mathit{Choice}\ S\ \ (\mathit{stay}) &: \quad M_{it} = 0, TC_{it} = TC_{i,t-1} \\ \mathit{Choice}\ MR\ (\mathit{move}\ \&\ \mathit{rent}) &: \quad M_{it} = 1,\ TC_{it} = 0 \\ \mathit{Choice}\ MO\ (\mathit{move}\ \&\ \mathit{own}) &: \quad M_{it} = 1,\ TC_{it} = 1 \end{aligned}$$

The observation rule for the multinomial probit model is given by

$$\begin{aligned} \mathit{Stay}\ \ (\ M_{it} &= 0),\ if\ U_{it}^{MR} < 0\ and\ U_{it}^{MO} < 0 \\ \mathit{Move}\ \&\ \mathit{own}\ \ (M_{it} &= 1,\ TC_{it} = 0),\ if\ U_{it}^{MR} > 0\ and\ U_{it}^{MR} - U_{it}^{MO} > 0 \\ \mathit{Move}\ \&\ \mathit{rent}\ \ (M_{it} &= 1,\ TC_{it} = 1),\ if\ U_{it}^{MO} > 0\ and\ U_{it}^{MO} - U_{it}^{MR} > 0 \end{aligned}$$

The utility of the alternatives *move & rent* (MR), and *move & own* (MO), are specified as follows:

$$\begin{aligned} U_{it}^{MR} &= V_{it}^{MR} + \varepsilon_{it}^{MR} = \left(x_{it}'\gamma^{MR} + \beta_1^{MR} M_{i,t-1} + \beta_2^{MR} TC_{i,t-1}\right) + \varepsilon_{it}^{MR} \\ U_{it}^{MO} &= V_{it}^{MO} + \varepsilon_{it}^{MO} = \left(x_{it}'\gamma^{MO} + \beta_1^{MO} M_{i,t-1} + \beta_2^{MO} TC_{i,t-1}\right) + \varepsilon_{it}^{MO} \end{aligned}$$

Here, V represents the systematic part and ε the error term. The systematic part is modeled to depend on observed characteristics (x_{it}) and past decisions ($M_{i,t-1}$ and $TC_{i,t-1}$). The past decisions ($M_{i,t-1}$ and $TC_{i,t-1}$) are included to represent state dependence. We expect that β_2^j, $j = MR,\ MO$, will be negative, because, in general, owners are less likely to move, both for *move & rent* and *move & own*. The signs of β_1^j, $j = MR,\ MO$, are not very clear. If we believe moving is costly, then β_1^j should be negative. But if we think there is "habit formation", people who moved often in the past are more likely to move and have lower moving cost than those who rarely move, then β_1^j might be positive. The sign of β_1^j depends on the magnitudes of these two effects.

Again following Heckman's (1981) approach, we formulate static ("reduced form") equations to approximate the initial conditions for M_{i0} and TC_{i0} at $t = 0$. This gives the following two binary probit equations for the initial conditions:

$$\begin{aligned} \text{Moving decision:} &\qquad M_{i0} = 1[x_{i0}'\beta^M + e_i^M > 0] \\ \text{Tenure choice decision:} &\qquad TC_{i0} = 1[x_{i0}'\beta^T + e_i^T > 0] \end{aligned}$$

Here e_i^M represents the unobserved factors in the tendency to move at $t = 0$, and e_i^T the unobserved factors in preferences for owning versus renting at $t = 0$.

We denote the vector of error terms of household i for the alternatives $j = MR$, and MO, in all time periods t, $t = 1, \cdots, T$, as $\varepsilon_i = (\varepsilon_{i1}^{MR}, \varepsilon_{i1}^{MO}, \varepsilon_{i2}^{MR}, \varepsilon_{i2}^{MO}, \cdots\cdots, \varepsilon_{iT}^{MR}, \varepsilon_{iT}^{MO})'$, and we write $e_i = \left(e_i^M, e_i^T\right)'$. As in the binary choice model, we decompose the error term ε_{it}^j, $j = MR,\ MO$, into two parts: $\varepsilon_{it}^j = \eta_i^j + \delta_{it}^j$. The random effects part, η_i^j, does not change over time. We assume that the time varying part δ_{it}^j has an autoregressive error AR(1) structure:

$$\delta_{it}^j = \rho_j \delta_{i,t-1}^j + \zeta_{it}^j.$$

Imposing distributional assumptions analogously to the binary choice case, we obtain that ε_i follows a $2T-$variate normal distribution, with zero mean vector and a covariance matrix with components

$$cov(\varepsilon_{it}^j, \varepsilon_{is}^k) = cov(\eta_i^j, \eta_i^k) + \rho_j^{t-s} \frac{cov(\zeta_{is}^j, \zeta_{is}^k)}{1 - \rho_j \rho_k}, \quad \text{if } t \geqslant s \geqslant 1,\ j, k \in \{MR, MO\}$$

The vector $(e_i', \varepsilon_i')'$ then follows a $2T + 2-$variate normal distribution, where we assume that the correlations between the ICs error terms and the other periods' error terms are decreasing over time, such that:

$$corr(e_i^\ell, \varepsilon_{it}^j) = \rho_{\ell j}^{t-1},\ \ell = M, T;\ j = MR, MO$$

where $\rho_{\ell j}$ are correlation coefficients between the error terms of ICs and the first period. For the sake of normalization we set the variances of e_i^M, e_i^T, and ζ_{it}^{MR} equal to 1.

Overall, we use model the covariance matrix of dimension 20×20 (using $T = 9$), with 12 parameters in total:

- the variances of η_i^{MR}, η_i^{MO}, and ζ_{it}^{MO} ($\sigma^2_{\eta^{MR}}$, $\sigma^2_{\eta^{MO}}$, and $\sigma^2_{\zeta^{MO}}$; $\sigma^2_{\zeta^{MR}} = \sigma^2_{e^M} = \sigma^2_{e^T} = 1$ are normalized);
- the autocorrelation coefficients of the error terms ζ_{it}^{MR}, and ζ_{it}^{MO} (ρ_{MR} and ρ_{MO});
- the correlation coefficients between the random effects η_i^{MR} and η_i^{MO} (ρ_η) and between ζ_{it}^{MR}, and ζ_{it}^{MO} (ρ_δ);
- the correlation coefficients between e_i^M and e_i^T ($\rho_{M,T}$);
- and the correlation coefficients between the error terms of the ICs e_{i0}^M and e_{i0}^T and the other periods errors ε_{it}^{MR} and ε_{it}^{MO} ($\rho_{M,MR}$, $\rho_{M,MO}$, $\rho_{T,MR}$, and $\rho_{T,MO}$).

Again, we use the GHK simulator and the method of maximum simulated likelihood (MSL) to estimate the model. Train (2003) gives a good introduction of how to use the GHK simulator and MSL to estimate the multinomial probit model. Also see Hajivassiliou and Ruud (1994).

5.4 Results

5.4.1 Binary Probit Models for Moving Decision

The results of binary probit models for the moving decision are presented in table 5.4. For the sake of comparison, we estimate the model with and without initial conditions (IC). When we estimate the model, we also use eight time dummy variables, which are dummies for different waves, in order to allow for a time effect. In the version without IC, 11,131 observations are used. For the version with ICs, more observations, namely, 15,320 are used because of the initial observations.[12]

Most results of the version without ICs are quite similar to the one estimated with the ICs, but the parameters of lagged moving (m_{t-1}) and ρ, the autocorrelation coefficient of the error terms δ_{it}, are different. The estimated parameter of m_{t-1} in the model with ICs (-0.49) is much less significant and has much smaller absolute value than in the model without ICs (-0.73). The same applies to ρ.

As discussed in the previous section, a negative effect of the lagged dependent variable m_{t-1} can reflect substantial moving costs that make it unattractive to adjust housing consumption shortly after it has been adjusted. Apparently, the evidence for such an effect is much weaker in the model that takes account of the ICs. It might also be the case that this effect is partly cancelled by the positive effect of habit formation. Households that recently moved may be able to move more efficiently (i.e., have lower cost of moving) than those who did not recently move.

Without ICs there seems to be a clear significant time persistence effect, both through state dependence and through individual heterogeneity via the autocorrelation in δ_{it}. Notice that the estimations of σ_η are close to zero, suggesting that the individual effect is much less important than the time varying errors δ_{it}, which have normalized variance 1. However, taking into account the ICs, the time persistence parameters are not significant anymore.

[12] Due to lack of data on earlier points in time, the lagged variables cannot be included among the regressors in the initial condition equation. This means that we do not include eight variables: m_{t-1}, tc_{t-1}, morc_{t-1}, retire_{t-1}, logr_{t-1}, ratio_{t-1}, loghv_{t-1}, and logw_{t-1} in the ICs equation of the model with ICs.

Possible inconsistency due to omitted ICs might thus particularly affect the understanding of state dependence and time persistency of the errors.

Table 5.4: Moving Decision Model

Variable	Without ICs		With ICs			
			Initial			
	Para.	t-st.	Para.	t-st.	Para.	t-st.
m_{t-1}	-0.733	-2.99	-	-	-0.491	-1.30
tc_{t-1}	-0.507	-8.83	-	-	-0.502	-8.74
age	-0.622	-5.12	-1.13	-7.02	-0.583	-4.64
age-sq.	0.442	3.64	0.915	5.30	0.410	3.32
school2	0.051	0.79	-0.085	-1.00	0.052	0.82
school3	0.275	4.52	-0.024	-0.31	0.273	4.50
morc_{t-1}	0.171	2.00	-	-	0.173	2.04
nchild	-0.053	-2.31	-0.075	-2.29	-0.051	-2.20
nadult	0.277	5.43	0.034	0.47	0.271	5.27
div	0.368	1.89	0.365	1.67	0.366	1.90
jobch	0.206	3.42	0.173	2.30	0.207	3.42
retire_{t-1}	0.164	1.23	-	-	0.166	1.23
logr_{t-1}	0.177	1.86	-	-	0.175	1.84
ratio_{t-1}	-0.167	-2.72	-	-	-0.167	-2.73
loghv_{t-1}	-0.149	-2.26	-	-	-0.149	-2.27
logw_{t-1}	0.002	0.46	-	-	0.002	0.47
constant	-0.140	-0.47	1.24	3.19	-0.265	-0.83
σ_η	0.001	0.02			0.00	0.01
ρ_1					-0.241	-1.39
ρ	0.243	2.00			0.155	0.84
log-likelihood	-1778.5		-2723.3			

Note: 11,131 obs. in the model without ICs, 15,320 with ICs.

Other than in the complete model discussed below, tenure status is assumed to be strictly exogenous here. As indicated by the regression coefficient of tc_{t-1} and consistent with our prior expectations, the probability to move is significantly lower for home owners than for renters, keeping other variables constant. This corresponds to the notion that moving costs, including costs of selling the current dwelling, are higher for owners than for renters. The small and insignificant parameter on retire_{t-1} shows that retirement has no direct effect on the probability of moving. The significant parameters on age (negative) and age-squared (positive) indicate that the probability of moving is higher when the household head is younger, and falls with age until the head of the household is about 71 years old. The model without ICs gives almost the same estimation of the point where the minimum is attained, 70 years old. This result is consistent with the notion that younger households more often adjust their housing consumption, due to, for example, the needs of the children, an increasing income pattern, or diminishing importance of liquidity constraints. After housing consumption is adjusted, the moving intensity decreases. Around 70 years old, the household is in its most stable phase of its life-cycle, and the probability of moving is the lowest. Ioannides and Kan (1996) obtain two significant parameters of age and age-squared with the same signs as ours. Normally, 70 years old is five to ten years after retirement, when income typically falls, so people of age around 70 might already have used some part of their saving to support their non-housing consumption. They are more likely to move to a cheaper place when they are getting older, and sell the former expensive dwellings if they are owners, e.g. to save money from housing consumption to support non-housing consumption in old age.

Job changes and changes in household composition also affect the probability of moving. If there is a job change, the household is more likely to move, as expected; a divorce also increases the probability of moving, but this is significant only at the 10% level. With more adults in the household, the probability of moving increases; and moving becomes less likely if the number of children is larger. The latter result may be explained by the size of the perceived moving costs, which can be higher if children have to move schools, etc. The dummy variable school3 points out that the households with a higher-educated head are more likely to move, perhaps because their income is higher. The parameter of morc_{t-1} shows that if the household faced binding mortgage constraints in the previous year, it is more likely to move and adjust its housing consumption.

If the moving cost is an increasing function of housing consumption, h_t, the parameters of logr_{t-1} and loghv_{t-1} should be negative. But we do not obtain this kind of result. The positive parameter of logr_{t-1} (0.18, significant at the 10% level) and the negative one of loghv_{t-1} (-0.15) show that renters and owners may have different structures of moving costs. A renter with higher rent is more likely to move than a renter with a low rent. An explanation may be that the low rent families often live in socially protected housing with an artificially low rent. If they move, they will either have to pay a much higher rent or they will have to buy a house. Alternatively, a high rent may indicate borrowing constraints, uncertainty, or expected changes. Moreover, renters with a high rent are probably more likely to buy a house and move into their own dwelling. We can investigate this in more detail in the next subsection.

For owners, it is quite clear that a higher house value means a higher cost of moving. This corresponds to what we find: with a more expensive house, the owner is less likely to move. The variable ratio_{t-1} shows the household's expected return on its dwelling, it has the same influence on the moving decision as the house value. The significant negative effect of ratio_{t-1} implies that with a higher expected return, the household might expect the price of its dwelling to continue to increase in the future, leading to a smaller probability to move. Ioannides and Kan (1996) get insignificant parameter estimates on the real rent and the real house value, indicating that there might be different moving cost functions for the U.S. and the Netherlands. One such a difference may be the tax structure: every home purchase in the Netherlands is taxed with 6% of the sales price.

5.4.2 Tenure Choice and Household Mobility

We use a dynamic multinomial probit model with the initial conditions to estimate the joint decision of housing tenure choice and household mobility. The results are presented in Tables 5.5 and 5.6. We also use eight time dummies in order to capture the fluctuations in the housing market over time.

As we expected, the parameters of TC_{t-1}, β_2^j, are significantly negative, both for *move & rent* and *move & own*; this implies that owners are less likely to move than renters. For M_{t-1}, the parameter in case of *move & own* is significantly negative, but in case of *move & rent* the parameter is close to zero. These outcomes mean that if a household moved in the previous period, this does not change the probability of *move & rent*, but the possibility *move & own* becomes less likely.

As seen in Table 5.6 the variances of the two random effects η_i^{MR} and η_i^{MO} are close to zero, like in the binary choice model for the moving decision, suggesting that the individual effect is much less important than the time-varying errors δ_{it}^{MR} and δ_{it}^{MO}. There is a highly significant autocorrelation structure for the error terms δ_{it}^{MO} but not for δ_{it}^{MR}.

Thus, time persistency seems to be clearly present, both via path dependence and via unobserved heterogeneity. However. in case of *move & rent*, the main channel of time persistency seems to be the path dependence represented by tenure choice, while in case of *move & own* the path dependence is via both path dependence and via unobserved heterogeneity. In both cases the unobserved heterogeneity is via the household and time specific idiosyncratic effects (δ), and not via the household specific effects (η). The highly positive correlation coefficient, 0.92, of the unobserved error terms between *move & rent* $\left(\delta_{it}^{MR}\right)$and *move & own* $\left(\delta_{it}^{MO}\right)$ is then what we expect: it means

Table 5.5: Result of multinomial probit model for MR and MO

	Initial conditions				Move		Move	
Variable	Moving		Own or rent		& rent		& own	
	Para.	t-st.	Para.	t-st.	Para.	t-st.	Para.	t-st.
m_{t-1}					-0.008	-0.03	-0.729	-5.68
tc_{t-1}					-0.750	-7.72	-0.273	-4.04
age	-1.16	-7.37	1.26	12.1	-0.448	-3.27	-0.186	-2.06
ages	0.943	5.59	-1.17	-11.2	0.375	2.76	0.102	1.17
school2	-0.106	-1.31	0.316	5.65	-0.018	-0.22	0.047	1.20
school3	-0.014	-0.19	0.487	9.22	0.108	1.35	0.112	2.69
$morc_{t-1}$					0.107	0.63	0.035	0.73
nchild	-0.063	-1.98	0.151	7.21	-0.040	-1.21	-0.012	-0.88
nadult	0.042	0.60	0.804	18.0	0.163	2.97	0.113	2.85
div	0.351	1.67	-0.207	-1.33	0.506	2.32	0.108	0.79
jobch	0.101	1.48	-0.215	-3.68	0.231	2.79	0.044	1.37
retire					0.001	0.01	0.102	1.18
$logr_{t-1}$					-0.174	-1.66	0.076	1.28
$ratio_{t-1}$					-0.099	-0.90	-0.031	-0.90
$loghv_{t-1}$					-0.165	-1.11	-0.045	-1.27
$logw_{t-1}$					-0.001	-0.20	0.000	-0.02
constant	1.27	3.34	-4.28	-17.0	-0.754	-2.15	-0.123	-0.66
Loglike.				-4876.8				

that the probabilities of *move & rent* and *move & own* always change in the same direction if there is an unexpected shock.

Table 5.6: The parameters of the variance matrix

ICs			Multinomial Probit		
Variable	Para.	t-st.	Variable	Para.	t-st.
$\rho_{M,T}$	-0.01	-0.18	$\sigma_{\eta^{MR}}$	0.00	0.01
$\rho_{M,MR}$	-0.09	-5.83	$\sigma_{\eta^{MO}}$	0.00	0.00
$\rho_{M,MO}$	0.57	26.9	$\sigma_{\zeta_{MO}}$	0.41	4.94
$\rho_{T,MR}$	0.05	0.60	ρ_{MR}	-0.05	-0.78
$\rho_{T,MO}$	0.24	5.18	ρ_{MO}	0.56	24.9
			ρ_{η}	0.00	0.00
			ρ_{δ}	0.92	25.4

Two kinds of moves can be distinguished: moving within the region (intra-regional moves) and moving between different regions (inter-regional moves). When the household wants to move to another region, it often cannot directly find a suitable dwelling to rent or buy in that region. If the incentive to move is strong enough, the household might then first move to an imperfect dwelling in the new region, and will then adjust its housing consumption again after it has found a more suitable dwelling, thus, moving within the same region. This implies that an inter-regional move is often followed by an intra-regional move[13]. According to our data, the Netherlands shows a similar job mobility as the United States. But the Netherlands is a small country, and commuting can be a good substitute for moving, even for inter-regional moving. This makes it less likely to observe an inter-regional move followed by an intra-regional move. This might explain why the moving rate in the Netherlands is so low. Only 0.89% of the households move more than once in our data, and only 0.33% households have moved twice in two subsequent years, providing some evidence that the Dutch households do not need to move twice if they want to change region. This makes the moving decision much more important than the housing tenure choice in the Netherlands, and might explain why the probabilities of *move & rent* and *move & own*

[13] In the US, around 60% of all moves are within county moves. For example, between March 1999 and March 2000, 56% of those moves were local (within the same county), 20% were between counties in the same state, and only 19% were moves to a different state. This data of moving is based on individuals, not household, but the results should be similar. Data comes from Current Population Reports, Geographical Mobility (P20), U.S. Census Bureau, http://www.census.gov/prod/2001pubs/p20-538.pdf.

change in the same directions, with a highly positive correlation coefficient of the error terms.

The parameters of age and age-squared are all significant except the one of age-squared for *move & own*. The initial conditions of moving show a similar age-pattern as in the binary moving model: the probability of moving is higher when the household head is younger, decreasing with age until the head of the household is retired, and increasing with age beyond retirement age. The initial conditions of owning or renting have an opposite age-pattern to moving. This is consistent with the finding that owners move less frequently than renters. The probability of owning a dwelling is low when the household head is younger, increases with age until the head of the household is around retirement age, and falls with age when the head of the household is beyond retirement age. The probabilities of *move & rent* changes with a similar age-pattern like the ICs of moving, but *move & own* has a slightly different pattern: the probability of *move & own* is always decreasing with age.

School3 makes the household more likely to *move & own*. Several variables, like nchild, morc_{t-1}, retire, logr_{t-1}, ratio_{t-1}, loghv_{t-1} and logw_{t-1}, they all have no significant effect both on *move & rent* and *move & own*. With more adults in the household, the household is more likely both *move & rent* and *move & own*. But divorce and job-changing make the household more likely to *move & rent* rather then *move & own*.

5.5 Conclusion

Where and how to live are important decisions for all households. In this paper we estimated, using the method of simulated maximum likelihood, a dynamic binary probit model with random effects to quantify the household moving decision and a dynamic multinomial probit model with random effects to quantify the household tenure choice and residential mobility jointly. We estimated the models using a Dutch unbalanced panel consisting of ten waves. The main methodological novelty of our paper is to account for the initial conditions in an appropriate way, as is required in such a relatively short panel, following the suggestion in Heckman (1981).

Our empirical findings can be summarized as follows. First, in case of the binary choice model, we compare the outcomes with and without initial conditions (ICs). Most results of the version without ICs are quite similar to the ones estimated with the ICs, but the inference on the parameter of lagged moving (m_{t-1}) as well as the autocorrelation of the individual and time specific error terms turn out to be clearly different. This suggests that the inference on time persistency and state dependence is sensitive to the presence or absence of the initial conditions, indicating the importance of including and modelling ICs in an appropriate way.

Secondly, the time persistency seems to be absent in the binary choice model with ICs. But in case of the multinomial probit model, time persistency seems to be clearly present, both via path dependence and via unobserved heterogeneity in case of *move & own* and mainly via path dependence represented by tenure choice in case of *move & rent*. The unobserved heterogeneity seems to be mainly via the time specific idiosyncratic effects, and not so much via the household specific effects.

Thirdly, and in line with our expectations, we find that the probability to move is significantly lower for home owners than for renters, keeping other variables constant. This corresponds to the notion that moving costs, including costs of selling the current dwelling, are higher for owners than for renters.

Finally, we get a robust age pattern of the households' housing decisions in our models. In the binary choice model, we find that the probability of moving falls with the age of the head of household, until the head reaches the age of 70, when the probability of moving starts increasing again. This age-pattern is confirmed in the multinomial probit model. So, when the head of the household is young, the moving rate is high, and the household is more likely to rent, while round retirement age, the household has the smallest probability of moving and the highest probability of owning a dwelling. After retirement age, the moving rate increases and the household becomes more likely to change to renting.

5.6 Appendix to Chapter 5

Appendix A: Variable Definitions

m	Dummy variable of moving: 0 stay; 1 move.
m_{t-1}	Lagged dummy variable of moving.
tc	Dummy variable of the housing tenure choice: 0 rent; 1 own.
tc_{t-1}	Lagged dummy variable of tenure choice.
age	Age of the head of household / 10.
age-sq	Age square of the head of household / 1000.
school2	Dummy variable of head-of-hh-s education level 2: 0 no, 1 yes.
school3	Dummy variable of head-of-hh-s education level 3: 0 no, 1 yes.
morc	Mortgage constraint, comes from the question: "Would you have chosen to buy a more expensive house if you had been able to receive a larger mortgage loan on the basis of your income at that time?" 0 no 1 yes.
$morc_{t-1}$	Lagged dummy of mortgage constraint: 0 no, 1 yes; 0 for all renter.
nchild	Number of children in the household.
nadult	Number of adults in the household.
div	Dummy variable of divorce in last three years: 0 no, 1 yes.
jobch	Dummy variable of job changing in the previous year: 0 no, 1 yes.
retire	Dummy variable of retirement in the previous year: 0 no, 1 yes.
logr	(log rent) – (mean of log rent) for renter; 0 for owner.
$logr_{t-1}$	Lagged logr.
ratio	(ratioexp) – (mean of ratioexp) for owner, 0 for renter. ratioexp=expected selling price of the house now / house value, using 3% as the appreciation rate of the house value
$ratio_{t-1}$	Lagged ratio.
loghv	(log house value) – (mean of log house value) for owner, 0 for renter.
$loghv_{t-1}$	Lagged loghv.
$logw_{t-1}$	Lagged variable of log total wealth.

Appendix B: Behavioral Model

In this appendix, we present a dynamic behavioral model in which a household maximizes its utility of housing and non-housing consumption and decides on the housing tenure choice and residential mobility simultaneously. We use the structure of expected utility over an infinite lifetime to model the household's problem of utility maximization. We assume that every household can have only two kinds of assets: one is the risk free financial asset with a fixed return, given by the interest rate r; the other one is the house with an uncertain return. We assume that preferences are intertemporally separable. In each period, every household makes a decision about whether to move or not and whether to rent or to own a dwelling, and decides on its housing and non-housing consumption at the same time.

We assume that the household's utility is only determined by its housing and non-housing consumption. The moving decision and housing tenure choice do not enter the utility function directly, they will affect the household's utility only through the budget constraint. In each period, the household will choose its non-housing consumption c_t and housing consumption h_t conditional on its moving decision and tenure choice together, so as to maximize the expected utility of its remaining lifetime. The household's expected utility of its remaining lifetime at time t is given by

$$\sum_{\tau=t}^{\infty}\delta^{\tau-t}E_tU(c_\tau,h_\tau;X_\tau)=U(c_t,h_t;X_t)+\sum_{\tau=t+1}^{\infty}\delta^{\tau-t}E_tU(c_\tau,h_\tau;X_\tau)$$

where $U(\cdot)$ is the utility in one period, X_t is a vector of characteristics of the household in period t, capturing shifts of the household's tastes, and δ is the discount factor of utility, representing the household's rate of time preference. We assume that the household can only change the housing consumption and tenure choice by moving. In each period, the household can decide to move to a rented or owned dwelling in order to adjust its housing and non-housing consumptions jointly and maximize the expected utility of its remaining lifetime. We also assume that the household can only move or change its tenure choice once in a period[14]. Following the standard approach, an intertemporal budget constraint with an infinite time horizon is used, abstracting from mortality risk and bequest motives.

In each period, combining the decisions of housing tenure choice and household mobility, there are three options for all households, both for renters and owners. A renter can choose among: staying at the same place and

[14] Households might move more than once in one period (a year), but it is not so common and we did not observe it in our data.

remaining as a renter, moving and renting a new place, or moving and becoming an owner. Similarly, a homeowner can also choose to stay at the same dwelling and remain as an owner, to move and rent a new place (and thus become a renter), or to move and own a new place. We assume that the household solves the problem of the housing tenure choice and moving decision simultaneously in two steps. In the first step, it solves the utility maximization problem for each possible choice (stay, move & rent, move & own) separately, conditional on its former tenure choice, moving decision, its characteristics and given the budget constraint (which we will discuss it in detail later on). In the second step, the household will choose one of its three available alternatives, which gives it the maximum utility.

Let R_t denote the renting price of per unit of housing consumption for the dwelling in period t, A_t is the household's total amount of financial assets in period t. If A_t is a negative number, it could be interpreted as a debt, for example, the household might have a mortgage loan for buying a house[15]. y_t is the labor income in period t, and r is the interest rate, which is assumed a constant over time. We use two dummy variables, TC_t and M_t, to represent the household's tenure choice and moving decision respectively. If the household is an owner in period t, the tenure choice $TC_t = 1$; otherwise $TC_t = 0$. If the household moves in period t, the moving dummy satisfies $M_t = 1$; otherwise $M_t = 0$.

Ioannides and Kan (1996) used a similar standard life-cycle model, but with a relatively simple budget constraint. Compared to their model, we take advantage of a more complicated budget constraint with dummy variables in order to model the decision procedure more clearly in a single equation. Ignoring the maintenance costs of owner-occupied housing and restrictions on the mortgage, the budget condition is given by a single equation as follows:

$$\begin{aligned} & \sum_{\tau=t}^{\infty}\frac{1}{(1+r)^{\tau-t}}[c_\tau + (1-[TC]_\tau)h_\tau R_\tau + M_{\tau-1}[TC]_\tau h_\tau P_\tau + M_\tau mc_\tau] \\ = \quad & A_{t-1} + \sum_{\tau=t}^{\infty}\frac{1}{(1+r)^{\tau-t}}[y_\tau + M_{\tau-1}[TC]_{\tau-1}h_{\tau-1}P_\tau] \end{aligned}$$

The left hand side of the equation is the total life-cycle consumption plus actual moving cost in case of moving. c_τ is the household non-housing consumption in period τ. $(1-[TC]_\tau)h_\tau R_\tau$ is the cost of renting if the household is a renter, 0 for an owner. $M_{\tau-1}[TC]_\tau h_\tau P_\tau$ is the price of buying a dwelling if the household moved in the previous time and is an owner now. Finally, $M_\tau mc_\tau$ is the moving cost if the household decides to move. The right hand

[15] We assume that households will have no assets left at the end of their lifetme.

side of the equation is the total life-cycle income. A_{t-1} is the initial assets in time $t-1$; y_τ is the household income in period τ, assumed to be given exogenously; and $M_{\tau-1}[TC]_{\tau-1}h_{\tau-1}P_\tau$ is the price of selling the previous dwelling if the household was an owner and moved.

At the time t, a former renter in previous period ($tc_{t-1} = 0$) will choose the option among the following three choices combining housing tenure choice and moving decision together: stay as a renter, move and rent, and move and own, which gives a maximum utility for the renter. Note that the budget constraint depends on which option is chosen.

1. Stay as a renter, then $h_t = h_{t-1}$, $tc_t = tc_{t-1} = 0$, $m_t = 0$,

$$\max_{\{c_t,h_t,tc_t,m_t\}} U(c_t, h_t = h_{t-1}) + \beta E_t V(c_t, h_t = h_{t-1}, tc_t = 0, m_t = 0; X_{t+1})$$

   where $E_t V(\cdot)$ is the household's expected utility of the remaining lifetime.

2. Move and rent, then $tc_t = tc_{t-1} = 0$, $m_t = 1$,

$$\max_{\{c_t,h_t,tc_t,m_t\}} U(c_t, h_t) + \beta E_t V(c_t, h_t, tc_t = 0, m_t = 1; X_{t+1})$$

3. Move and own, then $tc_t = 1$, $tc_{t-1} = 0$, $m_t = 1$,

$$\max_{\{c_t,h_t,tc_t,m_t\}} U(c_t, h_t) + \beta E_t V(c_t, h_t, tc_t = 1, m_t = 1; X_{t+1})$$

Similarly, a former owner in previous period ($tc_{t-1} = 1$) chooses the alternative among the following three alternatives that gives the highest utility subject to the budget constraint.

1. Stay as an owner, $h_t = h_{t-1}$, $tc_t = tc_{t-1} = 1$, $m_t = 0$

$$\max_{\{c_t,h_t,tc_t,m_t\}} U(c_t, h_t = h_{t-1}) + \beta E_t V(c_t, h_t = h_{t-1}, tc_t = 1, m_t = 0; X_{t+1})$$

2. Move and rent, $tc_t = 0$, $tc_{t-1} = 1$, $m_t = 1$

$$\max_{\{c_t,h_t,tc_t,m_t\}} U(c_t, h_t) + \beta E_t V(c_t, h_t, tc_t = 0, m_t = 1; X_{t+1})$$

3. Move and own, $tc_t = tc_{t-1} = 1$, $m_t = 1$

$$\max_{\{c_t,h_t,tc_t,m_t\}} U(c_t, h_t) + \beta E_t V(c_t, h_t, tc_t = 1, m_t = 1; X_{t+1})$$

The main link between the theoretical model and the econometric model is that moving costs enter the econometric model through state dependence: if moving is costly, tenure choice in this period will affect tenure choice and moving decision next period. Moreover, adjustment of housing consumption will be lumpy: instead of moving every year to adjust housing consumption immediately and gradually, it will be optimal to move less often and then adjust housing consumption substantially. This may lead to a negative state dependence in moving decisions: people who have just moved are less likely to move again in the next year.

Bibliography

Abdellaoui, Mohammed (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. Management Science, 46 (11), pp. 1497-1512.

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

Barberis, Nicholas, and Richard Thaler (2003). A Survey of Behavioral Finance, in the Handbook of the Economics of Finance, Edited by G.M. Constantinides, M. Harris and R. Stulz, Elsevier Science B.V.

Bateman, Ian, Alistair Munro, Bruce Rhodes, Chris Starmer, and Robert Sugden (1997). A Test of the Theory of Reference-Dependent Preferences. Quarterly Journal of Economics,112 (2), pp. 479-506.

Becker, Gary, and Casey Mulligan (1997). The Endogenous Determination of Time Preference. Quarterly Journal of Economics, 112, pp. 729-758.

Becker, Gary, and Kevin Murphy (1988). A Theory of Rational Addiction. Journal of Political Economy, 96, pp. 675-701.

Bell, D. E. (1982). Regret in Decision-making under Uncertainty. Operations Research, 30, pp. 961-981.

Benartzi, Shlomo, and Richard H. Thaler (1995). Myopic Loss Aversion and the Equity Premium Puzzle. Quarterly Journal of Economics, 110 (1), pp.73-92.

Berndt, E., B. Hall, R. Hall, and J. Hausman (1974). Estimation and Inference in Nonlinear Structural Models. Annals of Economic and Social Measurement, pp. 653-665.

Bleichrodt, Han, and Jose Luis Pinto (2000). A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis. Management Science, 46 (11), pp. 1485-1496.

Boehm, T. P. (1981). Tenure Choice and Expected Mobility: A Synthesis. Journal of Urban Economics, 10, pp. 375-389.

Bowman, David, Deborah Minehart, and Matthew Rabin (1999). Loss Aversion in a Consumption-Savings Model. Journal of Economic Behavior and Organization, 38 (2), pp. 155-178.

Brandstatter, E., A. Kuehberger, and F. Schneider (2002). A Cognitive-Emotional Account of the Shape of the Probability Weighting Function. Journal of Behavioral Decision Making, 15 (2), pp. 79-100.

Brendl, C. Miguel, and E. Tory Higgins (1996). Principles of Judging Valence: What Makes Events Positive or Negative. Advances in Experimental Social Psychology, 28, pp. 95-160.

Camerer, Colin, and George Loewenstein (2003). Behavioral Economics: Past, Present, Future, in book Advances in Behavioral Economics, edited by Colin F. Camerer, George Loewenstein, and Matthew Rabin, Princeton University Press.

Carrillo, Juan D. (1999). Self-control, Moderate Consumption, and Craving. CEPR Discussion Paper 2017.

Clark, William, and Allan Heskin (1982). The Impact of Rent Control on Tenure Discounts and Residential Mobility. Land Economics, 58 (1), pp. 109-117.

Daniel, Teresa S. (1994). Time Preference and Saving: An Analysis of Panel Data. VSB-Center Savings Project Progress Report no. 20, Tilburg: Center for Economic Research.

Donkers, Bas (2000). Explaining Time Preference Anomalies: a Quantification, in Ph.D. thesis, Subjective Information in Economic Decision Making, pp.101-129.

Donkers, Bas, and Arthur van Soest (1999). Subjective Measures of Household Preferences and Financial Decisions. Journal of Economic Psychology, 20, pp. 613-642.

Donkers, Bas, Bertrand Melenberg, and Arthur Van Soest (2001). Estimating Risk Attitudes using Lotteries: A Large Sample Approach. Journal of Risk and Uncertainty, 22 (2), pp. 165-195.

Duxbury, Darren, and Barbara Summers (2004). Financial Risk Perception Are Individuals Variance Averse or Loss Averse? Economics Letters, 84, pp. 21–28.

Ekamper, P., and L. van Wissen (2000). Regionale arbeidsmarkten, migratie en woon-werkverkeer. Report for the Dutch Planning Agency (RPD), The Hague, NIDI. ["regional labour markets, migration and commuting."; in Dutch]

Fishburn, Peter C. (1977). Mean-Risk Analysis with Risk Associated with Below-Target Returns. American Economic Review, 67, pp. 116-126.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue (2002). Time Discounting and Time Preference: A Critical Review. Journal of Economic Literature, 40, pp. 351-401.

Fu, Yuming (1991). A Model of Housing Tenure Choice: Comment. American Economic Review, 81 (1), pp. 381-383.

Fuchs, Victor R. (1982). Time Preference and Health: An Exploratory Study. Victor R. Fuchs, ed., Economic Aspects of Health. Chicago: University of Chicago Press.

Giordano, L.A., W. K. Bickel, G. Loewenstein, Eric A. Jacobs, Marsch Lisa, and Gary J. Badger (2002). Opioid Deprivation Affects How Opioid-dependent Outpatients Discount the Value of Delayed Heroin and Money. Psychopharmacology, 163, pp. 174-182.

Gonzalez, Richard, and George Wu (1999). On the Shape of the Probability Weighting Function. Cognitive Psychology, 38 (1), pp. 129-166.

Green, Richard K., and Michelle J. White (1997). Measuring the Benefits of Homeowning: Benefits to Children. Journal of Urban Economics, 41, pp. 441-61.

Gruber, Jonathan, and Botond Koszegi (2001). Is Addiction 'Rational'? Theory and Evidence. Quarterly Journal of Economics, 116, pp. 1261-1305.

Hajivassiliou, Vassilis A., and Paul A. Ruud (1994). Classical Estimation Methods for LDV Models using Simulation, in R. Engle and D. McFadden (eds.), Handbook of Econometrics, vol. IV, North-Holland, New York, pp. 2384-2443.

Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams (2002). Estimating Individual Discount Rates in Denmark: A Field Experiment. American Economic Review, 92, pp. 1606-1617.

Hausman, Jerry A. (1979). Individual Discount Rates and the Purchase and Utilization of Energy-using Durables. Bell Journal of Economics, 10, pp. 33–54.

Heckman, James J. (1981). Chapter 3, Statistical Models for Discrete Panel Data, and Chapter 4, The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process, in C. Manski and D. McFadden, eds, Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, Massachusetts, pp. 114-78 and 179-95.

Heckman, James J., and B. Singer (1986). Econometric Analysis of Longitudinal Data, Chapter 29 in Griliches, Z. and Intriligator, M., eds, Handbook of Econometrics, North-Holland, Amsterdam, pp. 1689 – 1763.

Henderson, J. V., and Y. M. Ioannides (1983). A Model of Housing Tenure Choice. American Economic Review, 73 (1), pp. 98-113.

Houston, Douglas A. (1983). Implicit Discount Rates and the Purchase of Untried, Energy-saving Durable Goods. Journal of Consumer Research, 10, pp. 236-46.

Hyslop, Dean R. (1999). State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women. Econometrica, 67 (6), pp. 1255-1294.

Ioannides, Yannis M. (1987). Residential Mobility and Housing Tenure Choice. Regional science and Urban Economics, 17, pp. 265-287.

Ioannides, Yannis M., and Stuart S. Rosenthal (1994). Estimating the Consumption and Investment Demands for Housing and Their Effect on Housing Tenure Status. Review of Economics and Statistics, 76 (1), pp. 127-141.

Ioannides, Yannis M., and Kamhon Kan (1996). Structural Estimation of Residential Mobility and Housing Tenure Choice. Journal of Regional Science, 36 (3), pp. 335-364.

Kahneman, Daniel, and Amos Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47, pp. 263-292.

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.

Kan, Kamhon (2000). Dynamic Modeling of Housing Tenure Choice. Journal of Urban Economics, 48 (1), pp. 46-69.

Kilka, Michael, and Martin Weber (2001). What Determines the Shape of the Probability Weighting Function Under Uncertainty? Management Science, 47 (12), pp. 1712-1726.

Kirby, Kris N., and Nino N. Marakovic (1996). Delay-discounting Probabilistic Rewards: Rates Decrease as Amounts Increase. Psychonomic Bulletin & Review, 3, pp. 100-104.

Kirby, Kris N., Nancy M. Petry, and Warren K. Bickel (1999). Heroin Addicts Have Higher Discount Rates for Delayed Rewards than Non-drug-using Controls. Journal of Experimental Psychology: General, 128, pp. 78-87.

Kirby, Kris N., Ricardo Godoy, Victoria Reyes-Garc, Elizabeth Byron, Lilian Apaza, William Leonard, Eddy Perez, Vincent Vadez, David Wilkie (2002). Correlates of Delay-discount Rates: Evidence from Tsimane' Amerindians of the Bolivian Rain Forest. Journal of Economic Psychology, 23, pp. 291–316.

Kooreman, Peter (1995). Individual Discounting and the Purchase of Durables with Random Lifetimes. Economics Letters, 48, pp. 29-32.

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

Loewenstein, George (1987). Anticipation and the Valuation of Delayed Consumption. Economic Journal, 97, pp. 666-684.

Loewenstein, George (1988). Frames of Mind in Intertemporal Choice. Management Science, 34, pp. 200-214.

Loewenstein, George, and Drazen Prelec (1991). Negative Time Preference. American Economic Review, 81 (2), pp. 347-352.

Loewenstein, George, and Drazen Prelec (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation. Quarterly Journal of Economics, 107 (2), pp. 573-597.

Loomes, G., and R. Sugden (1982). Testing Different Stochastic Specifications of Risky Choice. Economic Journal, 92, pp. 805-824.

Madden, Gregory J., Nancy M. Petry, Gary J. Badger, and Warren K. Bickel (1997). Impulsive and Self-control Choices in Opioid-dependent Patients and Non-drug-using Control Participants: Drug and Monetary Rewards. Experimental and Clinical Psychopharmacology, 5, pp. 256-262.

Markowitz, Harry (1952). The Utility of Wealth. Journal of Political Economy, 60, pp. 151-158.

McFadden, Daniel, and Paul A. Ruud (1994). Estimation by Simulation. Review of Economics and Statistics, 76 (4), pp. 591-608.

Neilson, William, and Jill Stowe (2002). A Further Examination of Cumulative Prospect Theory Parameterizations. Journal of risk and uncertainty, 24 (1), pp. 31-46.

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

Olsen, Edgar (1990). What is Known about the Effects of Rent Controls? Consulting Report for U.S. Dept. of Housing and Urban Development, Contract Number, No. NA 90-7272.

Pender, John L. (1996). Discount Rates and Credit Markets: Theory and Evidence from Rural India. Journal of Development Economics, 50, pp. 257-296.

Pennings, Joost M. E., and Ale Smidts (2003). The Shape of Utility Functions and Organizational Behavior. Management Science, 49, pp. 1251-1263.

Picone, Gabriel, Frank Sloan, and Donald Jr. Taylor (2004). Effects of Risk and Time Preference and Expected Longevity on Demand for Medical Tests. Journal of Risk and Uncertainty, 28, pp. 39-53.

Prelec, Drazen (1998). The Probability Weighting Function. Econometrica, 66 (3), pp. 497-528.

Priemus, H. (1998). Improving or Endangering Housing Policies? Recent Changes in the Dutch Housing Allowance Scheme. International Journal of Urban and Regional Research, 22 (2), pp. 319-330.

Priemus, H. (2000). Rent Subsidies in the USA and Housing Allowances in The Netherlands: Worlds Apart. International Journal of Urban and Regional Research, 24 (3), pp. 700-712.

Rabin, Matthew (1998). Psychology and Economics. Journal of Economic Literature, 36, pp. 11-46.

Rabin, Matthew (2000). Risk Aversion and Expected-utility Theory: A Calibration Theorem. Econometrica, 68 (5), pp. 1281-1292.

Rabin, Matthew, and Richard H. Thaler (2001). Anomalies: Risk Aversion. Journal of Economic Perspectives, 15 (1), pp. 219-232.

Rae, J. (1834). The Sociological Theory of Capital. London: McMillan, 1905 [1834].

Rizzo, John A., and Richard J. Zeckhauser (2003). Reference Incomes, Loss Aversion, and Physician Behavior. Review of Economics and Statistics, 85 (4), pp. 909-922.

Rossi P. (1955). Why Families Move. MacMillan, New York.

Rossi P. (1980). Why Families Move. Beverly Hills: Sage Publications, 2nd ed.

Samuelson, Paul (1937). A Note on Measurement of Utility. Review of Economic Studies, 4, pp. 155-161.

Schmidt, Ulrich, and Stefan Traub (2002). An Experimental Test of Loss Aversion. Journal of Risk and Uncertainty, 25 (3), pp. 233-249.

Shea, John (1995a). Union Contracts and the Life Cycle/Permanent Income Hypothesis. American Economic Review, 85 (1), pp. 186-200.

Shea, John (1995b). Myopia, Liquidity Constraints, and Aggregate Consumption: A Simple Test. Journal of Money, Credit, and Banking, 27 (3), pp. 798-805.

Shelley, Marjorie K. (1993). Outcome Signs, Question Frames and Discount Rates. Management Science, 39, pp. 806-815.

Thaler, Richard H. (1981). Some Empirical Evidence on Dynamic Inconsistency. Economic Letters, 8, pp. 201-207.

Train, Kenneth (2003). Discrete Choice Methods with Simulation. Cambridge University Press.

Turner, Bengt (1988). Economic and Political Aspects of Negotiated Rents in the Swedish Housing Market. Journal of Real Estate Finance and Economics, 1 (3), pp. 257-276.

Turner, Bengt, and Stephen Malpezzi (2003). A Review of Empirical Evidence on the Costs and Benefits of Rent Control. Working paper.

Tversky, Amos, and Craig R. Fox (1995). Weighting Risk and Uncertainty. Psychological Review, 102 (2), pp. 269–283.

Tversky, Amos, and Daniel Kahneman (1991). Loss Aversion in Riskless Choice: A Reference Dependent Model. Quarterly Journal of Economics, 106, pp. 1039-1061.

Tversky, Amos, and Daniel Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty, 5 (4), pp. 297-323.

Van der Pol, Marjon M., and John A. Cairns (1999). Individual Time Preferences for Own Health: An Application of a Dichotomous Choice Question with Follow up. Applied Economics Letters, 6, pp. 649-654.

van Kempen, R., and Priemus, H. (2002). Revolution in Social Housing in the Netherlands: Possible Effects of New Housing Policies. Urban Studies, 39 (2), pp. 237-254.

Viscusi, W. Kip, and Michael J. Moore (1989). Rates of Time Preference and Valuation of the Duration of Life. Journal of Public Economics, 38, pp. 297-317.

Warner, John T., and Saul Pleeter (2001). The Personal Discount Rate: Evidence from Military Downsizing Programs. American Economic Review, 91, pp. 33-53.

Wu, George, and Richard Gonzalez (1996). Curvature of the Probability Weighting Function. Management Science, 42 (12), pp. 1676-1690.

Samenvatting (Summary in Dutch)

Dit proefschrift is gebaseerd op vier onafhankelijke artikelen. De eerste drie gaan over de interacties tussen preferentiekenmerken, zoals verliesafkeer ('loss aversion'), referentiepunten, verdisconteringsvoeten, risico-aversie, en hoe demografische en socio-economische variabelen deze kenmerken be"invloeden.

Hoofdstuk twee is een artikel over de tijdsvoorkeur van winsten en verliezen. Speciaal ontwikkelde experimenten, maar ook empirische studies hebben laten zien dat mensen geneigd zijn winsten meer te verdisconteren dan verliezen, en dat er verschillen zijn in de wijze waarop uitstel van betalingen in vergelijking met vervroegde betalingen worden verdisconteerd. Zulke verschillen in verdisconteringsvoeten kunnen belangwekkende gevolgen hebben voor de analyse van de uitkomsten van economisch beleid, waardoor een beter begrip ervan belangrijk is. Gebruikmakend van een representatief huishoudpanelbestand worden de ge"impliceerde verdisconteringsvoeten van de volgende vier scenario's geanalyseerd: uitstel van winst, uitstel van verlies, vervroegde winst en vervroegd verlies. Eerst wordt er een samenvatting gegeven van de bestaande literatuur over de relaties die er bestaan tussen verdisconteringsvoeten en andere individuele karakteristieken. Vervolgens worden de verbanden onderzocht tussen de verdisconteringsvoeten en vaak geobserveerde demografische grootheden, zoals geslacht en leeftijd, maar ook subjectieve grootheden, zoals prijsverwachtingen. Veel van deze variabelen hebben een significante invloed op de verdisconteringsvoeten en –belangrijker– deze variabelen be"invloeden de verschillende verdisconteringsvoeten op verschillende wijze. Zulke verschillen bieden de mogelijkheid om scenario-specifieke individuele verdisconteringsvoeten te genereren. We laten niet-geobserveerde heterogeniteit toe. Dit verklaart voor een belangrijk deel de spreiding in de geobserveerde verdisconteringsvoeten. Interessant genoeg blijken de niet-geobserveerde heterogeniteit en de resterende storingstermen onderling gecorreleerd te zijn tussen de vier scenario's. De waargenomen relaties kunnen worden gebruikt voor zowel een beter begrip als het doen van voorspellingen van huishoudgedrag in het kader van het ontwikkelen van economisch beleid.

Om te komen tot een beter begrip van de tijdsvoorkeur van winsten en verliezen in de vier scenario's, wordt er in hoofdstuk drie een structureel model gepresenteerd dat verliesafkeer en referentiepunten in de modellering van de intertemporele keuze meeneemt, gebaseerd op inzichten van Loewensteins (1988) referentiepuntenmodel. Er wordt gebruik gemaakt van data afkomstig van een representatief Nederlands huishoudpanelbestand over de jaren 1997-2002. Deze data bevatten een keur aan informatie over individuele tijdsvoorkeuren en andere individuele karakteristieken. Er wordt een niet-lineair model met stochastische co"effici"enten gebruikt –geschikt voor paneldata– om gelijktijdig de referentiepunten van zowel uitgestelde als vervroegde betalingen te kunnen schatten, alsmede de risico-afkeerco"effici"ent en de verdisconteringsvoet. De uitkomst is dat gemiddeld gezien het referentiepunt van uitgestelde betalingen hoger ligt dan het referentiepunt van vervroegde betalingen, overeenkomstig de hypothese van Loewenstein; het gemiddelde niveau van de verliesafkeerco"effici"ent ligt rond de twee, in lijn met bevindingen van anderen. Dit houdt in dat het negatieve nut van een verlies twee keer zo groot is als het nut geassocieerd met een winstbedrag van dezelfde omvang; vrouwen hebben een grotere verliesafkeer dan mannen, en een hogere opleiding of leeftijd maken mensen verliesafkeriger; hoger opgeleiden en ouderen zijn ook geduldiger.

In hoofdstuk vier wordt de verliesafkeer samen met een kanswegingsfunctie geschat. 'Cumulative prospect' theorie is een standaardbenadering aan het worden voor het modelleren van individuele beslissingen onder risico. Tversky en Kahneman hebben een model geschat met een machtsnutsfunctie met verliesafkeer en tweedelige machtsfuncties als kanswegingsfuncties voor winsten en verliezen. In de onderhavige studie gebruiken we een vergelijkbare structuur om beslissingen van individuen te modelleren. De vijf belangrijke parameters zijn de twee machten van de nutsfuncties voor winsten en verliezen, de verliesafkeerparameter en de twee co"effici"enten van de kanswegingsfuncties voor winsten en verliezen. Om heterogeniteit binnen de populatie te modelleren, worden deze parameters in het empirische model als stochastische co"effici"enten behandeld, afhankelijk van zowel waargenomen demografische factoren als niet waargenomen karakteristieken. De data die we gebruiken zijn afkomstig uit een onderzoek dat representatief is voor de Nederlandse bevolking, met zeven vragen over één of twee weddenschappen. Onze resultaten laten zien dat gemiddeld genomen de machten van de nutsfuncties gelijk zijn aan 0,68 en 0,73 voor respectievelijk winsten en verliezen. Vrouwen hebben een lagere macht voor winsten dan mannen, hetgeen impliceert dat vrouwen risico-afkeriger zijn in geval van winsten. De gemiddelde waarde van de verliesafkeerparameter is 3,1 en de gemiddelde waarden van de kanswegingsfunctieco"effici"enten zijn 1,0 en 0,59 voor respectievelijk winsten en verliezen.

Het laatste artikel, hoofdstuk vijf, heeft een behoorlijk andere inhoud dan de andere drie artikelen, maar maakt gebruik van dezelfde econometrische technieken en modellen. Het onderwerp van het hoofdstuk is de vraag hoe de mobititeit (verhuisbeslissing) en de woonsituatie (beslissing om te huren of te kopen) van huishoudens gezamenlijk te modelleren, gebruikmakend van een multinomiaal probit model dat geschikt is voor paneldata. Hierbij wordt uitgegaan van de veronderstelling dat een verandering in de woonsituatie eigenlijk alleen kan worden waargenomen als een huishouden ook daadwerkelijk verhuist. De modellen worden geschat met de methode van de gesimuleerde maximale aannemelijkheid, waarbij het belang van het goed modelleren van de beginvoorwaarden benadrukt wordt. De schatting is gebaseerd op een niet-gebalanceerd panelbestand afkomstig van het CentER Panelbestand 1994-2003. Er wordt een negatieve toestandsafhankelijkheid in de verhuisbeslissing gevonden. Huizenbezitters verhuizen met kleinere kans naar (opnieuw) een koophuis of naar een huurhuis dan huurders, hetgeen kan worden verklaard uit de veel hogere verhuiskosten voor huizenbezitters.