Disability Insurance

Modeling the transition probabilities

by

C. van der Helm [s338897]

[BSc. Tilburg University 2011]

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Quantitative Finance and Actuarial Science

Tilburg School of Economics and Management
Tilburg University

Supervisors:

dr. R. van den Akker (Tilburg University)

dr. F.C. Drost (Tilburg University)

drs. R.P. de Jonge AAG (PwC)

Date: November 22, 2012

Abstract

This study is about modeling transition probabilities between different states of disability. The transition probabilities are used for estimating the distribution of the benefits to be paid to the policyholders, by the insurance company, for the subsequent twelve months to the measurement date. Since the distribution of the benefits to be paid is necessary to calculate the best estimate of the benefits and the Solvency Capital Requirement of the benefits (Solvency II regulations), the transition probabilities are therefore of importance for the insurance company. In this study different models which could be used to model transition probabilities are investigated. The different models are compared based on the area under the ROC curve, the uncertainty given in the coverage intervals and based on the outcomes of a backtest.

Keywords: Disability insurance, discrete choice models, survival analysis, competing risks analysis, ROC curve.

Table of contents

Acknowledgments

I would like to express my sincere gratitude to all who helped me with this master thesis.

First of all, I would like to thank my colleagues of PwC, who helped me with finding a topic to graduate. In specific, I would like to thank Rian de Jonge and Pieter Bultena. Thank you both for the effort you put in.
This study could not have been done without the help of Theo Beekman. Thank you for giving me the opportunity to do this study and for helping me with all my questions.
Furthermore I would like to thank dr. R. van den Akker and dr. F.C. Drost. Dr. R. van den Akker, thank you for the meetings we had and the support you gave me. You always gave me the feeling that one day I would succeed in finishing this study, and this moment is here. Also a special thanks to dr. F.C. Drost, not only for being my second supervisor and therefore being the chairman of the committee, but also for the flexibility which made it possible for me to finish this study in time.

Last, but definitely most important, I would like to thank my parents for the unconditional support they gave during my time at Tilburg University. They were the ones that supported me throughout all my decisions made and who always believed in me. I will never be able to thank them enough for just being my parents.

Chapter 1

Introduction

In 1901 the first social insurance law, the 'Ongevallenwet', has been introduced in the Netherlands. This law was only applicable to employees with dangerous professions, and the law only granted a benefit in case there had been an industrial accident. Time evolved and so did social insurance.

Currently all companies (in the Netherlands) are obliged to insure their employees against sickness, disability and unemployment. The distinction between an accident or just getting a disease vanished, and so it does not matter what lead to the situation but all that matters is that an employee is not able to perform the job (either fully or partly). Employees pay a premium which is withdrawn from their paycheck, and next to this the employer has to pay some premiums as well. For employees the system covers the risk of reduction in salary. But how is it arranged for self-employed?

Self-employed are own carriers for the risk of reduction in salary due to disability. Due to the abolishment of the "Wet Arbeidsongeschiktheidsverzekeringen" (WAZ) in 2004, the self-employed are no longer covered for this risk and so whenever they are not able to perform -part of- the work anymore they can only rely on the so called "bijstandsuitkering". The "bijstandsuitkering" will, in general, lead to a substantial reduction in income. This is why the private insurance market came up with the disability insurance for self-employed, namely the "arbeidsongeschiktheidsverzekering" (AOV). For self-employed it is their own choice to insure themselves against the risk of (losing income due to) disability.

1.1 Solvency II

Solvency II, the new regulatory framework for insurers, is approaching. There is much debate on when Solvency II is expected to become in force in the Netherlands. 2014, 2015 and 2016 are all referred to by different stakeholders. The new regulatory framework provides in a standard formula for calculating the required capital. In addition to this it offers insurance companies the opportunity to quantify the risks by calculating the required solvency capital by applying an internal model. The standard formula is a generic formula, which, as a result of that, does not necessarily capture the risks run by individual insurance companies well.

The estimated distribution of the benefits to be paid is used for calculating the Technical Provision and the Solvency Capital Requirement. The Technical Provisions of an insurance company consists of the best estimate of the benefits to be paid and a risk margin. The Solvency Capital Requirement is based on the 99.5% quantile over a one-year horizon. This 99.5% quantile comes down to the fact that the event that happens only once each two hundred years needs to be taken into account. The insurance company should have enough capital to cover the benefits to be paid in case this event happens. For an overview of the

Technical Provisions and the Solvency Capital Requirement, Figure 1.1 is included.

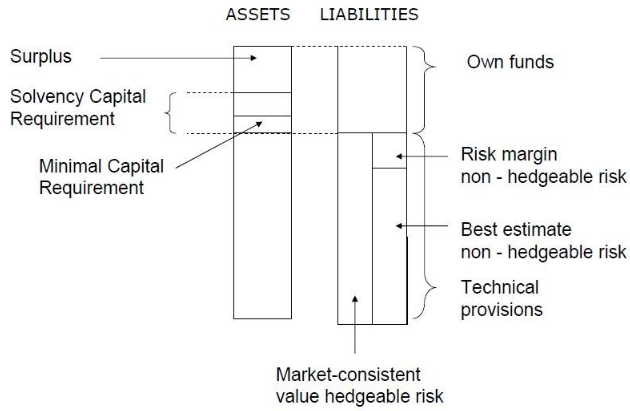

Figure 1.1: Overview of the Technical Provisions and the Solvency Capital Requirement

The Technical Provision and the Solvency Capital Requirement relate to the first pillar of the Solvency II regulation. The framework of Solvency II consists of the following three pillars:

- Pillar 1, Quantitative requirements;
- Pillar 2, Qualitative requirements;
- Pillar 3, Disclosure requirements.

The second pillar and the third pillar are not discussed here. The pillars are focusing on risk management and reporting, which is not of interest in this study (though it is an important part of Solvency II and though it is important for an insurance company).

In order to estimate the Technical Provision and the Solvency Capital Requirement, the distribution of the benefits to be paid need to be estimated. In turn, the transition probabilities between the different states of disability are necessary in order to estimate the distribution of the benefits to be paid. In the next section these transition probabilities are introduced.

1.2 Transition probabilities

Disability insurance products, offered by insurance companies, are complicated products. The products are complicated because they involve multiple uncertainties:
- What is the probability that a policyholder becomes disabled?
- If a policyholder becomes disabled, what disability percentage (percentage of work a policyholder cannot perform anymore) will the policyholder have?
- How long will the policyholder stay disabled? And as long as a policyholder is disabled, how does the disability percentage of this policyholder vary across time?

In this study, the focus is on transition probabilities. The policyholders are divided into different states (regarding their disability percentages) and the probabilities to go from one state to one another, the transition probabilities, are modeled. An example of a disability state is the active state, which indicates that a policyholder has a disability percentage of less than 25%. Based on the disability percentages at a certain time, policyholders are allocated to the different disability states at the various times. With the transition probabilities the issue of how long the policyholder will be disabled and how the disability percentage of this policyholder vary across time can be investigated. To be more precise, with the transition probabilities it is possible to generate multiple paths of how state a policyholder is in evolves over time.

Policyholders are allocated to different states: a policyholder is active, partially disabled or fully disabled. A policyholder is called active in case of a disability percentage of less than 25%, partially disabled in case of a disability percentage between 25% and 50% and fully disabled in case of a disability percentage of more than 50%. Since there are three different states a policyholder could be in, there are nine different transition probabilities which are of importance in this study. These transition probabilities are summarized by the following equation

$$p_t^{ij} = \Pr(\text{being in state i at time t, being in state j at time t+1}) \quad i, j \in \{0, 1, 2\},$$

with state 0 in case of the active state, state 1 in case of the partially disabled state and state 2 in case of the fully disabled state.

1.3 Problem definition

For an insurance company it is important to know the Technical Provision and the Solvency Capital Requirement of the benefits to be paid. For this, the distribution of the benefits to be paid need to be estimated. In order to do this, the transition probabilities between the different states of disability need to be modeled. Modeling transition probabilities in this study is done based on a dataset received from a large Dutch insurance company. To determine how to estimate these transition probabilities best, the probabilities are modeled using different models. Currently, the discrete choice models (dynamic binary and dynamic multinomial logit models) and survival analysis models (Cox proportional hazard model) are frequently used models to model transition probabilities, and therefore these models are investigated in this study. The focus of this study on the transition probabilities is also reflected in the title, ***Disability Insurance*** *Modeling the transition probabilities*. In this study the different models which are used to model the transition probabilities are compared and a preferred model is tried to find.

The transition probabilities are modeled based on data. Data is given up to a certain measurement date. For the subsequent twelve months to the measurement date, paths are generated which represent the disability state a policyholder will be in at each of the twelve months. The time horizon of twelve months is chosen due to the Solvency Capital Requirement which is based on a one-year horizon. Based on the generated paths, an estimated distribution of the benefits to be paid can be given.

1.4 Outline

In this study transition probabilities are modeled by different models. Based on the transition probabilities, the distribution of the benefits to be paid to the policyholders within the subsequent twelve months to the measurement date can be estimated. However, before being able to model the transition probabilities used to estimate the distribution of the benefits, the disability insurance need to be further explained. How did disability insurance evolve over time and which aspects are of importance in this study. Disability insurance is explained in Chapter 2.

In order to estimate transition probabilities, data are necessary. Data are provided by a large Dutch insurance company and are discussed in Chapter 3. In Chapter 3 modifications to the dataset are discussed as well as some summary statistics of the dataset.

In Chapter 4 and Chapter 5, the models used to estimate the distribution of the benefits to be paid to the policyholders within the subsequent twelve months to the measurement date are presented. The results of the distribution, as well as some sensitivity analysis and a test of backtesting, are discussed in Chapter 6.

In Chapter 7 the conclusions concerning the quality of the various models are drawn and compared to each other. Furthermore some recommendations for further research are given.

Chapter 2

Disability models

As mentioned in Chapter 1, the insurance for disability exists for a long time already. The time passing from the beginning of this insurance made the circumstances evolve and that is why the federation of Dutch insurers (the Actuarial Committee) set up a committee to advice on the disability insurance. This committee, the "Kontaktcommissie arbeidsongeschiktheid-, ziekengeld- en ongevallenverzekering" (KAZO), became known all over the world due to their recommendations in 1991, which are used for a long time. Gregorius (1993) summarizes their main results; Section 2.1 gives a brief review. In Section 2.2, the focus is on the model as used in this study.

2.1 KAZO-model

2.1.1 Types of disability insurance

The KAZO Committee distinguishes two types of individual disability insurance, namely the A-cover and the B-cover. The A-cover is the first year's risk which starts paying out a benefit after a waiting period (a period in which income reduction is the policyholder's own risk, for most of the policyholders this waiting period is 30 days). The B-cover is the after-first year's risk which starts paying out a benefit after one year. This benefit is an annuity, and it will be paid out until recovery, until death or until the end age of the contract is reached (at latest the first day of retirement).

In the years the KAZO committee investigated disability insurance, the so called "Algemene Arbeidsongeschiktheids Wet" (AAW) was in effect and this is why the Actuarial Committee was splitting the disability insurance into two covers. The AAW was a law which provided all self-employed an annuity after a waiting period of one year. The split between the A and B-cover made it possible to choose different insured amounts. Another possible difference between the two covers as discussed in Gregorius (1993) is that in case of the A-cover, it is examined whether (and if yes, in what proportion) the policyholders can perform their own tasks while in the B-cover it is checked whether (and if yes, in what proportion) the policyholders can perform other tasks which provide them with some income. Note that in none of the covers it is checked whether a policyholder can perform a job in general.

2.1.2 Mortality rates

As explained in Section 2.1.4, Gregorius (1993) assumes that a policyholder is either active or disabled. Furthermore, it is anticipated that policyholders can die. In case it is anticipated that policyholders can die, an assumption has to be made on mortality rates. Gregorius (1993) mentioned that there is a good reason to believe that the mortality rates of disabled and actives are different. This since a substantial part of the disabled have a decease, and sick people have higher mortality rates compared to healthy people. However, for both groups they use the mortality rates of the overall population.

2.1.3 Model of the A-cover

All mentioned in Gregorius (1993) about the A-cover is that it is very similar to other types of non-life insurance. Disability and recovery probabilities are calculated based on observations and the estimates benefits to be paid is being calculated based on these probabilities. The A-cover is no annuity and the A-cover only has a one-year horizon, which makes the estimation of the distribution of the benefits to be paid relatively easy.

2.1.4 Model of the B-cover

In Gregorius (1993) the model distinguishes between active policyholders and disabled policyholders. To define the various states, the disability percentage is not taken into account. Only the duration of the disability is translated into different states. Furthermore the mortality probabilities are taken into account and thus a state is defined for death. So, a policyholder could be either in the active state (A), in the death state (D) or the policyholder is disabled for x years (I(x)).

The bases used for estimating the distribution of the benefits to be paid, which is done based on actuarial methods, are the (conditional) disability rate, the (conditional) recovery rate (which depends on the lapsed time of the disability) and the mortality rate. For more information and a more detailed explanation, we would like to refer to Gregorius (1993).

The model as described by Gregorius (1993) assumes that whenever a policyholder gets disabled and stays disabled for six years, the probability to recover is nil. This is why the recovery probabilities (going from state disabled to state active) are set equal to zero for policyholders disabled for at least six years.

A graphical overview of the model used in Gregorius (1993) can be found in Figure 2.1. Note that A stands for active policyholder, D stands for dead policyholder and I(x) stands for being disabled for x years (and in case x=6, being disabled for at least x=6 years).

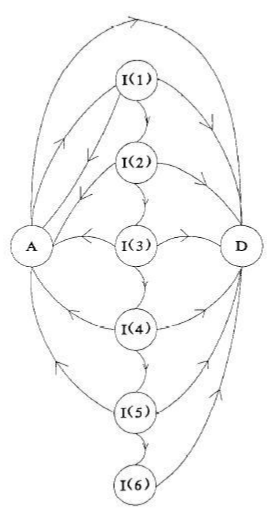

Figure 2.1: Graphical overview of the model in Gregorius (1993)

2.1.5 Recovery probabilities

The way the recovery probabilities are modeled is as explained before. In the B-cover the calculation is based on actuarial principles, where $r_i(x)$ stands for the recovery probability of a policyholder with age x at the start of the disability and with lapse time i of the disability. $r_i(x)$ is given by the equation $r_i(x) = a_i - b_i \cdot x$, and a the values for a_i and b_i, for $i \in \{1, 2, 3, 4, 5\}$ are given in Table 2.1. Note that $x \in \{22, 27, 32, \cdots, 57, 62\}$ (9 groups).

i	a_i	b_i
1	1.24111	0.02219
2	0.66499	0.01153
3	0.27394	0.00532
4	0.23547	0.00470
5	0.14166	0.00319

Table 2.1: The values for a_i and b_i, $i \in \{1, 2, 3, 4, 5\}$, in order to calculated the recovery probability of a policyholder with age x and with lapse time i of the disability

Gregorius (1993) does not mention anything about the negative recovery probabilities which appear at higher ages at the start of the disability (for example, $1.24111 - 0.02219 \cdot 57 = -0.02372$). The assumption has been made that the recovery probabilities are taken as the maximum of the recovery probability as defined above and 0. The recovery probability of a policyholders with age x at the start of the disability and with lapse time 6 years or longer, $r_6(x)$, is equal to zero.

The model as summarized in Gregorius (1993) and the recovery probabilities as listed above, are not completely representative for disability insurance at this moment of time. Next to the fact that the recovery probabilities given in Section 2.1.5 cannot be used due to different model definitions, it is plausible that the recovery probabilities as given are not representative anymore. In the late 90's of the last century the number of disabled persons (employees) in The Netherlands reached the million. This led to the idea that the insurance laws had to be updated such that they gave disabled persons an incentive to recover. In 2004, also the regulations regarding the disability of self-employed changed.

2.2 Current models

The model as explained in Section 2.1 is not the model used in this study. The main difference is that in this study, the focus is on models that include the disability percentage in the definition of the states. A policyholder could be active, partially disabled or fully disabled. However, only in case the policyholder survives to the next period (mortality rates are taken into account and so again a state is defined for death). A state is added to the alternatives "active" and "disabled". Adding a state to the existing states leads to a model which models the transition probabilities more precisely. Adding a disability state implies that, after the initial incidence not just recovery probabilities (the probability to -partially- recover) are of interest, but also partial incidence probabilities (the probability of a partially increasing disability percentage) are important. This is why transition probabilities in general are discussed in the remainder of this study, instead of recovery probabilities only. Due to the extra state a policyholder could be in, more transitions probabilities are created. This is explained later in Section 2.2.4.

2.2.1 Types of disability insurance

In order to perform this study, a dataset is received from a large Dutch insurance company. This dataset, which is further explained in Chapter 3, gives the disability percentage of the policyholders from the start of their disability. This implies that the dataset does not distinguish between the A and B cover. In the dataset received, an insurance product which pays out from the first payment till the end date (which lies between the age of 35 and 65 for the policyholders) is given.

2.2.2 Mortality rates

In this study mortality rates for the policyholders conform the "Prognosetafel AG2012-2062" [1] are assumed. Besides age, a distinction has been made in mortality rates regarding male and female policyholders. Furthermore a "prognosetafel" is assumed which implies that mortality rates may differ across years. The mortality rates which are applied in this study are included in Appendix F.

2.2.3 Model

In this study the focus is on a disability model with three different disability states: a policyholder can be active, partially disabled or fully disabled. In case a policyholder is active the disability percentage is less than 25%, in case a policyholder is partially disabled the disability percentage is between 25% and 50% and in case a policyholder is fully disabled the disability percentage is 50% or more. Again, as explained in the previous section, mortality rates are taken into account. Next to this, the assumption that any transition probability is equal to zero in case of being disabled for at least six year is rejected. A graphical overview of the model used in this study can be found in Figure 2.2. Note that A stands for active policyholder, D stands for dead policyholder, P(x) stand for being partly disabled for x years and F(x) stand for being fully disabled for x years.

As one can see when comparing both Figure 2.1 and Figure 2.2, the model used in this study is more complicated than the model in Gregorius (1993).

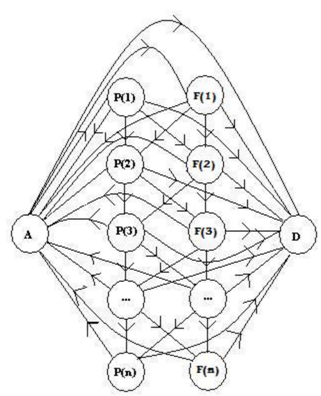

Figure 2.2: Graphical overview of the model used in this study

2.2.4 Transition probabilities

In contrast to Section 2.1.5, in this study not only recovery probabilities are discussed but incidence probabilities are discussed as well. Incidence probabilities can be seen as the probability to deteriorate, and so it could be either the probability of becoming (fully / partially) disabled as well as the probability of becoming fully disabled after being partially disabled. Both the recovery probabilities and the incidence probabilities

[1] http://www.ag-ai.nl/view.php?action=view&Pagina.Id=478

are called the transition probabilities. Whenever talking about the transition probabilities, either the recovery probabilities or the incidence probabilities are meant. The states (active, partially disabled or fully disabled) and the transition probabilities corresponding to these states are given in Figure 2.3. Note that the state of being active is given by state 0, the state of being partially disabled is given by state 1 and the state of being fully disabled is given by state 2.

As can be seen in Figure 2.3, there are four solid lines and two dashed lines. The two dashed lines are the transition probabilities which cannot be modeled on the basis of the dataset received. These transition probabilities are the probabilities of becoming disabled (either partially disabled or fully disabled) when being active. Since the dataset provided only contains policyholders which are already disabled, the transition probability of the active state to the partially disabled state and the transition probability of the active state to the fully disabled state cannot be observed.

The transition probabilities are the main focus in this study. Different models are used to model the transition probabilities and with these modeled transition probabilities it is possible to estimate the distribution of the benefits to be paid. Given the estimated distribution of the benefits to be paid, it is possible to derive the expected value of the benefits and the 99.5% quantile of the benefits.

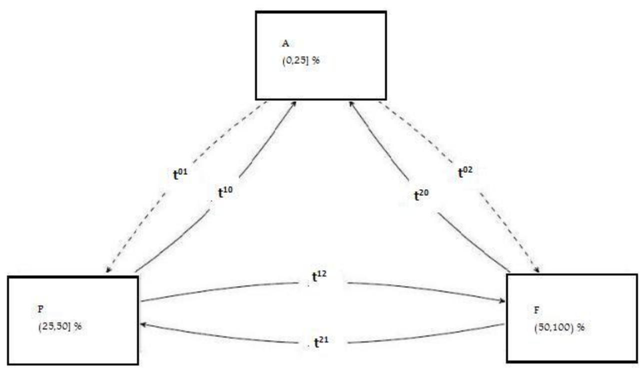

Figure 2.3: The transitions investigated in this study (solid lines)

2.2.5 Average disability percentages

The focus in this study is on transition probabilities which are used to estimate the distribution of the benefits to be paid to the policyholders within the next twelve months. In case a policyholder is in the active state, no benefit is paid out. In case of both the partially disabled and fully disabled state, a benefit is paid out. Normally, a benefit is paid out which equals the actual disability percentage of the policyholder times the insured amount of the policyholder. However, it is assumed that in case a policyholder is in the state of being partially disabled, the policyholder will have the average disability percentage of being in the state of partially disabled. The same holds for policyholders which are in the state of being fully disabled.

The average disability percentage in the partially disabled state is 33.86% and the average disability percentage in the fully disabled state is 77.49%. More specific it is assumed that a policyholder who is in the state of being partially disabled receives a benefit of 33.86% times the insured amount, while a policyholder who is in the state of being fully disabled receives a benefit of 77.49% times the insured amount. So all that matters is the state the policyholder will evolve in, not the disability percentage the policyholder will have in that specific state.

In Section 3.3.8, some more information regarding the disability percentages are discussed.

2.3 Markov process

Most models on disability insurance are built on the idea of the Markov chain. Using the definition of Yin and Zhang (2005)

> *"Suppose that α_k is a stochastic process taking values in M, which is at most countable (i.e., it is either finite $M = 1, 2, ..., m_i$ or countable $M = 1, 2, ...$). We say that α_k is a Markov chain if*
>
> $$P^{ij}_{k,k+1} = P(\alpha_{k+1} = j|\alpha_k = i) = P(\alpha_{k+1} = j|\alpha_0 = i_0, ..., \alpha_{k-1} = i_{k-1}, \alpha_k = i).''$$

As one can imagine, the Markov property is a quit stringent assumption since it assumes that the future state of the policyholder does not depend on the past states of the policyholder, only on the current state. This assumption can be tested by including the previous state as an explanatory variable. In case the explanatory variable turns out to be significant, it indicates that the assumption is not applicable (on the data).

As the results show, in both Chapter 4 and Chapter 5, the Markov property does not seem to hold in this study. In both Appendix A, Appendix B and Appendix C, respectively Table A.1, Table B.1 and Table F.1, it can be seen that the explanatory variable previous state (x_8 respectively x_7) is significant for almost all transitions for both the logit models as for the Cox PH model. This implies that a Markov chain model does not fit the underlying process well: the past does play a role in the future transition probabilities.

2.4 Summary

In this chapter it is shortly discussed how the disability models worked in the nineties of the last century and the changes that have been made since then. Not only the world a policyholder lives in changed, also the law changed (which caused that it is more attractive to recover) and the insurance contracts offered changed. Now that the basics of the disability models are known, the data used for this study is discussed in the next chapter. Once introduced to the data, the focus is on the models to approach the transition probabilities.

Chapter 3

Data analysis

The data used for this study are provided by a large Dutch insurance company. A dataset has been provided with almost 56,000 policyholders (self-employed, because the dataset is a dataset of policyholders who have an "arbeidsongeschiktheidsverzekering") who reported a disability, with the claim date of the disability (start date) ranging from January 2000 up to and including June 2012. When a policyholder reports a disability, this is also referred to as a claim. Therefore policyholders who have reported a disability (all policyholders in the dataset) are sometimes referred to as claimants.

In Chapter 2 an overview is provided of the basis of disability insurance. The chapter furthermore elaborates with some important changes that have been made since the publication of Gregorius (1993). With this information it is possible to investigate the models which are used to approach the transition probabilities. Before these models (discrete choice models and survival analysis models) are discussed in Chapter 4 and Chapter 5, the dataset used for modeling transition probabilities is discussed in this chapter. In Section 3.1 the adjustments are discussed that had to be made in order to have a dataset which could be used in modeling transition probabilities. Both adjustments that lead to a reduction of the dataset as well as adjustments that had to be made manually are discussed. Next to these adjustments some remarks are mentioned regarding the dataset. In Section 3.2 the covariates which are used in this study are listed and the expected impact on the transition probabilities is discussed. Summary statistics of the most important characteristics of the policyholders are given in Section 3.3.

3.1 Adjustments of the dataset

The original dataset of 55,958 policyholders is reduced to a dataset of only 29,756 policyholders. This means that the dataset is reduced by almost 47%. In Section 3.1.1 the adjustments which led to this reduction are described and motivated. Furthermore some manual adjustments had to be made. These manual adjustments, plus some remarks on the data, are explained in Section 3.1.2.

3.1.1 Reduction of the dataset

The reduction of the dataset by almost 47% is partly caused by policyholders who reported a disability but never received a benefit from the insurance company. For all policyholders a disability percentage is given for each month the claim has not finalized (yet), with the disability percentage being the percentage of disability of the policyholder and thus the percentage of work a policyholder cannot perform anymore. It is assumed that a benefit is only paid out in case a policyholder has a disability percentage of 25% or more. The interest of this study lies in the probabilities of the transitions t^{10}, t^{12}, t^{20} and t^{21}, which indirectly states that the policyholders who are either in the state of being partially disabled or in the state of being fully disabled need to be taken into account. Policyholders who never received a benefit, never had a disability percentage of 25% or more and thus never were in the state of being partially disabled or in the state of being fully disabled. These policyholders are of no effect in this study and therefore can be removed from the dataset without any further implications. A part of these removed policyholders are pregnant women who did not experience any complications during their pregnancy. These women receive a fixed benefit from the insurance company (comparable to maternity leave in case of being employed) and their disability percentage is set to 0%.

Another inconsistency which reduced the original dataset is given by policyholders with multiple claims. Policyholders with multiple claims have a disability percentage of 0% in a specific month, with this month being between months with a disability percentage higher than 0%. The reason for this can be diverse and, as explained by the insurance company, does not mean that the policyholder really turned into the active state. It could, for example, be the case that the insurance company still needs to receive some documents from the policyholder and that it forces the policyholder to hand in these documents by setting the benefit at zero until it receives the documents. These policyholders may incorrectly influence the estimation of the transition probabilities. That is why it is decided to remove the policyholders with multiple claims from the dataset.

The dataset of 30,702 remaining policyholders, after the two most important reductions explained above, could be used in the remainder of this study. The dataset is reduced by 42.9% (23,981 policyholders) due to removing the policyholders who never had a disability percentage of 25% or more, and the dataset is reduced by another 2.3% (1,275 policyholders) due to removing the policyholders with multiple claims. However, since the covariates duration, sex, age, insured amount, class of profession and previous state are used later in this study, the policyholders for which there is missing information for either one or more of these variables are removed from the dataset as well. The reduction due to missing information based on age, sex or insured amount led to an extra reduction of 946 (= 1.7%) policyholders. The removal of these policyholders leads to a dataset of 29,756 policyholders.

3.1.2 Manual adjustments and remarks

Next to the adjustments which led to a reduction in the dataset, a manual adjustment had to be made as well. This adjustment is the result of policyholders who were born in a leap year. Due to a shortcoming of Microsoft Excel, a problem exists with part of the policyholders who were born on February 29^{th}. Microsoft Excel works well with the date of February 29^{th} in case of a leap year, but Microsoft Excel malfunctions in case February the 29^{th} does not exist in a specific year. The problem existed with the calculation of the maximum amount of months benefits of the policyholders (calculated as the date or birth plus the end age, minus the age at the start of the disability). This is why the date of birth, of those policyholders who were born on February 29^{th} and who have an end age unable to divide by four (in order to get an integer), are changed to the 28^{th} of February.

Other important remarks that need to be made regarding the dataset are that it involves right censoring and missing data. The dataset received consists of policyholders who reported a disability between January 2000 up to and including June 2012. Some of these policyholders are still disabled at the measurement date of June 30th 2012. When analyzing the data, it is given that 18.49% of the dataset (5,501 policyholders) is still disabled at the measurement date of 30 June 2012 (and so it is unknown how these policyholders will evolve over time). The data of these policyholders are right censored. Regarding the missing data, from the 29,756 claims, there are 9,257 claims not reported during (at least) the first month of benefit payment. This is more than 31% of the dataset.

In this study the so called IBNR claims are not taken into account. These IBNR claims are Incurred But Not Reported, and are those policyholders who turned disabled before the measurement date, but who did not report their disability at that measurement date yet.

The dataset only consists of policyholders who actually reported a disability. So no policyholders who have an insurance but never were disabled are present in the dataset. This means that, as is explained in Section 2.2.4, it is not possible to identify the probabilities of transitions t^{01} and t^{02} based on the dataset. In case it is of interest to model the probabilities of transitions t^{01} and t^{02}, a dataset need to be provided with all policyholders.

3.2 Covariates

In this study the following explanatory variables regarding the transition probabilities are taken into account:

- duration of the disability so far. This variable differs over time;
- sex of the policyholder. Sex is a dummy variable (either value one or value zero). Female is taken as base level;
- age of the policyholder at the start of the disability;
- insured amount of the policyholder;
- class of profession of the policyholder. Class of profession could be one, two, three or four and is a dummy variable. Class of profession one is taken as base level (more information about the class of profession and the difference between the classes of profession is given below);
- previous state of the policyholder. This variable differs over time.

It is investigated whether the explanatory variables are significant. If not, they are left out of the model.

The different covariates may have a different impact on the transition probabilities, and even the same covariate may have a different impact on the different transition probabilities. The expected impact the covariates have is shortly discussed below. Note that this is just the expectation and the sign of the estimated coefficients by the different models eventually state the real impact of the covariates in the different models for each transition probability.

duration

It is expected that the covariate duration has a negative impact on all transition probabilities. This is clear since the longer a policyholder is disabled, the less likely that the disability percentage of the policyholder will change. Even if the disability percentage changes, most likely it will only be a small change which could occur within a disability state.

sex

In practice the probability to become disabled is highly dependent upon the sex of a policyholder, however the transition probabilities once disabled are less dependent upon this covariate. Therefore the expected sign of the covariate is conditional on being significant.

As seen in Van Waarden (2012), is expected that the coefficient of the covariate sex is positive for transition probability p_{12} and negative for transition probability p_{20}. For the other two transition probabilities, since they both involve recovery, it is expected that the coefficient of the covariate sex is negative as well.

age

For age, a negative coefficient for the transitions which involve recovery (transition probabilities p^{10}, p^{20} and p^{21}) is expected. The younger a policyholder, the faster the recovery of the disability. For transition probability p^{12} the opposite sign is expected.

insured amount

For the insured amount it is not known whether policyholders who have a higher insured amount, and so who probably have a higher expected income, are less likely to stay disabled or are more likely to stay disabled.

The reasoning could be both ways and so the impact of the insured amount should become clear from the estimated coefficients by the different models

class of profession

Just as mentioned in case of the covariate sex, the covariate class of profession seems to have a substantial influence on the probability to become disabled in practice, but the influence of the transition probabilities once disabled turns out to be insignificant. The expectation as discussed is therefore only in case the covariate turns out to be significant.

Since class of profession three has the average probability to become disabled, whereas class of profession one and class of profession two have a lower probability and class of profession four has a higher probability, the expected sign for the estimated coefficients regarding class of profession two, three and four for transition probability p^{12} is positive and for transition probabilities p^{10}, p^{20} and p^{21} is negative.

previous state

It is important to note that the previous state of the policyholder is not taken as a dummy variable. In case a policyholder is in the state of being partially disabled, the previous state of the policyholder could be either the active state (state 0) or the fully disabled state (state 2). In case a policyholder is in the state of being fully disabled, the previous state of the policyholder could be either the active state (state 0) or the partially disabled state (state 1). The previous state of the policyholder must be viewed from a different perspective. The specific state could be of importance, however it probably is of more importance whether the disability percentage of the policyholder deteriorates of improves. This is why for the transition probability p^{10} it is expected that the sign of the estimated coefficient is positive. The transition probability p^{20} probably will have a negative sign, this since the disability percentage of a policyholder who was in the state of being partially disabled before turning fully disabled deteriorates. For the other transition probabilities it is not clear what the expected sign of the estimated coefficients will be.

In the dataset more information is given about the policyholders. For instance, specific information regarding job, illness / disease and insurance contract. Not all the information is included in the model, since either an overall (less detailed) explanatory variable is included (for example the class of profession is included and so the target group, group of profession and code of profession are excluded) or due to some assumptions that have been made in this study (for example it is assumed that the interest rate is equal to 0 and so there is no such thing as discounting or indexation).

3.3 Summary statistics

The dataset contains extensive details on the policyholders like contract details, personal characteristics and claim characteristics. The dataset has been analyzed and in the following sections the most important characteristics of the policyholders are summarized and discussed.

3.3.1 Sex of the policyholders

27,040 policyholders out of the 29,756 policyholders (90.87%) in the dataset are male. No conclusions can be drawn about disability probabilities for male and female based on this percentage since no information is available on the proportion of male to female policyholders in the complete dataset.

3.3.2 Class of profession of the policyholders

The distinction of the classes of profession is regarding the probability to become disabled. Class of profession three is taken those professions which have an average probability to become disabled. Classes of profession one and two both are less heavy, and so the probability to become disabled in those two classes of profession should be lower. Class of profession four is the class of profession which is most heavy. By far, most policyholders have class of profession four.

class of profession	
one	1.14
two	7.42
three	10.28
four	81.16

Table 3.1: Class of profession of the policyholders (in percentages)

3.3.3 End age of the policyholders

The end age the policyholders have chosen lies between the age of 35 and 65 years old. From all policyholders, 74.32% has an end age of sixty years. Most policyholders use the option to insure themselves for disability until pension, however there is also the option to choose to insure for a selected number of years. This is the reason why there are policyholders with an end age below fifty years. In Table 3.2 the empirical distribution is given for different (groups of) end ages.

end age of policyholder	
unknown	0.03
< 55	0.71
55	5.06
56 - 59	0.46
60	74.32
61 - 64	1.78
65	17.63

Table 3.2: Empirical distribution of the end age of the policyholders (in percentages)

3.3.4 Cohort of the policyholders

In Table 3.3 information of the cohorts of the policyholders at the start of their disability is given.

The 4.54% of policyholders which turned disabled in 2012 all became disabled in the first six month of the year 2012. Off course, since the measurement date of the dataset is the 30th June 2012, it is assumed that in the subsequent six months more policyholders will become disabled. The percentage which is low compared to the years 2003 up to and including 2011 is therefore understandable. In the years 2000, 2001 & 2002 the percentages are less understandable, and after a check with the insurance company these low percentages can be explained. The reason for this is that the system as used by the insurance company is in use from the beginning of 2003, and the policyholders who became disabled in 2000, 2001 or 2002 and who were still disabled in 2003 were included in the system. However, these policyholders are those policyholders who are already disabled for one, two or three and so are only a small percentage of all the policyholders who turned disabled on those three year.

cohort of start disability	
2000	1.22
2001	1.58
2002	3.05
2003	9.96
2004	9.72
2005	9.71
2006	9.34
2007	9.35
2008	9.46
2009	10.62
2010	11.16
2011	10.28
2012	4.54

Table 3.3: Empirical distribution of the cohorts of the policyholders at the start of the disability (in percentages)

3.3.5 Transitions of the policyholders

In total there are 36,957 transitions. These transitions occurred during the entire duration ($t = 1, \cdots, 132$) of the dataset. In Table 3.4 the empirical distribution of the different transitions is presented. Some policyholders transferred multiple times between states, while other policyholders did not transfer at all.

	active	**partially disabled**	**fully disabled**
partially disabled	21.40	-	8.72
fully disabled	44.20	25.67	-

Table 3.4: Empirical distribution of the transitions of the policyholders (in percentages)

3.3.6 Age of the policyholders at the start of the disability

In Table 3.5 some information can be found regarding the age of the policyholders at the start of the disability. Figure 3.1 shows a histogram of this variable. Note that it is quite remarkable that there is a policyholder who has an age of more than 65 at the start of the disability, while the end age of this policyholder is less or equal than 65. This implies that at the time the benefit payments to this policyholder started, the policyholder was not insured anymore.

age	
average	43.33
minimum	18.57
maximum	65.15
median	43.28
standard deviation	8.85

Table 3.5: Summary statistics of the age of the policyholders at the start of the disability

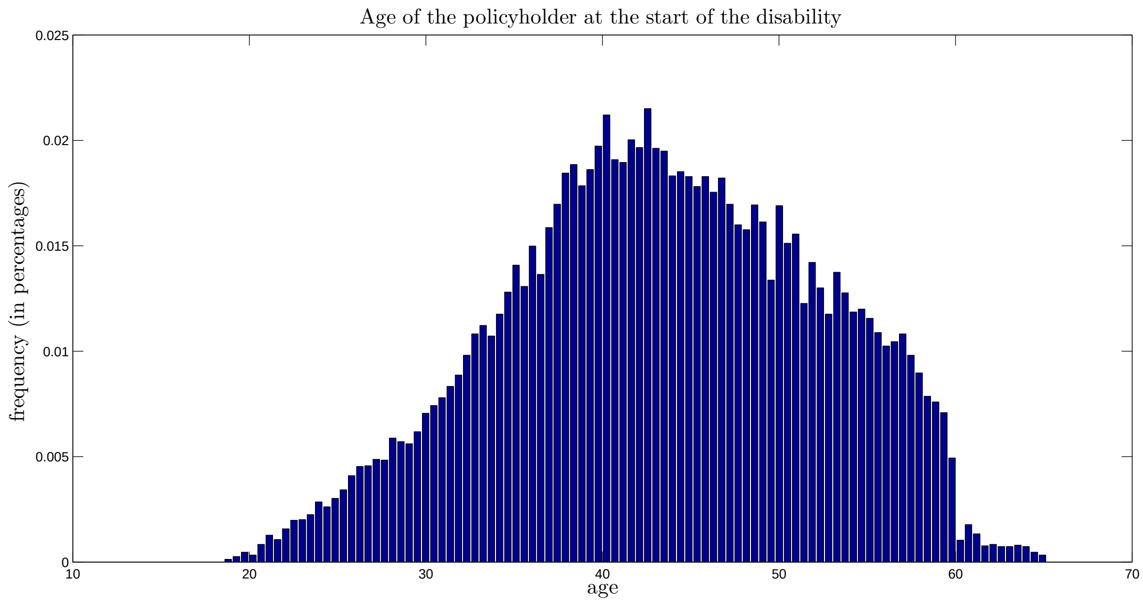

Figure 3.1: Histogram of the age of the policyholders at the start of the disability

3.3.7 Insured amount of the policyholders

The insured amount of a policyholder depends on the income of the policyholder. It is expected that the higher the income of the policyholder, the higher the standard of living. A higher insured amount results in a higher premium that needs to be paid to the insurance company. The insured amount may be chosen up to 80% of the income of the self-employed, with €250,000 as the maximum. Information regarding the variable insured amount is given in Table 3.6. Figure 3.2 shows a histogram of the insured amount of the policyholders.

insured amount	
average	22,821
minimum	463
maximum	139,365
median	22,323
standard deviation	10,135

Table 3.6: Summary statistics of the insured amount of the policyholders (in Euros)

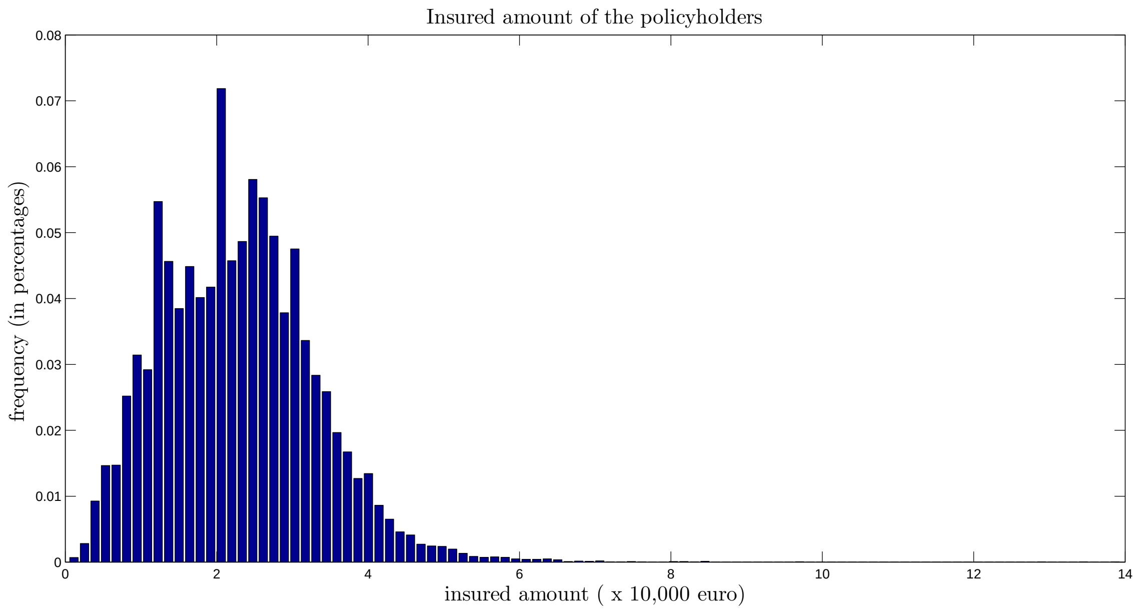

Figure 3.2: Histogram of the insured amounts of the policyholders

3.3.8 Disability percentage of the policyholders

The average disability percentage of being in the partially disabled state is 33.86% and of being in the fully disabled state 77.49%. This average is calculated over all policyholders and all durations given. A histogram of the disability percentages in case of being in the partially disabled state is given in the left graph of Figure 3.3, and the histogram of the disability percentages in case of being in the fully disabled state is given in the right graph of Figure 3.3.

Investigating the average disability percentages in more detail, it is possible to calculate the average disability percentage per cohort. In Table 3.7 the mean and the standard deviation of the average disability percentages of the partially disabled state per cohort are given, just as the mean and the standard deviation of the average disability percentages for the fully disabled state for all cohorts. In Figure 3.4 the average disability percentage for both the partially disabled and the fully disabled state are given for the different cohorts. As can be seen in Figure 3.4, the average disability percentage for the partially disabled state is stable, while the average disability percentage for the fully disabled state is increasing over the cohorts.

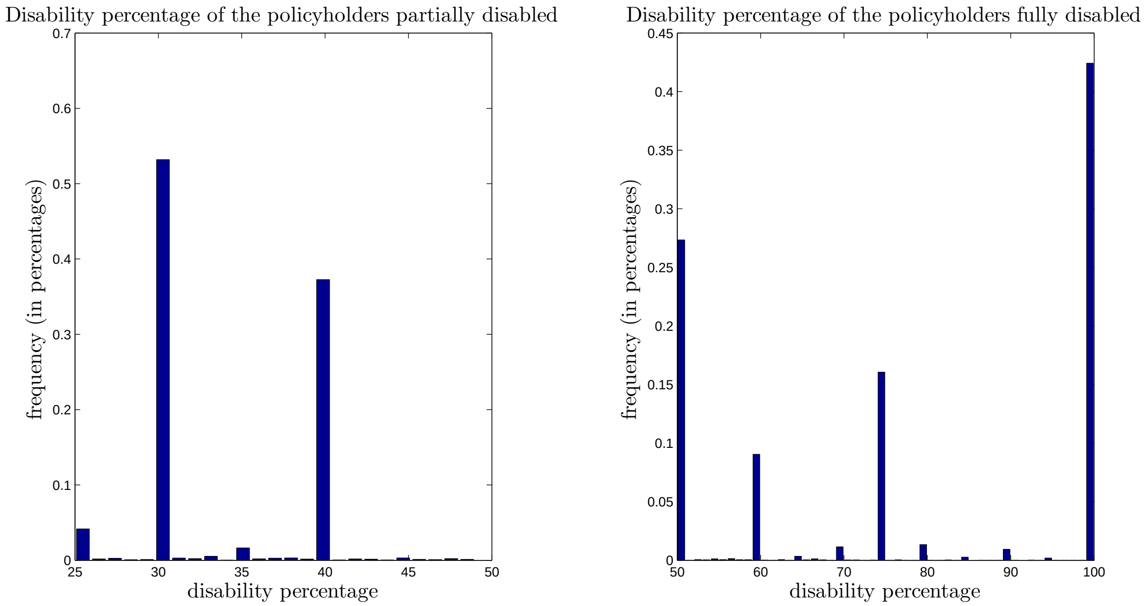

Figure 3.3: Histograms of the disability percentages in the partially disabled state (left) and in the fully disabled state (right)

	partially disabled		**fully disabled**	
cohort	**mean**	**standard variance**	**mean**	**standard variance**
2000	35.2688	5.1540	72.8639	20.6204
2001	34.2259	4.9506	73.9100	21.1771
2002	33.9813	4.8965	74.3947	20.6331
2003	33.1408	5.0486	76.1699	21.6021
2004	34.4764	5.2040	76.2623	21.4366
2005	34.0215	5.3679	76.4202	21.5656
2006	33.9287	5.3688	77.9669	21.1119
2007	33.6432	5.3518	77.8082	21.2038
2008	33.3526	5.2617	79.3487	21.3244
2009	33.1375	5.2502	80.6109	21.1141
2010	33.2271	5.4388	80.1852	21.2980
2011	32.7504	5.5901	82.3195	20.9977
2012	32.3022	5.6184	84.6116	20.2397
all cohorts	33.8583	5.2284	77.4869	21.3629

Table 3.7: The mean and standard deviation of the average disability percentages for both the partially disabled state and the fully disabled state

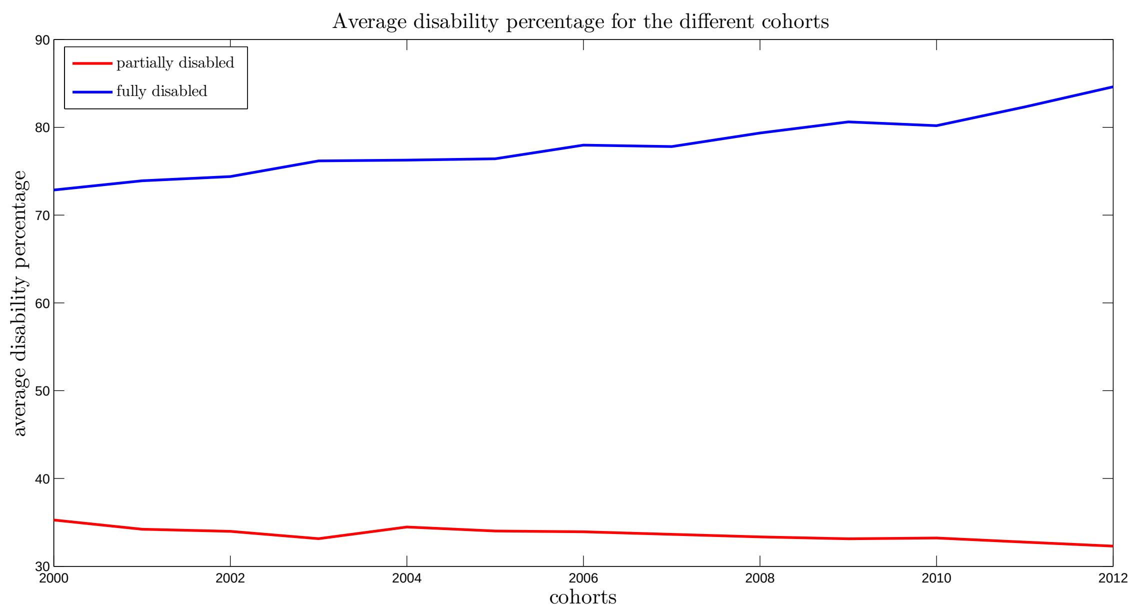

Figure 3.4: Plot of the average disability percentages evolving over the cohorts (in percentages)

Chapter 4

Discrete choice models

In this chapter discrete choice models are discussed: a dynamic binary logit model as well as a dynamic multinomial logit model. Before starting with the explanation of discrete choice models and elaboration on the dynamic binary logit model in Section 4.3 and on the dynamic multinomial logit model in Section 4.4, the difference between cross-sectional data and panel data is discussed and the Receiver Operating Characteristic curve is explained.

4.1 Panel data, a dynamic logit model

The data received from the insurance company consist of panel (or longitudinal) data, i.e. repeated observations are given for multiple policyholders. In more detail, the data consist of unbalanced panel data. This because the policyholders in the dataset are not all observed at the same time (they are neither observed at the same moments in time, neither for the same time periods).

Panel data handles the problem of heterogeneity. Heterogeneity between policyholders is handled by including covariates into the model. This can be done both in case of panel data as well as in case of cross-sectional data (data based on multiple policyholders in just one time aspect). However, in case of cross sectional data, heterogeneity over time could not be taken into account. Heterogeneity over time can be handled by panel data. A disadvantage of panel data could be that the data suffer from attrition. Attrition in the data could be due to death of the policyholder or ending the contract before the end date was reached (for example policyholders who quit the business and start being employed again). However, since the dataset contains information on the reason of ending of the claim, attrition is not a problem in this study.

In a logit model, the main assumption made is that the policyholders are independently distributed. This assumption implies that the policyholders behave independently of one another.

The most used term of "the logit model" is used in case the dataset would be a cross-sectional dataset. This is why the name "the logit model" is not used in case of panel data. The term "dynamic logit model" fits better in case of panel data, since the model includes dynamic variables: previous state and duration. Dynamic variables are variables that differ across time, and are discussed in Section 4.3.2 and Section 4.4.2.

4.2 Receiver Operating Characteristic curve

In order to estimate the correctness of the models, the Receiver Operating Characteristic (ROC) curves of the models are calculated. This ROC curve is a frequently used tool in the field of medical decision making.

Based on Fawcett (2006), the ROC curve can be explained according to Figure 4.1.

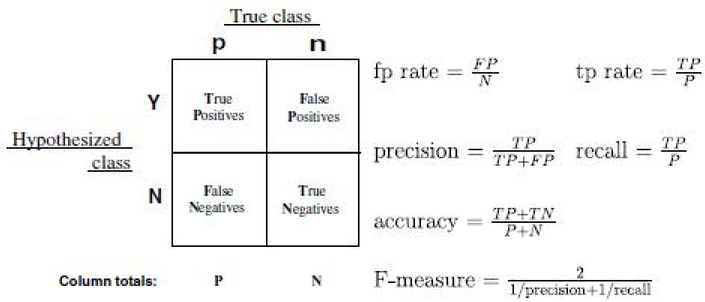

Figure 4.1: Principle of the ROC curve

As can be seen in Figure 4.1, the observations of both the true class (the observations available in the dataset) and the hypothesized class (the observations available based on the model) are taken into account when considering at the ROC curve. It is possible to calculate a so called *true positive rate* as the number of observations that re both positive in the true class as well as in the hypothesized class divided by the number of observations that are positive in the true class. Positive implies that the specific transition has occurred. The true positive rate is also called sensitivity. The *false positive rate* is calculated by dividing the number of observations that are positive in the hypothesized class but negative in the true class, by the number of observations that are negative in the true class. One minus the false positive rate is called specificity. The ROC curve is now given by plotting the false positive rate (the x-axis) against the true positive rate (the y-axis). In case the plot equals the 45-degree line, a random choice is considered. There is always 50% possibility that the correct answer is chosen. The further upwards from the 45-degree line, the better the model.

The ROC curve itself is interesting, but the most interesting part of this curve is the surface below this curve. The area under the ROC curve (mostly referred to as AUC) gives an accuracy of the model, which makes it obvious that the further the plot upwards from the 45-degree line, the better the model. In case the area equals one, the model is perfect. In case the area equals a half, the model is worthless.

The areas under the ROC curve of both discrete choice models as well as the areas under the ROC curves regarding some sensitivity analysis are given and discussed in Section 6.3.

4.3 Dynamic binary choice model

The focus is on the question how the disability percentages of policyholders evolve over time, and thus whether (and if so, when) a policyholder goes from one state to one another. The focus is therefore on binary outcome models. As already mentioned, it is assumed that policyholders are independently distributed.

In total there are nine possible "transitions" which could be investigated, which are all summarized in Table 4.1. These transitions are subtracted from Figure 2.3, with the "transitions" to stay in the state one already was (t^{00}, t^{11} and t^{22}) included. The focus is on the transitions of the partially disabled state to both the active state and the fully disabled state (t^{10} and t^{12}) and on the transitions of the fully disabled state to both the active state and the partially disabled state (t^{20} and t^{21}). This because of the reason explained in Section 2.2.4. The models as discussed in this chapter and in the following chapter, are used separately to model p^{10}, p^{12}, p^{20} and p^{21}.

In case of transition $t^{ab}_{i,t}$, the question that arises is whether a policyholder (with covariates x_i and time of

	active	**partially disabled**	**fully disabled**
active	t^{00}	t^{01}	t^{02}
partially disabled	t^{10}	t^{11}	t^{12}
fully disabled	t^{20}	t^{21}	t^{22}

Table 4.1: Nine possible transitions

disability t) transfers from state a to state b. It could be either a success (there is a transfer from state a to state b) or there could be a failure (the policyholder who is in state a at time t will not be in state b at time $t+1$). This is why focusing on binary choice models.

In order to make use of a binary choice model, the transition of the partially disabled state to the active state and the transition of the partially disabled state to the fully disabled state are modeled independently. Just as the transition of the fully disabled state to the active state and the transition of the fully disabled state to the partially disabled state. This modeling would only be correct if the policyholders are either possible to transfer to one state, or are possible to transfer to another state. However, given a policyholder who is in the partially disabled state at time t, this policyholder could be either in the active state, in the partially disabled state or in the fully disabled state at time $t+1$. This implies that modeling the transition probabilities p^{10}_{t-1} and p^{12}_{t-1} separately, is thus taking a binary choice model for modeling the transition probabilities, cannot be seen as a good model to model the transition probabilities upfront. However, in this chapter and later in Chapter 6 and Chapter 7, the results of this model will be discussed. This since the binary choice model could be a good method to estimate the distribution of the benefits taken into account, given not only the correctness of the model but also referring to the ease of this method.

4.3.1 Probabilities

In case the transfer from state a to state b is a success, $t^{ab}_{i,t}$ equals one, in case the transfer is a failure, $t^{ab}_{i,t}$ equals zero. It is assumed that $t^{ab}_{i,t}$, conditional on the information available at time $t-1$, is a random variable which follows the Bernoulli distribution with probability $p^{ab}_{i,t-1}$. Written in an equation, $t^{ab}_{i,t}$ is given by

$$t^{ab}_{i,t} = \begin{cases} 1 & \text{with probability} \quad p^{ab}_{i,t-1}; \\ 0 & \text{with probability} \quad 1-p^{ab}_{i,t-1}, \end{cases}$$

where $p^{ab}_{i,t-1}$ is the probability to transfer from state a to state b for a policyholder with covariates x_i and time of the disability t.

A binary choice model approaches the probability $p^{ab}_{i,t-1}$ (in the case $t^{ab}_{i,t}$) based on a vector of covariates x_i (which are introduced in Section 4.3.2), time t and an error term ($\epsilon^{ab}_{i,t}$). The binary choice model is given by the following equation:

$$t^{ab}_{i,t} = \mathbb{1}\{\beta^{ab'} x_{i,t-1} \geq \epsilon^{ab}_{i,t}\}, \tag{4.1}$$

where $\mathbb{1}\{\cdot\}$ equals the indicator function. $t^{ab}_{i,t}$ has value one in case the argument ($\beta^{ab'} x_{i,t-1} \geq \epsilon^{ab}_{i,t}$) is true, and value zero in case the argument in not true. Transition $t^{ab}_{i,t}$ is modeled by means of maximizing the (log-)likelihood of $t^{ab}_{i,t}$. The (log-)likelihood is given at the end of this section.

The difference between binary choice models is about the distribution in the error term. $F_{\epsilon^{ab}_{i,t}}$ has to be defined, which is the cumulative distribution function of $\epsilon^{ab}_{i,t}$. This cumulative distribution function is given by the following equation

$$F_{\epsilon^{ab}_{i,t}}(y) = Pr\left(\epsilon^{ab}_{i,t} \leq y\right)$$

Since $Pr\left(\epsilon^{ab}_{i,t} \leq \beta^{ab'} x_{i,t-1}\right) = Pr\left(t^{ab}_{i,t} = 1\right)$, it is given that $F_{\epsilon^{ab}_{i,t}}\left(\beta^{ab'} x_{i,t-1}\right) = p^{ab}_{i,t-1}$.

For the dynamic binary logit model the assumption regarding the error term is that it follows a logistic distribution. Therefore $F_{\epsilon^{ab}_{i,t}}(y)$ is given by

$$F_{\epsilon^{ab}_{i,t}}(y) = \frac{\exp(y)}{\exp(y)+1} \tag{4.2}$$

This implies the following formula for the transition probability of a policyholder (with covariates x_i and time of disability t) from state a to state b:

$$p^{ab}_{i,t-1} = \frac{\exp(\beta^{ab'} x_{i,t-1})}{\exp(\beta^{ab'} x_{i,t-1})+1} \tag{4.3}$$

In order to calculate the transition probabilities, it is necessary to estimate the coefficients of the dynamic binary logit model. The method to estimate the coefficients of the dynamic binary logit model is by means of the maximum likelihood estimator. Given the assumptions, the log-likelihood is given by

$$\log\left(L(\beta^{ab})\right) = \sum_j \sum_t t^{ab}_{i,t} \cdot \log\left(\frac{\exp(\beta^{ab'} x_i)}{1+\exp(\beta^{ab'} x_i)}\right) + (1-t^{ab}_i) \cdot \log\left(1 - \frac{\exp(\beta^{ab'} x_i)}{1+\exp(\beta^{ab'} x_i)}\right)$$

where j stand for the policyholders ($j = 1, \cdots, 29756$) and t equals the time ($t = 1, \cdots, 132$). The coefficients of the covariates are calculated by maximizing this log likelihood function.

4.3.2 Covariates

Based on the information given in the dataset, eight covariates are included in modeling the transition probabilities:

- $x_{1,t}$, duration of the disability so far. This variable differs over time;
- x_2, sex of the policyholder. Sex is a dummy variable (either value one or value zero). Female is taken as base level;
- x_3, age of the policyholder at the start of the disability;
- x_4, insured amount of the policyholder;
- x_5, x_6 & x_7, class of profession of the policyholder. Class of profession could be one, two, three or four and is a dummy variable. Class of profession one is taken as base level;
- $x_{8,t}$, previous state of the policyholder. This variable differs over time.

Next to the covariates stated above, the model also includes a constant, which is marked as β^{ab}_0. For this variable, x_0 is just a vector of ones. Note that in Section 3.2 the expected influences of these covariates on the different transition probabilities are discussed.

There are two possible ways to define the duration of the disability. There is the so called *clock forward* approach and the *clock reset* approach. The difference between the two possibilities is that in case of clock forward the duration is defined as the duration of the disability from the start of the disability, while in case of clock reset the duration is defined as the duration of the disability from the time the policyholder entered the specific state. In this study, whenever discussing duration, the duration from the start of the disability is meant and so in this study the clock forward approach is used.

In Section 4.3.3 is discussed which covariates are significant for which transition probability.

4.3.3 Estimated coefficients based on a dynamic binary logit model

In this section the transition probabilities are discussed that are modeled by the dynamic binary logit model. In Appendix A, Table A.1, the estimates of the coefficients for each covariate are given. From this table it is possible to distinguish between significant and insignificant covariates, for which the value of $\alpha = 0.05$ is used.

In the subsequent subsections the results are discussed for each transition probability separately. It has to be noted that those covariates are omitted which turned out to be insignificant, and the transition probabilities are modeled again without the insignificant variables. Furthermore note that all the coefficients given in Appendix A, Table A.1, are rounded to 4 decimals. The covariate insured amount is given in Euros, and despite the fact that a change of a single Euro only has a small influence, a change of thousands of Euros may have a big influence. This is why the estimated coefficient (just as the standard error of the coefficient) of the covariate insured amount is displayed times 10^{-4}.

A coefficient worth mentioning is the coefficient of the covariate previous state. The covariate previous state has an impact on the transition probabilities. For all transitions, this covariate is significant. As explained in Section 2.3, this implies that the Markov property does not hold.

4.3.3.1 Transition probability of transferring from partially disabled to active

For transition probability p^{10} it can be seen in Appendix A, Table A.1, that the covariate insured amount is insignificant based on $\alpha = 0.05$. Also the constant in the model turned out to be insignificant. The estimated coefficient for the covariate class of profession two is also insignificant. The assumption in made that in case at least one dummy variable turns out to be insignificant, the dummy covariate is omitted from the model. This is why the covariate insured amount as well as the covariate class of profession are omitted from the model. The model with duration, sex, age and previous state as covariates remained. This model is used to approach the probability to transfer from the state of being partially disabled to the state of being active. The constant turned out to be insignificant, but is not removed from the model. The estimated coefficients of the significant variables of this model can be found in Table 4.2.

p^{10}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-0.0623 (0.0803)	0.9396
duration	-0.0990 (0.0014)*	0.9057
sex	-0.1350 (0.0440)*	0.8738
age	-0.0232 (0.0016)*	0.9771
previous state	0.4814 (0.0155)*	1.6183

* significant at $\alpha = 0.01$

Table 4.2: Estimated coefficients including standard error of the modeled transition probabilities of the transition partially disabled to active

Focusing on the transition probability from the state of being partially disabled to the state of being active (Table 4.2), it can be shown that the variables duration and age have a negative influence. As explained in Section 3.2, this is intuitively clear. The coefficient of sex tells us that being a male (sex = 1) leads to a lower probability to transfer to the active state. the sign of the covariate previous state is positive. The positive influence of the covariate previous state is reasonable, due to the fact that the policyholders who were in the state of being fully disabled before they turned into the state of being partially disabled are recovering.

An overview of the transition probabilities, for policyholders with different sex and different previous states, can be found in Figure 4.2. In this figure it can be seen that policyholders with previous state 2 have a higher probability to transfer to the active state than policyholders with previous state 0. Furthermore it can be seen that female policyholders have a slightly higher probability to transfer compared to male policyholders.

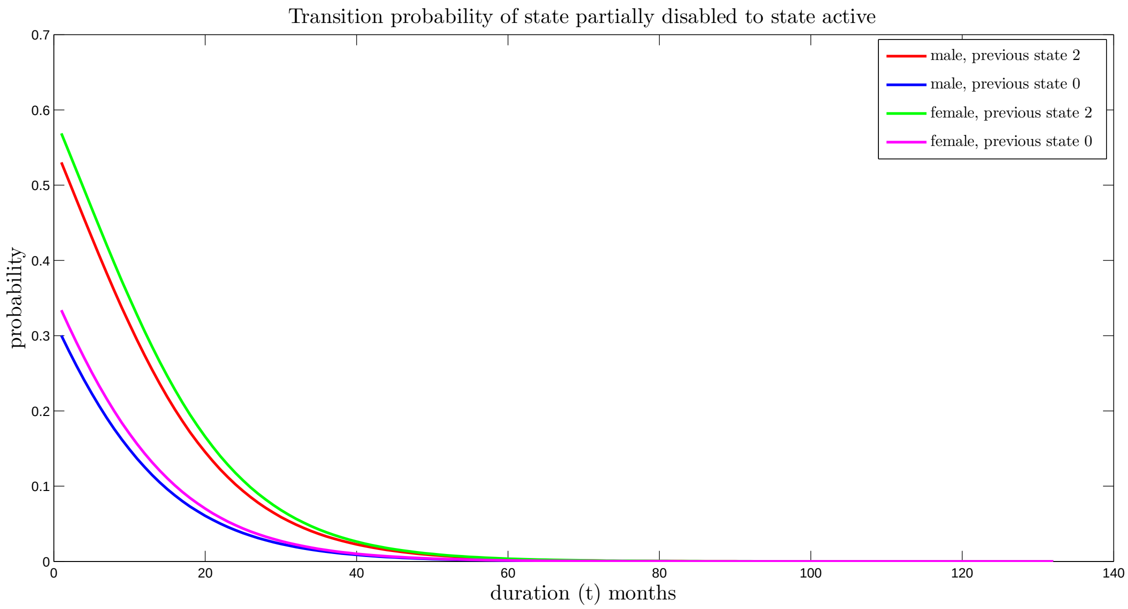

Figure 4.2: Transition probabilities for the transition of the state partially disabled to the state active, for policyholders with age 25 at the start of the disability

4.3.3.2 Transition probabilities of transferring from partially disabled to fully disabled

In case of the transition probability to go from the state of being partially disabled to the state of being fully disabled, the model remains which only includes a constant and the covariates duration, age and previous state. See Table 4.3. The other covariates are omitted because they turned out to be insignificant (Appendix A, Table A.1).

p^{12}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-2.3878 (0.1088)*	0.0918
duration	-0.0197 (0.0007)*	0.9805
age	-0.0068 (0.0023)*	0.9932
previous state	-0.0690 (0.0192)*	0.9333

* significant at $\alpha = 0.01$

Table 4.3: Estimated coefficients including standard error of the modeled transition probabilities of the transition partially disabled to fully disabled

Just as when modeling the transition probability p^{10}, also when modeling the transition probability of the transition from the partially disabled state to the fully disables state, the covariates duration and age have a negative impact. Since it is expected that the covariates duration and age have a negative influence on all recovering transition probabilities (duration even on all transition probabilities), this outcome is reasonable.

The coefficient of the covariate previous state is negative, which is caused by the fact that it is likely that the policyholders are dealing with deterioration of the disability.

An overview of the transition probabilities, for policyholders with different ages (at the start of the disability) and different previous states, can be found in Figure 4.3. From this figure it can be seen that the probability of a policyholder with age 25 (at the start of the disability) and previous state 0, is about the same as the probability of a policyholder with age 45 and previous state 2. The same holds for a policyholder with age 35 and previous state 0 and a policyholder with age 55 and previous state 2. A decrease in age of twenty years has therefore the same impact as having as previous state the active state instead of the fully disabled state.

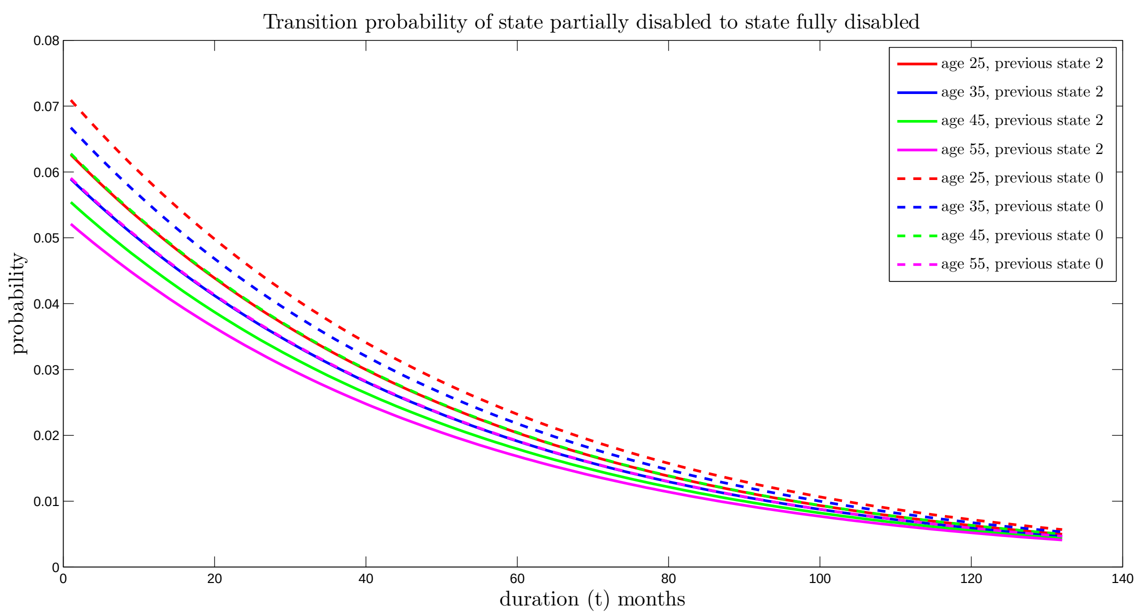

Figure 4.3: Transition probabilities for the transition of the state partially disabled to the state fully disabled

4.3.3.3 Transition probabilities of transferring from fully disabled to active

As can be seen in Appendix A, Table A.1, all variables in the case of approaching the transition probability from the state of being fully disabled to the state of being active (p^{20}) are significant. The results as discussed are also displayed in Table 4.4.

A remarkable outcome is the fact that the coefficients of the covariate sex are opposite for transition probabilities p^{10} and p^{20}. Furthermore, the sign of the coefficient of the covariate previous state of transition probability p^{20} is opposite to the sign for this coefficient for transition probability p^{10}, namely negative. This is clear since in case policyholders were in the partially disabled state before turning into the fully disabled state, their disability percentages are increasing which make it is plausible that these policyholders have a lower probability to recover to the active state. The positive sign for the covariate sex corresponds to the outcome as given in Van Waarden (2012)

An overview of the probabilities of transition t^{20}, for policyholders with different sex and different class

p^{20}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-1.5527 (0.1064)*	0.2117
duration	-0.0717 (0.0010)*	0.9308
sex	0.2963 (0.0312)*	1.3448
age	-0.0166 (0.0009)*	0.9835
insured amount	$-0.0922 \cdot 10^{-4}$ $(0.0086 \cdot 10^{-4})$*	1.0000
class of profession two	0.3070 (0.0980)*	1.3594
class of profession three	0.3081 (0.0968)*	1.3608
class of profession four	0.4952 (0.0933)*	1.6409
previous state	-0.1919 (0.0384)*	0.8254

* significant at $\alpha = 0.01$

Table 4.4: Estimated coefficients including standard error of the modeled transition probabilities of the transition fully disabled to active

of profession, can be found in Figure 4.4. Since the probabilities for the policyholders with class of profession two and class of profession three overlay, it can be concluded that there is no difference between the probabilities to transfer from the fully disabled state to the active state for policyholders with both classes of profession. Since the probabilities for the female policyholders are below the probabilities for the male policyholders (comparing male and female policyholders with the same classes of profession), it is given that male policyholders have a higher probability to transfer to the active state. In Figure 4.5, the transition probabilities p^{20} are given for policyholders with different ages (at the start of the disability), different insured amounts and different previous states. From Figure 4.5 it is possible to see that a change in the covariate age of 30 years has a larger impact on the transition probabilities than a change of the previous state (from previous state partially disabled to previous state active) and than a change in the insured amount of a policyholder of about €12,000 to €13,000. This since all the probabilities of the policyholders with age 25 (at the start of the disability) are higher than the probabilities of the policyholders with age 55.

.

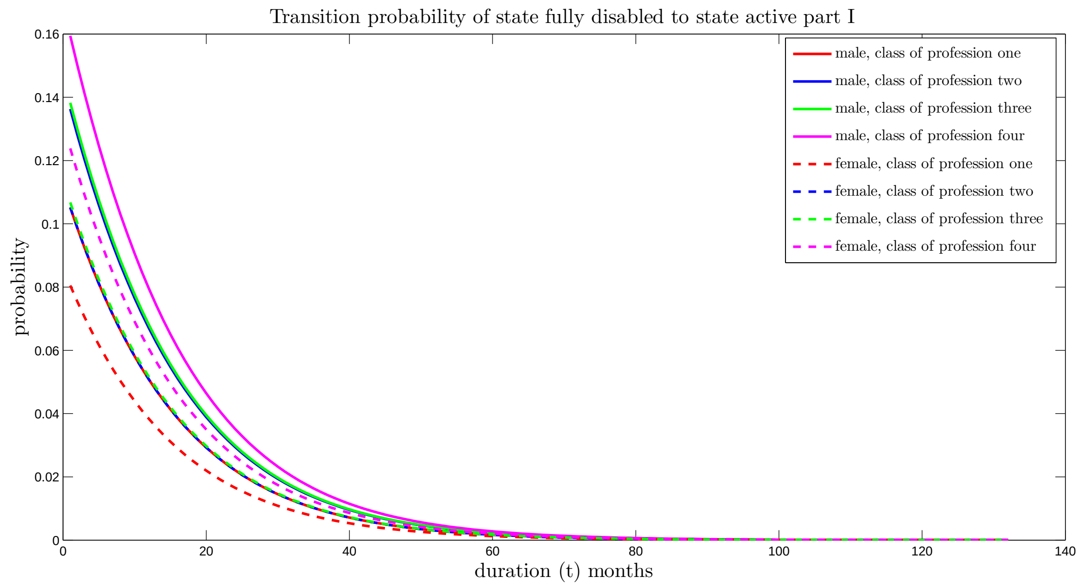

Figure 4.4: Transition probabilities for the transition of the state fully disabled to the state active, for policyholders with age 25 at the start of the disability, insured amount of €25,000 and previous state 1

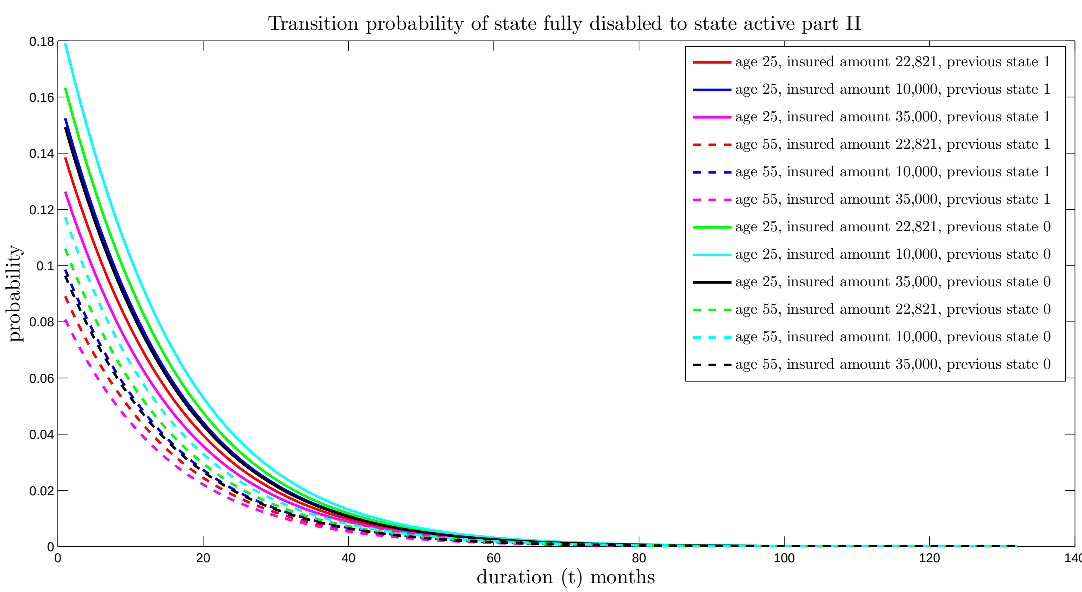

Figure 4.5: Transition probabilities for the transition of the state fully disabled to the state active, for male policyholders with class of profession three

4.3.3.4 Transition probabilities of transferring from fully disabled to partially disabled

The last transition probability modeled by the dynamic binary logit model is the transition probability from the fully disabled state to the partially disabled state of being partially disabled. The only covariate omitted in this transition probability is the variable sex.

p^{21}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-2.5891 (0.1316)*	0.0751
duration	-0.0346 (0.0007)*	0.9660
age	-0.0171 (0.0012)*	0.9831
insured amount	$0.0582 \cdot 10^{-4}$ $(0.0103 \cdot 10^{-4})$*	1.0000
class of profession two	0.2962 (0.1199)**	1.3447
class of profession three	0.3370 (0.1188)*	1.4008
class of profession four	0.3904 (0.1145)*	1.4776
previous state	1.0227 (0.0292)*	2.7807

* significant at $\alpha = 0.01$

** significant at $\alpha = 0.05$

Table 4.5: Estimated coefficients including standard error of the modeled transition probabilities of the transition fully disabled to partially disabled

From Table 4.5 it can be seen that the odds of the policyholders with class of profession three are 40.08% higher than for policyholders with class of profession one. The sign of the coefficient of the covariate previous state is positive. The estimated coefficients of the covariates age and duration have, as already predicted in Section 3.2, a negative sign.

In Figure 4.6 an overview of the transition probabilities, for policyholders with different classes of profession and different previous states, is given. In this figure it can be seen that the previous state 1 gives a higher probability than the previous state 0. This implies that the sign of the coefficient previous state is positive. In Section 3.2 it is discussed that the classes of profession are distinguished between the probability to become disabled (policyholders with class of profession three have average probability, policyholders with either class of profession one or class of profession two have low probability and policyholders with class of profession four have high probability). In Figure 4.6 it can be seen that the same ordering holds for the probability to transfer from the fully disabled state to the partially disabled states. Policyholders with class of profession four have the highest probability to transfer. Figure 4.7 gives the transition probabilities p^{20} for policyholders with different ages (at the start of the disability) and different insured amounts. As expected, younger policyholders have a higher probability to transfer from the fully disabled state to the partially disabled state. The covariate insured amount also has a positive effect in modeling the transition probabilities. The higher the insured amount of a policyholder, the higher the probability to transfer.

.

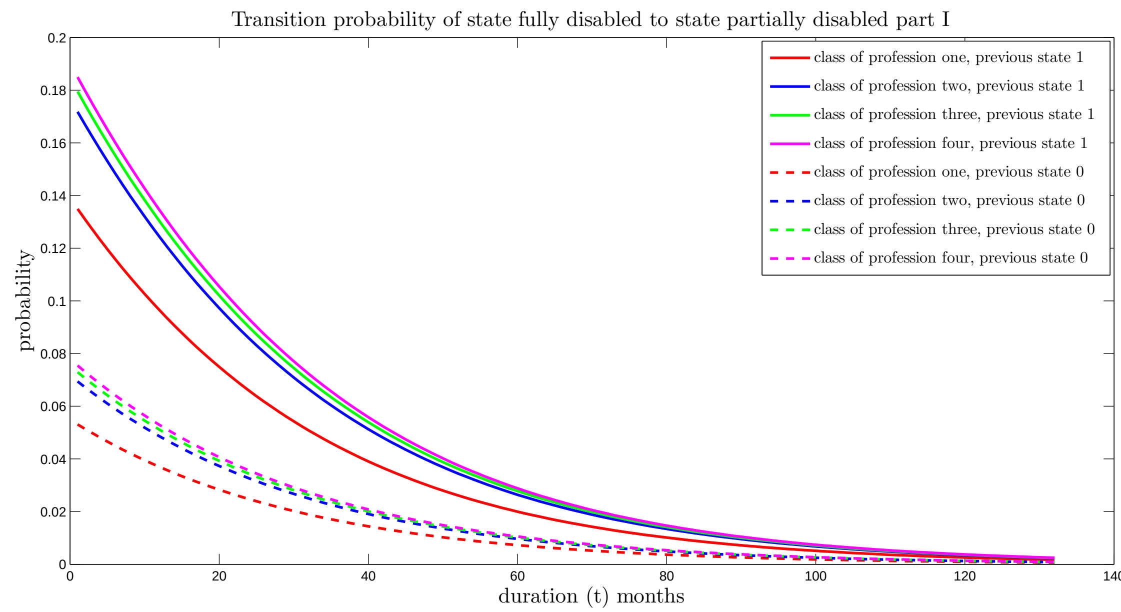

Figure 4.6: Transition probabilities for the transition of the state fully disabled to the state partially disabled, for policyholders with age 25 at the start of the disability and an insured amount of €25,000

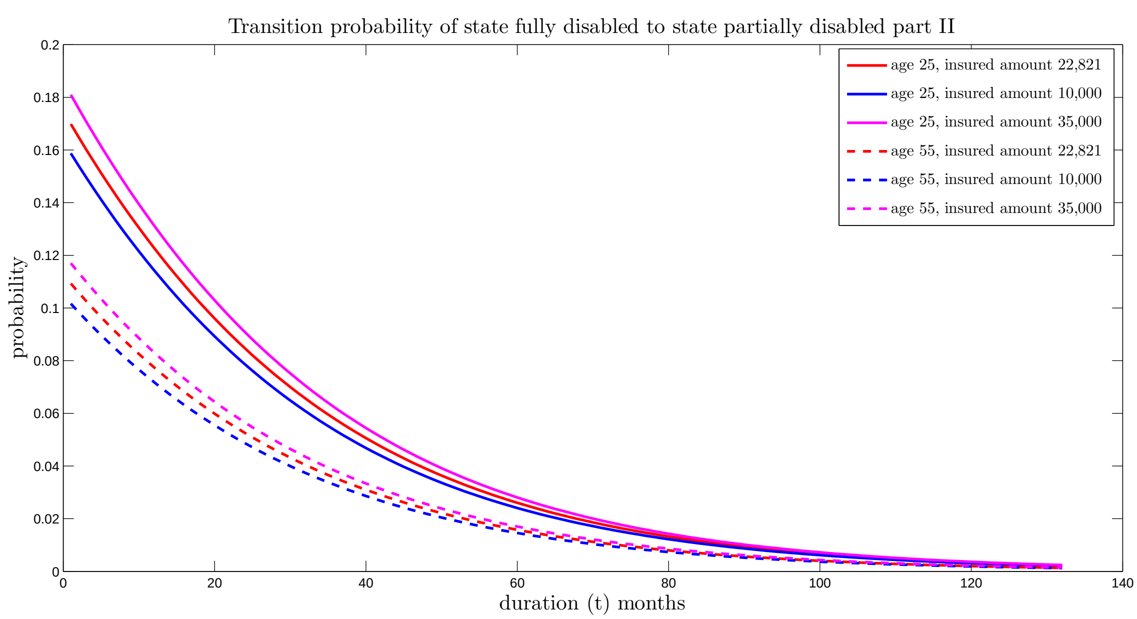

Figure 4.7: Transition probabilities for the transition of the state fully disabled to the state partially disabled, for policyholders with class of profession three and previous state 1

4.4 Dynamic multinomial choice model

In the dynamic binary logit model as discussed in the previous sections, the transition probabilities p^{10}, p^{12}, p^{20} and p^{21} are modeled separately. This is the result of choosing a binary choice model. In case of transition t^{10}, a transition from the partially disabled state to the active state is called a success, and a failure in case the policyholder would stay in the partially disabled state. However, it would also be possible to investigate the policyholders who are in the state of partially disabled at this moment (state 1), and model the probabilities that these policyholders will either go to state 0 (active), will stay in state 1 (partially disabled) or will go to state 2 (fully disabled) in one model. This could be done based on a multinomial choice model. In this section the dynamic multinomial logit model is discussed.

4.4.1 Probabilities

The multinomial logit model is summarized by the following equation

$$t^a_{i,t} = \begin{cases} 0 & \text{with probability} \quad p^{a0}_{i,t-1}; \\ 1 & \text{with probability} \quad p^{a1}_{i,t-1}; \\ 2 & \text{with probability} \quad 1 - p^{a0}_{i,t-1} - p^{a1}_{i,t-1}. \end{cases}$$

for $a \in \{1,2\}$. This because the transition probabilities to transfer from the initial state 0 (the active state, and so the probabilities to become disabled) cannot be estimated from the dataset (as explained before).

In case of facing a dynamic multinomial logit model, the question could be asked whether this model is an ordered model, a model in which the alternatives can be ordered, or an unordered model, a model in which the alternatives cannot be ordered. In modeling transition probabilities, since the different states can be ordered based on their disability percentages, an ordered model could be used. However, due to the default setting in Matlab, the unordered model is used.

In case of the multinomial logit model, the focus is on the utility function $u^{ab}_{i,t}$. The utility function for a transition of state a to state b, given covariates x_i and duration t, is given by

$$u^{ab}_{i,t} = \beta^{ab'} x_{i,t-1} + \epsilon^{ab}_{i,t} \tag{4.4}$$

Since a policyholder could be either in the active state, in the partially disabled state or in the fully disabled state, there are three alternatives in this multinomial logit model. The alternative chosen, or the disability state transferring to, is that utility function which gives the highest value above the utility functions of the other alternatives. Lets define $t^a_{i,t}$ as the transition from the initial state a given the covariates x_i and duration t, it is given that

$$t^a_{i,t} = j \quad \text{if} \quad u^{aj}_{i,t} \geq u^{aj\prime}_{i,t} \quad \text{for all} \quad j\prime \in \{0,1,2\}, j\prime \neq j \tag{4.5}$$

Just as in case of the binary choice model, the differences between multiple multinomial choice models are based on the distribution of the error term. The dynamic multinomial logit model assumes that the error terms (note that there are multiple error terms due to the definition of the transition $t^a_{i,t}$) are distributed independently according to the type I extreme value distribution. This distribution is given by the following equation

$$F_{\epsilon^{ab}_{i,t}}(y) = \exp(-\exp(-y)). \tag{4.6}$$

Due to the assumptions of the type I extreme value distribution of the error terms, the transition probabilities to transfer from state a ($a \in \{1,2\}$) to state b ($b \in \{0,1,2\}$) for policyholders with covariates x_i and duration t are given by

$$p^{ab}_{i,t-1} = \frac{\exp(\sum_{i=0}^{8} \beta^{ab}_i \cdot x_i)}{\sum_{j=0}^{2} \exp\left(\sum_{i=0}^{8} \beta^{aj}_i \cdot x_i\right)}, \tag{4.7}$$

with $j \in 0, 1, 2$ referring to the different states the policyholder transfer to. Note that in this study $a \in \{1, 2\}$, since it is not possible to model the probabilities of transitions t^{01} and t^{02} based on the dataset.

A restriction that has to be made is about setting β^{ab} equal to zero for one transition per different state a policyholder can be in. In this study, β^{12} and β^{22} are equal to zero.

4.4.2 Covariates

The covariates used in the dynamic multinomial logit model are the same as the covariates that have been used in the dynamic binary logit model. Since including the covariates duration and previous state, which differ across time, the multinomial logit model again can be seen as a dynamic multinomial logit model.

4.4.3 Estimated coefficients based on a dynamic multinomial logit model

The results of modeling the transition probabilities according to the dynamic multinomial logit model are given in Appendix B, Table B.1. One of the disadvantages of the (unordered) dynamic multinomial logit model is that in case a covariate is insignificant for one transition and significant for another transition, you cannot omit the variable for only one of both transitions. Hence you either have to omit a significant variable in one transition, or you have an insignificant variable in one transition. This problem is handled consistently by omitting the variable in case it turned out to be insignificant for either one or both transitions.

4.4.3.1 Transition probabilities from the partially disabled state

For the transition probabilities from the state of being partially disabled, the first model included all the covariates. The variables sex, insured amount and the dummy variables class of profession turned out to be insignificant, which leads to the results presented in Table 4.6.

p^{10}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	1.6855 (0.1285)*	5.3951
duration	-0.0783 (0.0016)*	0.9247
age	-0.0132 (0.0027)*	0.9869
previous state	0.4459 (0.0190)*	1.5619
p^{11}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	1.8761 (0.1115)*	6.5277
duration	0.0238 (0.0008)*	1.0241
age	0.0109 (0.0023)*	1.0110
previous state	0.0730 (0.0158)*	1.0757

* significant at $\alpha = 0.01$

Table 4.6: Statistics, of transition from state partially disabled, excluding insignificant covariates

Focusing on the transition of the partially disabled state to the active state, it can be seen that just as in the binary choice model, the variables duration and age have a negative impact on the transition probabilities. The variable previous state has a positive sign. Just as explained in Section 4.3.3.1, these outcomes seem reasonable.

The transition probability of staying in the partially disabled state (so actually no transition) has been modeled in the dynamic multinomial logit model. For this it is shown that all the significant covariates are positive. So the longer a policyholder is in the state of being totally disabled, and the older the policyholder is at the start of the disability, the higher the probability that the policyholder will not transfer to another state (either active or fully disabled). Furthermore, in case the policyholder was in the state of being fully disabled before turning partially disabled, there is a higher probability that the policyholder will stay in the state of being partially disabled. Also this seems reasonable since the time to recover could be larger in case one was fully disabled than in case one was active.

In the left graph of Figure 4.8, an overview of the probabilities of transition t_{10} is given for policyholders with different ages (at the start of the disability) and different previous states. As can be seen from this figure, a change in the age at the start of the disability of 30 years (from age 25 to age 55) has a smaller impact than a change in the previous state from the fully disabled state to the active state for a policyholder with age 25 (at the start of the disability). In the right graph of Figure 4.8 the probabilities to stay in the partially disabled state one already was are given for policyholders with different ages (at the start of the disability) and different previous states. This figures shows that a change in previous state (from the state of being active to the state of being fully disabled) has a lower impact than a change of age at the start of the disability of 30 years (from age 55 to age 25) for a policyholder with the active state as previous state.

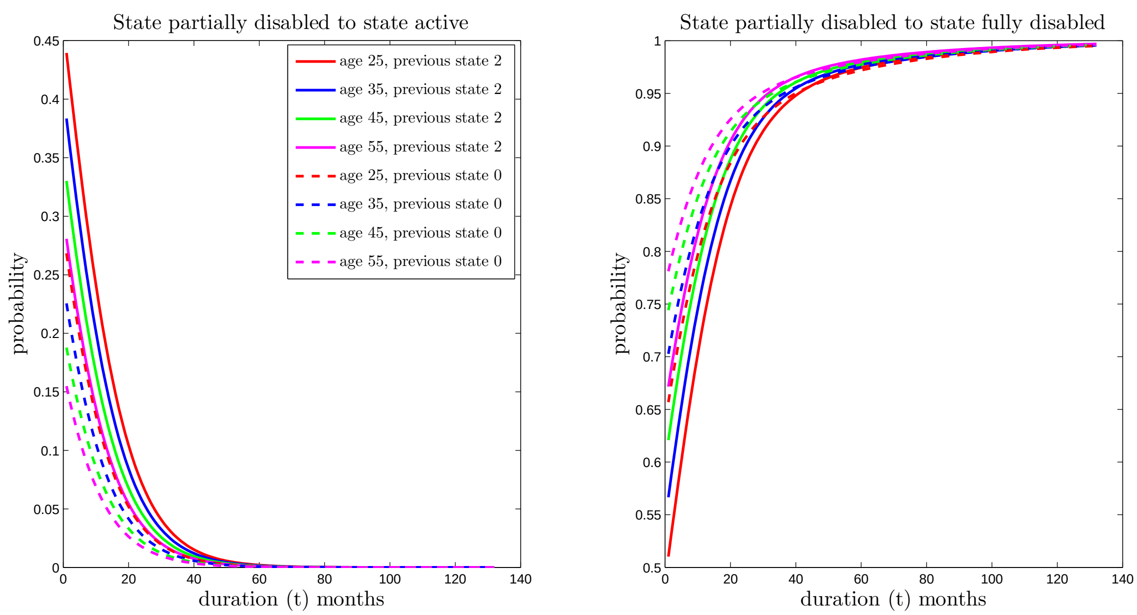

Figure 4.8: Transition probabilities for the transition of the state of being partially disabled to the state of being active (left) and for the transition of the state of being partially disabled to the state of being fully disabled (right)

4.4.3.2 Transition probabilities from the fully disabled state

Since all the variables are significant, no covariate is omitted from the model. The results of Appendix B, Table B.1 are given as well in Table 4.7.

p^{20}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-1.3910 (0.1068)*	0.2488
duration	-0.0704 (0.0010)*	0.9320
sex	0.2989 (0.0313)*	1.3483
age	-0.0176 (0.0010)*	0.9825
insured amount	$-0.0905 \cdot 10^{-4}$ $(0.0086 \cdot 10^{-4})$*	1.0000
class of profession two	0.3016 (0.0982)*	1.3520
class of profession three	0.3117 (0.0970)*	1.3658
class of profession four	0.4780 (0.0935)*	1.6129
previous state	-0.1632 (0.0154)*	0.8494
p^{21}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-2.6107 (0.1327)*	0.0735
duration	-0.0361 (0.0007)*	0.9645
sex	0.0971 (0.0375)*	1.1020
age	-0.0198 (0.0012)*	0.9804
insured amount	$0.0721 \cdot 10^{-4}$ $(0.0104 \cdot 10^{-4})$*	1.0000
class of profession two	0.3663 (0.1199)*	1.4424
class of profession three	0.4336 (0.1189)*	1.5427
class of profession four	0.4917 (0.1149)*	1.6351
previous state	0.3352 (0.0161)*	1.3982

* significant at $\alpha = 0.01$

Table 4.7: Statistics of dynamic multinomial logit model

For the transition probability of the state of being fully disabled to the state of being active, it can be seen that, compared to the dynamic binary logit model, all variables have the same direction. Also for the transition probability of the state of being fully disabled to the state of being partially disabled, the signs of the significant covariates are the same. However, note that the covariate sex was not included in the final model of transition probability p^{21}.

The coefficients of the covariates duration and age (at the start of the disability) have a negative sign. All other covariates are positive. Policyholders with class of profession four have a higher probability of transferring from the state of being fully disabled to the state of being partially disabled compared to policyholders with class of profession one.

In the left graph of Figure 4.9, an overview of the probabilities of transition t^{20} is given for policyholders with different insured amounts and different classes of profession. In the right graph of Figure 4.9 the transition probabilities of transition t^{21} are given again for policyholders with different insured amounts and different classes of profession.

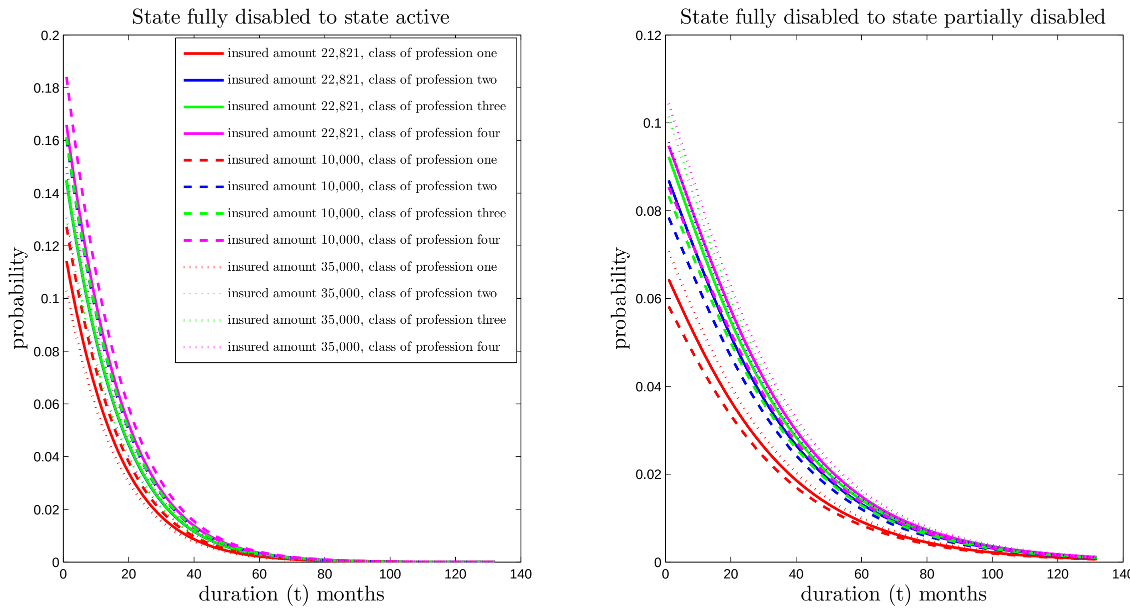

Figure 4.9: Transition probabilities for the transition of the state of being fully disabled to the state of being active (left) and for the transition of the state of being fully disabled to the state of being partially disabled (right), for male policyholders with age 25 at the start of the disability and previous state 1

4.5 Summary

In this chapter discrete choice models are introduced. These discrete choice models are dynamic logit models, which gave us the transition probabilities of four transitions: t^{10}, t^{12}, t^{20} and t^{21} in case of the dynamic binary logit model and t^{10}, t^{11}, t^{20} and t^{21} for the dynamic multinomial logit model. The model includes six covariates and those variables are omitted that turned out to be insignificant. Based on the approached transition probabilities it is possible to calculate the estimate of the benefits that need to be paid to the policyholders within the next twelve months. This will be discussed in Chapter 6. In the next chapter survival analysis models are discussed in detail. The Cox Proportional Hazard model is a well known model which is frequently investigated regarding the transition probabilities of disability. The Cox Proportional Hazard model is discussed in Section 5.3.

Chapter 5

Survival analysis models

The logit models as explained in Chapter 4, are discrete choice models. In this chapter a semi-parametric survival analysis model is discussed. A survival analysis model creates a survival function, where discrete choice models directly creates a probability. This is further explained in detail in this chapter.

The inspiration for this chapter is based on Koning and Spierdijk (2011) and Bultena (2009). Both Koning and Spierdijk (2011) and Bultena (2009) worked with the mixed proportional hazard model in order to model the disability percentages of the policyholders who are disabled. Furthermore, a lot of information about survival analysis is explained in both Cameron and Trivedi (2005) and Angrist and Pischke (2009).

In this chapter the details of the theory behind survival analysis models are explained in Section 5.1. This is done by introducing the terms hazard rate and survival function. After the theory of survival analysis models the focus is on the semi-parametric Cox proportional hazard model.

5.1 Survival analysis

Survival analysis is about the question how long a certain policyholder will "survive" in a specific state. In case one focus on a disability model with two disability states, namely active and disabled, the question is how long a policyholder who is in the disabled state will stay in this state. Since this study focuses on a multiple state model, survival times between the states of actives, partially disabled and fully disabled are considered. In Section 4.3.2 the different probabilities to define the duration of the disability are discussed and the assumption is made to use the clock forward way. This assumption is also made in this chapter. Note that this assumption is discussed in Section 7.2. Furthermore, just as in the discrete choice models, the assumption is made that the policyholders are independently distributed.

An important equation in the context of survival analysis is the survival function. The survival function is the probability that the duration T exceeds t, given by equation (5.1).

$$\begin{aligned} S(t) &= Pr(T > t) \\ &= 1 - F(t) \\ &= 1 - Pr(T \leq t) \end{aligned} \tag{5.1}$$

Note that $F(t)$ is the cumulative distribution function, which is the probability that t exceeds the duration T.

The hazard rate, which is closely related to the survival function, plays an important role in the survival analysis models. The hazard rate is the "instantaneous probability of leaving a state conditional on survival to time t", see, e.g., Cameron and Trivedi (2005). This means that it is the instantaneous rate that a failure occurs in the time interval $(t, t+h)$ for policyholders with no failure until $T = t$. A failure is defined as a policyholders who does not survives, and so a policyholder who transfers from one state to one another. The definition for the hazard rate, conditional on the characteristics of the policyholder (x_i), is given by

$$\lambda(t) \quad = \quad \lim_{h \to 0} \frac{Pr(t \leq T < t+h | T \geq t)}{h} \tag{5.2}$$

Equation (5.2) can be rewritten into

$$\begin{aligned} \lambda(t) &= \lim_{h \to 0} \frac{Pr(t \leq T < t+h | T \geq t)}{h} \\ &= \frac{f(t)}{S(t)} \\ &= \frac{f(t)}{1 - F(t)} \end{aligned} \tag{5.3}$$

The function $f(t)$ in equation (5.3) is the probability density function, which equals the derivative of the cumulative distribution function $(F'(t))$. The proof of equation (5.3) can be found at page 245 of Verbeek (2004).

When investigating survival analysis models, an important definition is duration dependence. Duration dependence exists if the hazard rate changes (either decreases or increases) over time. In this study, since the covariates age, sex, insured amount, class of profession and previous state are included, the question of duration dependence has to be about two policyholders which have the same characteristics regarding the covariates.

As discussed in Chapter 4, duration proves to be a significant covariate in discrete choice models. This covariate has a coefficient with a negative sign for almost each transition probability (except for the probability to stay in the partially disabled state in case of the multinomial logit model). This gives an intuitive feeling that in case of disability, there is negative duration dependence. This implies that policyholders who are disabled for a longer period (either partially or fully disabled) have a lower probability to transfer to the active state. Also in case of the transition from the state of being partially disabled to the state of being fully disabled and in case of transition from the state of being fully disabled to the state of being partially disabled, negative duration dependence is expected.

5.2 Kaplan-Meier estimate

The Kaplan-Meier estimate (hereafter KM estimate) is a non-parametric estimator which can can used to estimate the survival function $S(t)$ as given in equation (5.1). The fact that it is a non-parametric estimator implies that it does not assume a pre-specified form of the survival function.

The general equation for the KM estimate for the transition of state a to state b is given in equation (5.4).

$$\widehat{S}^{ab}(t) \quad = \quad \prod_{j | t_j \leq t} \frac{r_j - d_j}{r_j}, \tag{5.4}$$

where $t_1 \leq t_2 \leq t_3 \leq \cdots \leq t_N$. In the dataset used for this study, $N = 132$.

In equation (5.4) r_j is the amount of policyholders which are in state a at time t_j and d_j is the amount of policyholders which will be in state b at time t_{j+1}. For r_j the policyholders which are in state a at time

t_j but who are right-censored in time t_{j+1} are excluded. The variance of the KM estimate for transition of state a to state b is given in equation (5.5).

$$\widehat{var}\left(\widehat{S}^{ab}(t)\right) = \widehat{S}^{ab}(t)^2 \sum_{t_j<t} \frac{d_j}{r_j(r_j-d_j)} \tag{5.5}$$

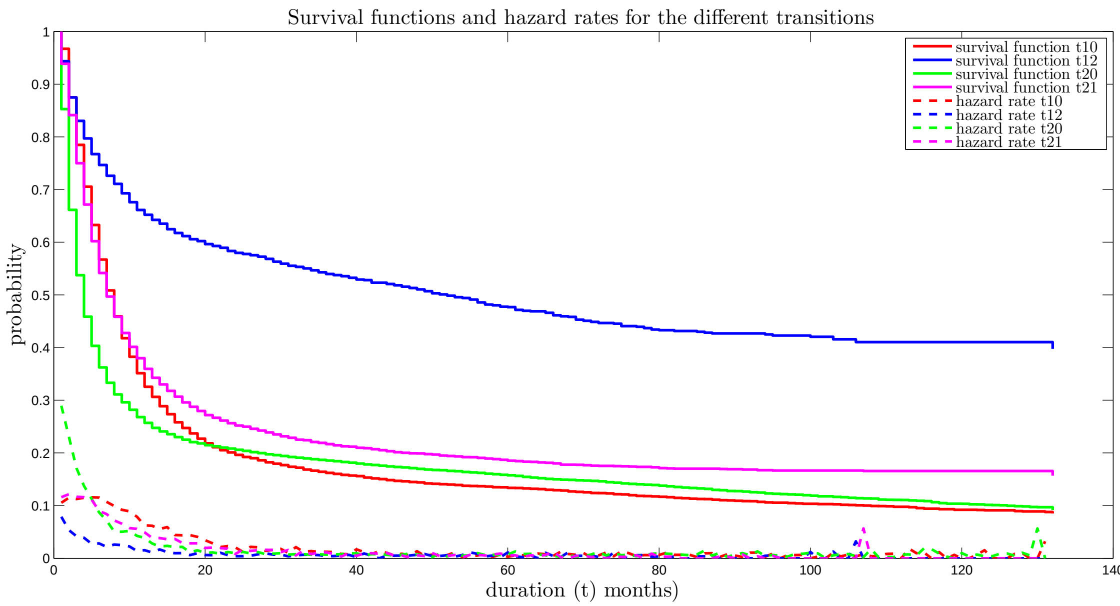

Figure 5.1: Overview the survival function and hazard rates

Note that the data are discrete time (monthly disability percentages are given in the dataset), and so equation (5.4) and equation (5.5) are given in discrete time formulation. In Figure 5.1 the KM estimates of $\widehat{S}^{10}(t)$, $\widehat{S}^{12}(t)$, $\widehat{S}^{20}(t)$ and $\widehat{S}^{21}(t)$ can be found as well as the hazard rates belonging to these estimates of the survival function. Equation (5.3) and equation (5.1) are used to get the hazard rate out of the KM Estimate. Due to the fact that the following equation holds in discrete time,

$$\begin{aligned} F(t) - F(t-1) &= Pr(T \leq t) - Pr(T \leq t-1) \\ &= Pr(T = t) \\ &= f(t) \end{aligned}$$

the hazard rate is given by:

$$\begin{aligned} \widehat{\lambda}(t) &= \frac{\widehat{f}(t)}{\widehat{S}(t)} \\ &= \frac{(1-\widehat{S}(t)) - (1-\widehat{S}(t-1))}{\widehat{S}(t)} \end{aligned}$$

Figure 5.1 shows that the survival function for the transition of the partially disabled state to the fully disabled state still has high values after 132 months. This leads to low values for transition probability p^{12}. When the values of the survival function decrease fast, the probability to transfer is high (for example for transition t^{10}, t^{20} and t^{21}). The shape of the survival functions implies that the probabilities to transfer decrease over time, because the slope of the survival functions decreases over time.

The KM estimate is not used for the estimation of the benefits that need to be paid to the policyholders in the coming twelve months. The reason for this is that the KM estimate only contains duration as covariate, while in this study it is preferred to include more covariates into the model. One of the reasons for this is that including covariates into the model handles with heterogeneity bias.

Due to the fact that policyholders may have different characteristics, it could be that the wrong information is given when time evolves. For example assume that male policyholders are so called "leavers", and that female policyholders are "stayers". With stayers it is meant that female policyholders have naturally high survival probabilities and that male policyholders have naturally low survival probabilities. Due to the low survival probabilities of men, they will exit faster than the women and this will lead to a group leftovers with proportional more female. Due to the resignation of the men, it seems like there is some negative duration. This does not necessarily have to be the case.

The fact that policyholders may have different characteristics is called heterogeneity bias. This heterogeneity bias is not taken into account in this section. Heterogeneity bias could be taken into account in proportional hazard models, which are discussed in Section 5.3.

5.3 Cox PH Model

In equation (5.2), the equation of the hazard rate is given. In case heterogeneity bias has to be modeled, a proportional hazard model could be taken into consideration. For proportional hazard models, the assumption has to be made that the covariates are multiplicatively related to a baseline hazard. The hazard rate $\lambda(t|x_i)$ for a proportional hazard model is given by:

$$\lambda(t|x_i) \quad = \quad \lambda_0(t)\exp(x_i'\beta), \tag{5.6}$$

where $\lambda_0(t)$ is the baseline hazard.

The baseline hazard is the part of the hazard rate which applies to all policyholders. The fact that it is a proportional model implies that the hazard rate for an individual policyholder is based on the baseline hazard, and moves proportionally to the baseline hazard with the individual characteristics of the policyholder. Hence the multiplicative relation that is mentioned above.

Based on the hazard rate, the survival function is then:

$$S(t|x_i) \quad = \quad \exp\left(-\int_0^t \left(\lambda_0(u)\exp(x_i'\beta)\right)du\right) \tag{5.7}$$

The Cox PH model is a so called semi-parametric model. As explained in Section 5.2, a non-parametric model does not have a pre-specified distribution. Focusing on equation (5.6), the baseline hazard ($\lambda_0(t)$) has a non-parametric form whereas $\exp(x_i'\beta)$ has a parametric form. Therefore the Cox PH model is referred to as a semi-parametric model.

An overview of the transition probabilities based on the Cox PH model are given in Section 5.4.1, Section 5.4.2, Section 5.4.3 and Section 5.4.4. In Section 5.3.1 the covariates which are used in the Cox PH model are summarized and in Section 5.3.2 the method of competing risks analysis is discussed. Competing risks analysis is the method used in this study, this since in case a policyholder transfers, there are two possible states to transfer to.

5.3.1 Covariates in case of the Cox PH model

The covariates that are considered in the Cox PH model are in accordance with the covariates as used in the discrete choice models:

- x_1, sex of the policyholder. Sex is a dummy variable (either value one or value zero). Female is taken as base level;
- x_2, age of the policyholder at the start of the disability;
- x_3, insured amount of the policyholder;
- x_4, x_5 & x_6, class of profession of the policyholder. Class of profession could be one, two, three or four and is dummy variable. Class of profession one is taken as base level (more information about the covariate class of profession if given in Section 3.2);
- x_7, previous state of the policyholder. This variable differs over time.

In case of the Cox PH model, the constant is included in the baseline hazard. Duration is excluded as covariate, but is not excluded from the model since the baseline hazard depends on the duration ($\lambda_0(t)$).

5.3.2 Competing risks

Survival analysis considers a survival time to a specific event. In case of the transition probability from the state of being partially disabled to the state of being fully disabled, the time a policyholder stays in the partially disabled state before transferring to the state of being fully disabled is called the survival time. In case a policyholder stays in the state of being partially disabled until the end of the dataset, the data is right-censored. In this survival time, only the policyholders who are in the partially disabled state are considered, and either stay in that state or go to the fully disabled state. The policyholders who are in the state of being partially disabled and who transfer to the state of being active, are construed as right-censored.

Another way to model the transition probabilities with survival analysis is by means of competing risks analysis. In this section competing risks analysis is explained. This is done based on Bakoyannis and Touloumi, Beyersmann et al. (2009), Calle et al. (2007) and Van Waarden (2012).

Competing risks analysis differs from survival analysis in taking into account multiple events that could occur. The idea behind the competing risks analysis can be seen in Figure 5.2. The number of the different states (the initial state, the event of interest and the competing event) in Figure 5.2 differs from the number of the states in this study.

Figure 5.2 shows that in case a policyholder transfers from the initial state, there are two possible paths the policyholder can follow, the policyholder moves into the event of interest or the policyholders moves into the competing event. In case the initial state of the policyholder is the partially disabled state, the competing event would be the fully disabled state and in case the initial state of the policyholder is the fully disabled state, the competing event would be the partially disabled state. The event of interest is always the active state.

The states the policyholders could be in are either state 0 (active), state 1 (partially disabled) and state 2 (fully disabled). The initial state (I) a policyholder is in is referred to state a, i.e., $I = a$. The initial state can be either state partially disabled (state 1) or state fully disabled (state 2). State 0 (the active state) is referred to as the event of interest. The competing event is either the partially disabled state (in case $I = 2$) or the fully disabled state (in case $I = 1$).

Following Calle et al. (2007), the pair (T, E) is introduced. This pair gives us both the failure time as

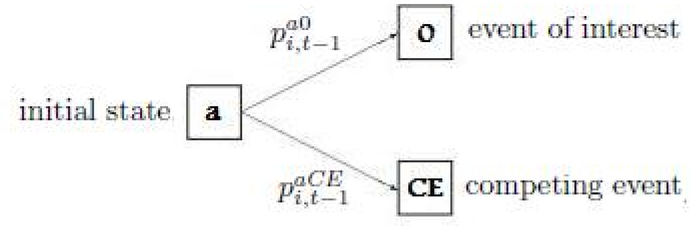

Figure 5.2: Idea of competing risks

well as the failure event. T is the failure time which is defined as the time at which a transition occurs, and E is the event which occurred at the failure time T. In case the transition is from the initial state to the active state, the failure time is the failure time to transition 0 ($T = T_0$). In case the transition is to the competing event, T equals T_{CE}. Here T_{CE} stand for the failure time the competing event occurs. It could be that no transition occurs at all. In case no transition occurs at all the data is right-censored and hence the event is censored ($E = C$ and $T = T_C$). The observed failure time T equals the minimum of T_0, T_{CE} and T_C, i.e., $T = \min\{T_0, T_{CE}, T_C\}$.

The competing risks process of the initial state of being partially disabled is determined through both λ^{10} and λ^{12} and of the initial state of being fully disabled by λ^{20} and λ^{21}. λ^{ab}, the hazard rate of the transition from the initial state a ($I = a$) to the state b (either the event of interest 0 or the competing event CE), is given by equation (5.8).

$$\lambda^{ab}(t) = \lim_{h\to 0} \frac{Pr(t \le T < t+h, E = b|I = a, T > t)}{h}, \quad b \in \{0, CE\} \tag{5.8}$$

Equation (5.2) can be rewritten as

$$\begin{aligned}\lambda^{a}(t) &= \lim_{h\to 0} \frac{Pr(t \le T < t+h|I = a, T \ge t)}{h} \\ &= \lim_{h\to 0} \frac{Pr(t \le T < t+h, E = 0|I = a, T \ge t)}{h} + \lim_{h\to 0} \frac{Pr(t \le T < t+h, E = CE|I = a, T \ge t)}{h} \\ &= \lambda^{a0}(t) + \lambda^{aCE}(t)\end{aligned} \tag{5.9}$$

The survival function (as given in equation (5.7)) can now be rewritten based on the event-specific hazard rates. This is done by:

$$S^a(t|x_{i,t-1}) = \exp\left(-\int_0^t (\lambda_{a0}(u|x_{i,u-1}) + \lambda_{aCE}(u|x_{i,u-1}))\, du\right) \tag{5.10}$$

$S^a(t|x_{i,t-1})$ is the survival function of the initial state a. The interest lies in the probability of surviving in the initial state up to duration t, conditional on surviving up to duration $t-1$. This probability, p^{aa}_{t-1}, is given in the following equation

$$\begin{aligned}p^{aa}(t-1|x_{i,t}) &= \frac{S^a(t|x_{i,t-1})}{S^a(t-1|x_{i,t-2})} \\ &= \exp\left(-\int_{t-1}^t \left(\lambda^{a0}(u|x_{i,u-1}) + \lambda^{aCE}(u|x_{i,u-1})\right) du\right)\end{aligned} \tag{5.11}$$

It is known that, in case a failure occurs, the event of interest will occur or the competing event will occur. From the point of view of the policyholder, a failure implies that the policyholder transfers from the initial state to either the active state or the competing event state. Again the competing event state depends on the initial state a policyholder is in. As can be seen in Beyersmann et al. (2009), the probability that a policyholder transfers to another state at duration t, $1-p_t^{aa}$, has to be assigned to both the probability that the policyholder will transfer to the active state and the probability that the policyholder will transfer to the competing event state. In equation (5.12) it is shown how $1-p_t^{aa}$ is assigned to state b ($b \in \{0, CE\}$).

$$\begin{aligned} Pr(E=b|t \leq T < t+h,) &= \frac{Pr(t \leq T < t+h, E=b|T \geq t)}{Pr(t \leq T < t+h|T \geq t)} \\ &= \frac{\int_0^t \left(\lambda^{ab}(u|x_{i,u-1})du\right)}{\int_0^t \left(\left(\lambda^{a0}(u|x_{i,u-1}) + \lambda^{aCE}(u|x_{i,u-1})\right)du\right)} \end{aligned} \tag{5.12}$$

Equation (5.12) gives the basis for both $p^{a0}(t)$ and $p^{aCE}(t)$ which are given in the equations below.

$$p^{a0}(t-1|x_{i,t-1}) = (1-p^{aa}(t-1|x_{i,t-1})) \cdot \frac{\int_0^t \left(\lambda^{a0}(t|x_{i,u-1})du\right)}{\int_0^t \left(\left(\lambda^{a0}(u|x_{i,u-1}) + \lambda^{aCE}(u|x_{i,u-1})\right)du\right)} \tag{5.13}$$

$$p^{aCE}(t-1|x_{i,t-1}) = (1-p^{aa}(t-1|x_{i,t-1})) \cdot \frac{\int_0^t \left(\lambda^{aCE}(u|x_{i,u-1})du\right)}{\int_0^t \left(\left(\lambda^{a0}(u|x_{i,u-1}) + \lambda^{aCE}(u|x_{i,u-1})\right)du\right)} \tag{5.14}$$

$p^{a0}(t-1|x_{i,t-1})$ equals the probability to transfer from the initial state a to the active state conditional on x_i and $p^{aCE}(t-1|x_{i,t-1})$ equals the probability to transfer from the initial state a to the competing event state conditional on x_i.

Again, just as explained in Section 4.3.1, the likelihood is maximized in order to get the parameter coefficients. Before it is possible to state the likelihood, the indicator function which indicates whether a specific transition is made has to be introduced. In case at time t the event of interest occurs for policyholder i, it is known that $d_i^0 = 1$ and if the competing event occurs $d_i^{CE} = 1$. At every point in time, for every policyholder, either the event of interest occurs, the competing event occurs or the policyholder is censored. In case a policyholder does not transfer to the event of interest or to the competing event, it is given that $d_i^C = 1 - d_i^0 - d_i^{CE} = 1$. The likelihood to observe the event on a time T is given by

$$L_i(t) = Pr(t=T_0)^{d_i^0} \cdot Pr(t=T_{CE})^{d_i^{CE}} \cdot Pr(t=T_C)^{(1-d_i^0-d_i^{CE})} \tag{5.15}$$

This log likelihood is given in equation (5.16). Elaboration on the process from equation (5.15) to equation (5.16) is given in Appendix E.

$$\begin{aligned} \ell_i(t) &= \log L_i(t) \\ &= d_i^0 \cdot \log\left(\lambda^{a0}(t)\right) + \log\left(S^{a0}(t)\right) + d_i^{CE} \cdot \log\left(\lambda^{aCE}(t)\right) + \log\left(S^{aCE}(t)\right) \end{aligned} \tag{5.16}$$

Equation (5.16) shows that the maximum of the log likelihood, $\ell_i(t)$, is reached by maximizing the two separate parts of equation (5.16). Because the two separate parts are identical to the log likelihood of transition probability p^{a0} and the log likelihood of transition probability p^{aCE}, it is possible to model the transition probabilities p^{a0} and p^{aCE} separately.

5.4 The hazard rates of the Cox PH model

Just as in Chapter 4, the outcomes of the transition probabilities are discussed in this chapter. In Chapter 6, the focus is on the estimated distributions of the benefits to be paid to the policyholders. In the sections below, the hazard rates of the transition probabilities for each of the transitions t^{10}, t^{12}, t^{20} and t^{21} are discussed separately.

In Figure 5.3, an overview is given for the baseline hazards for the different transitions. It can be seen that the transitions t^{10}, t^{20} and t^{21} all show duration dependence, since the baseline hazards decrease as duration increase. The baseline hazard for the transition of the partially disabled state to the fully disabled state shows a fluctuating path. This is partly caused by the fact that there are less observations of policy-holders transferring from the state of being partially disabled to the state of being fully disabled, than there are observations of the other transitions. Another reason why the fluctuations in the upper right graph of Figure 5.3 are more visible comes from the fact that the scale is smaller. Since the transition probabilities in the Cox PH model are proportionally to the baseline hazards, the same fluctuating pattern is observed for the transition probability of the partially disabled state to the fully disabled state.

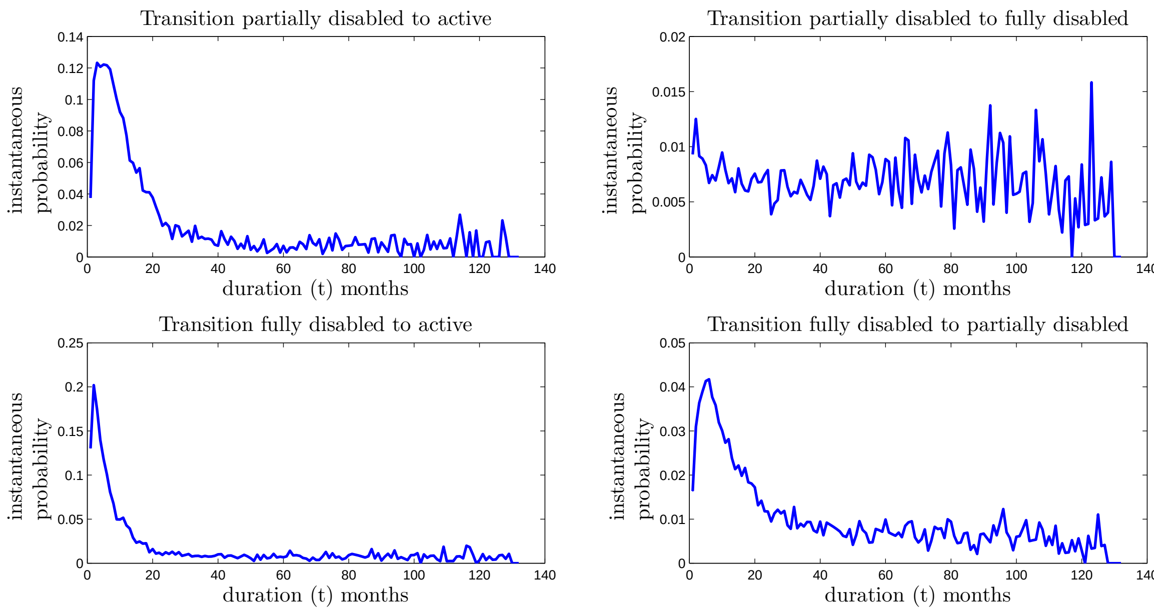

Figure 5.3: The baseline hazard rates for the different transitions

5.4.1 Hazard rate of transferring from partially disabled to active

As can be seen in Appendix C, Table C.1, the coefficients of covariate class of profession are not significant. Neither is the covariate sex. These covariates are therefore omitted from the model. A model with age, insured amount and previous state as covariates is used to model the transition probabilities of the partially disabled state to the active state. The results of the coefficients for these covariates are given in Table 5.1.

p^{10}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
age	-0.0134 (0.0014)*	0.9867
insured amount	$0.1706 \cdot 10^{-4}$ $(0.0105 \cdot 10^{-4})$*	1.0000
previous state	-0.1806 (0.0137)*	0.8348

* significant at $\alpha = 0.01$

Table 5.1: Statistics, of transition partially disabled to active, excluding insignificant covariates

Table 5.1 shows that both the covariates age and previous have a negative effect on the transition from the state of being partially disabled to the state of active. The coefficient of the covariate insured amount is positive. This implies that the higher the insured amount of the policyholder, the higher the probability that a policyholder will transfer to from the state of being partially disabled to the state of being active.

In Figure 5.4 the hazard rates are given for policyholders with different ages at the start of the disability (either age 25 or age 55) and different insured amounts (either insured amount €22,821 (average), €10,000 or €35,000). As expected, it can be seen that the younger the policyholder (at the start of the disability), the higher the probability to transfer from the partially disabled state to the active state. A decrease of the insured amount by €25,000 (from €35,000 to €10,000) gives a lower probability than a decrease of age at the start of the disability by 30 years (from 25 years to 55 years).

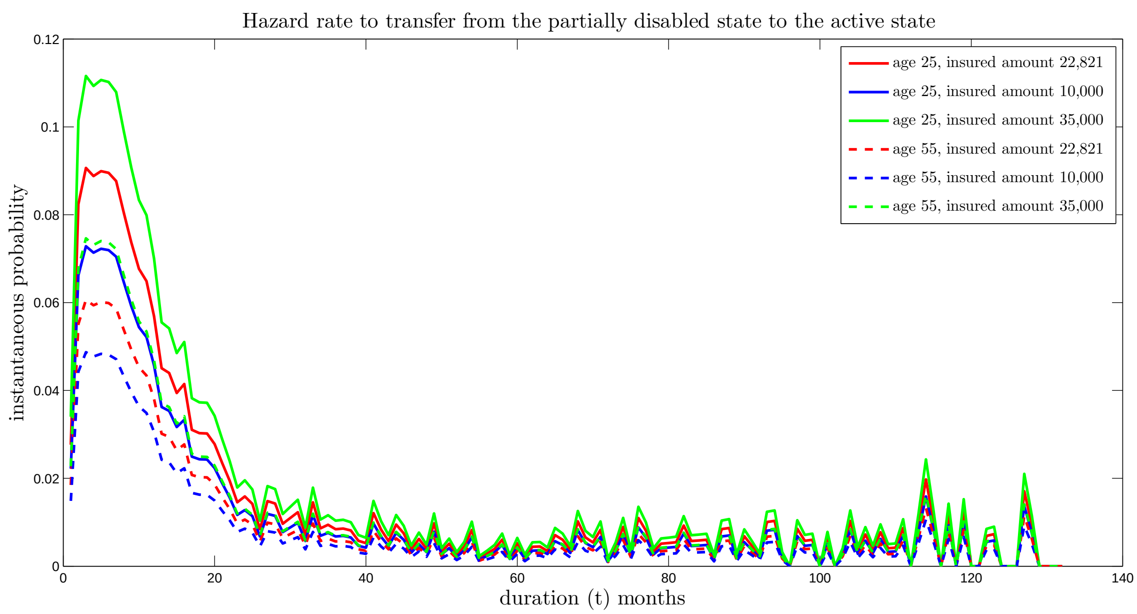

Figure 5.4: Hazard rates for the transition of the state partially disabled to the state active, for male policyholders with previous state 2 and class of profession one

5.4.2 Hazard rate of transferring from partially disabled to fully disabled

For the transition probabilities of the transition from the state of being partially disabled to the state of being fully disabled, all covariates turned out to be significant. The results, as given in Appendix C, Table C.1, are also given in Table 5.2.

In contradiction to the transition probability p^{10}, in case of the transition probability p^{12}, the covariate age has a positive influence on the transition probability. This is also as expected and discussed in Section 3.2. The previous state still has the same sign as in transition probability p^{12} and both the dummy covariates class of profession and the covariate insured amount have a positive influence.

In Figure 5.5, the hazard rates are given for policyholders with different ages at the start of the disability (either age 25 or age 55) and different insured amounts (either insured amount €22,821 (average), €10,000

p^{12}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
sex	-0.1418 (0.0661)**	0.8678
age	0.0060 (0.0024)**	1.0061
insured amount	$0.1670 \cdot 10^{-4}$ $(0.0185 \cdot 10^{-4})$*	1.0000
class of profession two	0.5591 (0.2348)**	1.7490
class of profession three	0.4740 (0.2330)**	1.6065
class of profession four	0.6964 (0.2273)*	2.0064
previous state	-0.4012 (0.0191)*	0.6689

* significant at $\alpha = 0.01$
** significant at $\alpha = 0.05$

Table 5.2: Statistics, of transition partially disabled to fully disabled, excluding insignificant covariates

or €35,000). As already mentioned in Section 6.3, the path of the transition probability from the state of being partially disabled to the state of being fully disabled is fluctuating a lot. The different probabilities as given in Figure 5.5 are not very far apart from each other. Note that the axes of the figure are from 0% to 1.6% only.

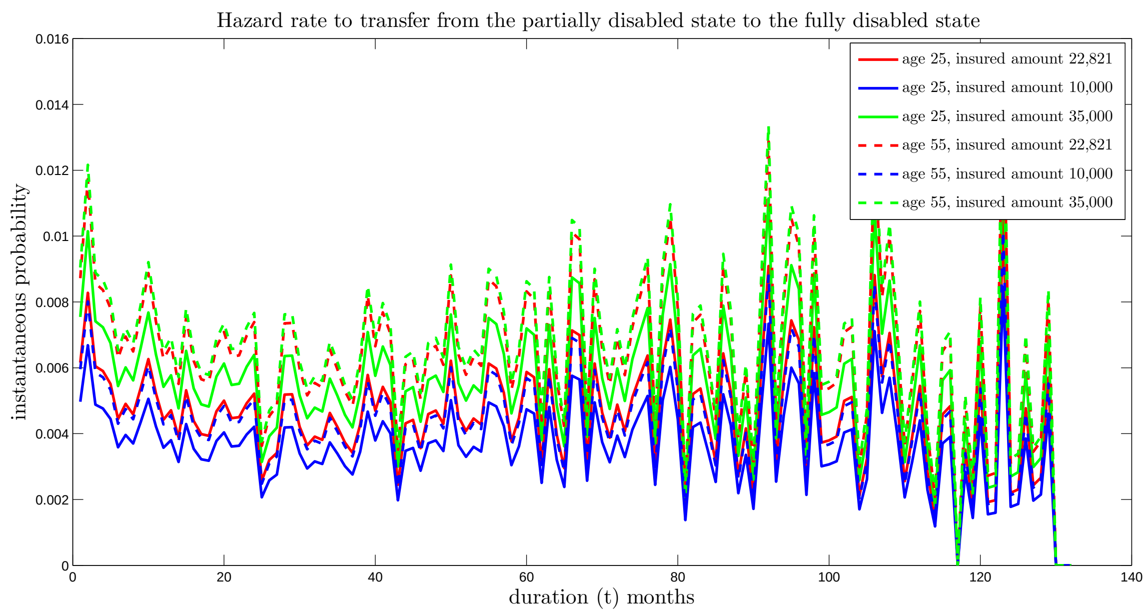

Figure 5.5: Hazard rates for the transition of the state partially disabled to the state fully disabled, for male policyholders with previous state 2 and class of profession one

5.4.3 Hazard rate of transferring from fully disabled to active

As well as in the transition of the state of being partially disabled to the state of being fully disabled, again in the transition from the state of being fully disabled to the state of being active, all covariates turn out to be significant. The result of modeling this transition probability are given in Table 5.3.

p^{20}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
sex	0.1649 (0.0297)*	1.1792
age	-0.0165 (0.0009)*	0.9836
insured amount	$0.0301 \cdot 10^{-4}$ ($0.0078 \cdot 10^{-4}$)*	1.0000
class of profession two	0.4727 (0.0950)*	1.6044
class of profession three	0.4523 (0.0939)*	1.5719
class of profession four	0.6999 (0.0906)*	2.0136
previous state	-1.1082 (0.0367)*	0.3302

* significant at $\alpha = 0.01$

Table 5.3: Statistics, of transition fully disabled to active, excluding insignificant covariates

In case of the transition from the state of being fully disabled to the state of being active, the coefficients of the covariates age and previous state are negative whereas the coefficients of the covariates sex, insured amount and class of profession are positive. All three (dummy) covariates of class of profession are compared to class of profession one. Since the coefficient of β_5 (of class of profession three) is smaller than the coefficient of β_4 (of class of profession two) there can be concluded that the in case class of profession two was taken as base level, the coefficient of class of profession three would be negative. Males tend to transfer faster to the active state and older policyholders (higher ages at the start of the disability) tend to transfer slower to the active state. Furthermore, in case the policyholder was partially disabled before turning fully disabled, the probability to transfer to the active state is smaller than in case the policyholder was active before turning fully disabled. These results may sounds reasonable.

In Figure 5.6 the hazard rates are given for policyholders with different class of profession (either class of profession two, three or four) and different previous states (either state partially disabled or state active). As can be seen in this figure, the probability to transfer from the fully disabled state to the active state is higher in case of previous state 0. Especially in the first months of the disability this gives a higher probability. The probabilities for policyholders with both class of profession two and class of profession three are close to each other. Policyholders with class of profession four have a higher probability to transfer.

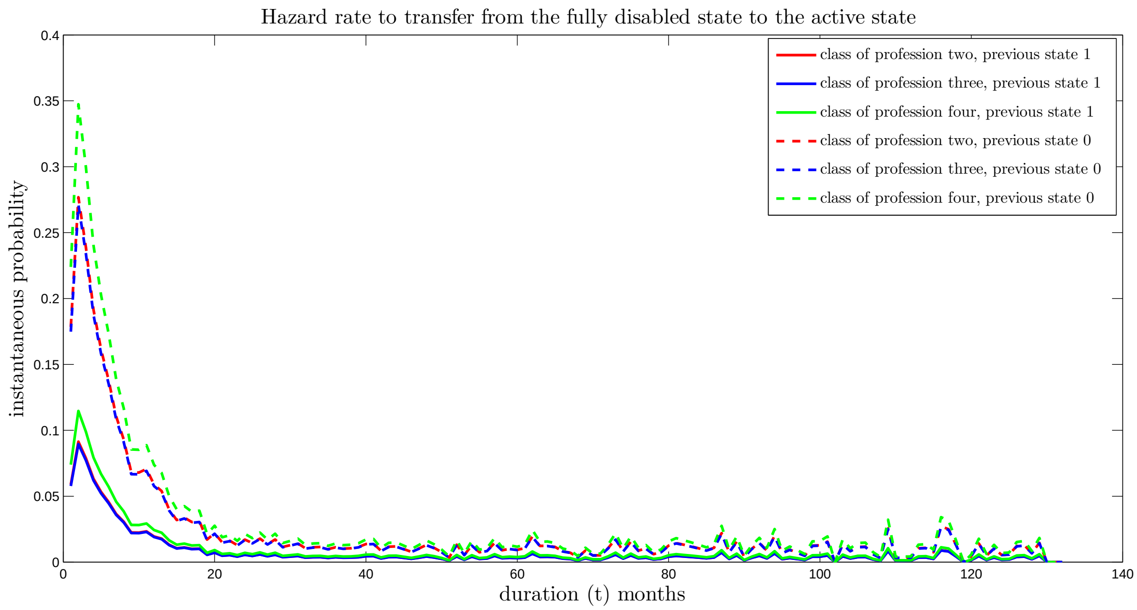

Figure 5.6: Hazard rates for the transition of the state fully disabled to the state active, for male policyholders with age 25 at the start of the disability and an insured amount of €30,000

5.4.4 Hazard rate of transferring from fully disabled to partially disabled

As can be seen in Appendix C, Table C.1, the only covariate that is not significant for transition probability p^{21} is sex. The covariate sex is omitted from the model and the remaining covariates are used to model the transition probabilities of the transition from the fully disabled state to the partially disabled state. The results of the coefficients are given in Table 5.4.

p^{21}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
age	-0.0158 (0.0012)*	0.9843
insured amount	$0.1662 \cdot 10^{-4}$ $(0.0096 \cdot 10^{-4})$*	1.0000
class of profession two	0.4467 (0.1180)*	1.5631
class of profession three	0.5020 (0.1170)*	1.6520
class of profession four	0.5903 (0.1130)*	1.8045
previous state	-0.3850 (0.0272)*	0.6804

* significant at $\alpha = 0.01$

Table 5.4: Statistics, of transition fully disabled to partially disabled, excluding insignificant covariates

In case of transition probability p^{21}, the probability to transfer in case the previous state equals 1 (partially disabled) is lower than the probability to transfer in case the previous state equals 0 (active). Furthermore, the probability to transfer from the fully disabled state to the partially disabled state is higher for policyholders with either class of profession two, three or four compared to policyholders with class of profession one.

In Figure 5.7 the hazard rates are given for policyholders with different class of profession (either class of profession two, three or four) and different previous states (either state partially disabled or state active). Just as in the transition from the fully disabled state to the active state, the probability to transfer is higher for policyholders with previous state 0.

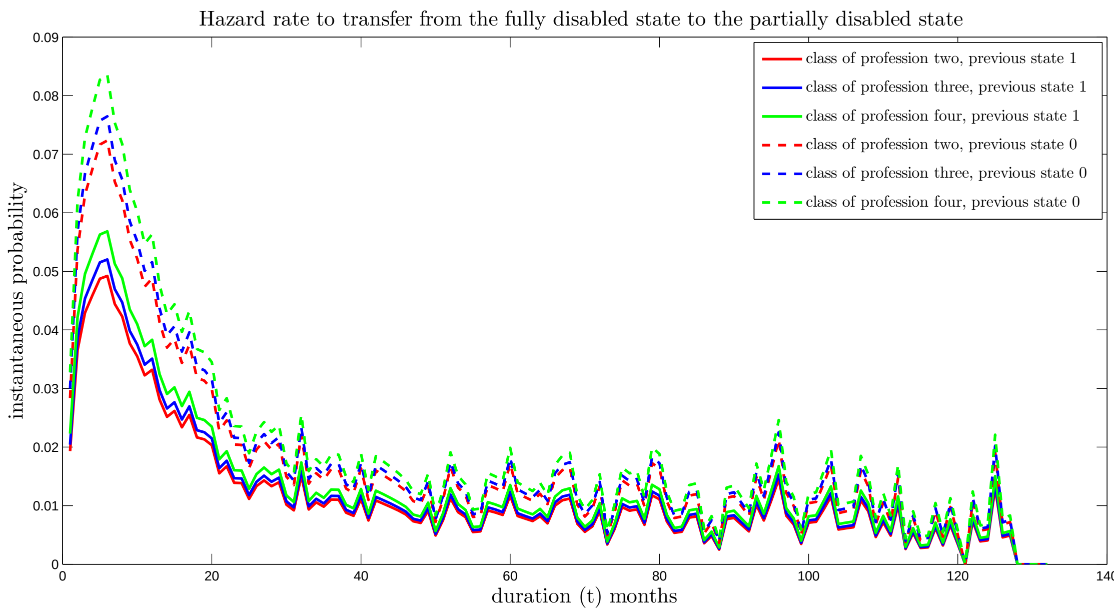

Figure 5.7: Hazard rates for the transition of the state fully disabled to the state partially disabled, for male policyholders with age 25 at the start of the disability and an insured amount of €30,000

5.5 Testing proportionality

As stated in Section 5.3, the assumption has to be made that the covariates are multiplicatively related to the baseline hazard. The model is called the Cox proportional hazard model, since the hazard rates should be proportional. More specific, the hazard rate for an individual with covariates x_{1i} should be proportional to the hazard rate for an individual with covariates x_{2i}. Since the hazard rate for an individual with covariates x_{1i} is given by $\lambda(t|x_{1i}) = \lambda_0(t)\exp(x_{1i}'\beta)$, and the hazard rate for an individual with covariates x_{2i} is given by $\lambda(t|x_{2i}) = \lambda_0(t)\exp(x_{2i}'\beta)$, the proportionality assumption implies that the ratio between the two different hazard rates, given by

$$\begin{aligned} \frac{\lambda(t|x_{1i})}{\lambda(t|x_{2i})} &= \frac{\lambda_0(t)\exp(x_{1i}'\beta_i)}{\lambda_0(t)\exp(x_{2i}'\beta_i)} \\ &= \frac{\exp(x_{1i}'\beta_i)}{\exp(x_{2i}'\beta_i)} \\ &= \exp((x_{1i}-x_{2i})'\beta_i), \end{aligned} \tag{5.17}$$

is constant.

One possibility to test whether $\exp((x_{1i}-x_{2i})'\beta_i)$ is constant, is by plotting $\log(S(t|x_{1i}))$ and $\log(S(t|x_{2i}))$.

This because

$$\begin{aligned}\log\left(S(t|x_{1i})\right) - \log\left(S(t|x_{2i})\right) &= x_{1i}'\beta_i + \log(\lambda_0(t)) - x_{2i}'\beta_i + \log(\lambda_0(t)) \\ &= \exp\left((x_{1i} - x_{2i})'\beta_i\right)\end{aligned}$$

In order for $\exp((x_{1i} - x_{2i})'\beta_i)$ to be constant, $\log\left(S(t|x_{1i})\right)$ and $\log\left(S(t|x_{2i})\right)$ have to be parallel.

One of the drawbacks of testing the proportionality graphically is that in case of a lot of covariates, or in case of continuous covariates, a lot of combinations need to be tested. This since the hazard rates need to be proportional for all individuals with certain characteristics x_i. In this study it would have been easier if the covariates age (at the start of the disability) and insured amount are taken in groups instead of taken continuous. Because the policyholder can have any age at the start of the disability and any insured amount (these are continuous covariates), it is impossible to check whether all possible hazard rates are proportional. Another drawback could be the fact that it is difficult to decide whether $\log\left(S(t|x_{1i})\right)$ and $\log\left(S(t|x_{2i})\right)$ are parallel or not. It is known for sure however that in case the survival function of different groups cross, that the proportionality assumption should be rejected. This is why in Appendix D some figures are included which show the survival functions for different groups of policyholders for the different transition probabilities. In Chapter 7 the figures as given in Appendix D are discussed in more detail.

5.6 Clock forward vs. clock reset

As already shortly mentioned in the beginning of Section 5.1, the clock forward definition of duration is assumed in the Cox PH model. This implies that the duration of the disability is taken from the start of the disability of a policyholder, and not only from the start of the specific state of the policyholder. Assume that a policyholder is disabled for thirteen months already, but transfered from the fully disability state to the partially disability state after nine months, the clock forward duration is thirteen months but the clock reset duration is only 4 months. In case of survival analysis, it is prescribed to use the clock reset definition of duration. As mentioned, this definition is not used in this study. In Section 7.2 it is discussed why in this study the clock forward definition of duration is assumed.

5.7 Summary

In this chapter the Kaplan Meier estimate and the Cox PH model are discussed and explained in detail. Because the Kaplan Meier estimate could not handle heterogeneity of the policyholders, the Cox PH model was introduced. Within the Cox PH model competing risks analysis is applied. Competing risks analysis is applied since in case a policyholder transfers to another state, there are two possible states the policyholder can transfer to. Either the event of interest occurs (always the active state) or the competing event occurs (the fully disabled state or the partially disabled state, dependent on the initial state). Given the modeled transition probabilities it is possible to estimate the distribution of the benefits to be paid to the policyholders within the subsequent twelve months to the measurement date. The results of these estimated distributions for both the Cox PH model as discussed in this chapter as well as for the discrete choice models as discussed in Chapter 4, are given in the Chapter 6.

Chapter 6

Results

In this chapter the results are given of the estimated distribution of the benefits that need to be paid to the policyholders that have right-censored data at the measurement date of the dataset, within the subsequent twelve months to this date. This implies that the dataset of disabled policyholders is considered at the time of the measurement date. The distribution of the benefits that need to be paid to the policyholder is estimated by means of 10,000 generated paths which are based on the transition probabilities as modeled by different models. The different models are as discussed in Chapter 4 and Chapter 5. Since the different models have different transition probabilities, different estimated distributions are anticipated. From the 29,756 claims in the dataset, 24,255 claims are finalized at the measurement date (June 30th, 2012). The remaining 5,501 policyholders still received a benefit in the month of the measurement date. For these 5,501 policyholders, 10,000 paths are generated of how their disability state evolves in the subsequent twelve months to the measurement date (up to and including June 2013). The generated paths take into account mortality rates ("Prognosetafel AG2012-2062") as well as the end age of the policyholder. Next to the expected value, the 95% quantile, the 99.5% quantile and the 95% coverage interval of the benefits are given.

In addition to the expected values and the quantiles which are given in this chapter, some sensitivity analyses and backtests are discussed in this chapter. In Section 6.2, the parameter uncertainty is modeled. In Section 6.4 the models are tested on their accuracy based on the results of a backtest.

6.1 Benefits to be paid

Before being able to generate 10,000 paths of 5,501 policyholders, the transition probabilities had to be modeled. Modeling the transition probabilities has been done in Chapter 4 and Chapter 5. With the transition probabilities it is possible to generate for each policyholders 10,000 paths of how the disability state evolves in the subsequent twelve months to the measurement data. For this it is assumed that the policyholders are independently distributed, which implies that policyholders behave independently one another. This allows the calculation of expected values and quantiles of the benefits. The task of generating paths is performed with Matlab[1].

Before presenting the results of the generated paths for the different models, the assumptions that have been made in generating paths are shortly discussed. In Section 6.1.2 a comparison between the different models is made. The significant covariates and the signs of the significant covariates as well as the transition probabilities of the different models are compared.

[1] Matlab 7.6.0 (R2008a) / Matlab R2011a.

6.1.1 Assumptions

In this section some (implicit) assumptions are discussed which have been made during the process of the distribution of the benefits to be paid (generating paths). After stated the assumptions made in the general overview of this study (assumptions applying to both the logit models as the Cox PH model) as well as comparing the different models based on the covariates as on the transition probabilities, the results of the estimated distribution of the benefits to be paid to the policyholders within the next twelve months are discussed.

6.1.1.1 Waiting period

The waiting period of a policyholder can be defined by the period the policyholder have to wait, after the start of the disability, before receiving a benefit. This waiting period can be chosen based on the preference of the policyholder. The longer the waiting time chosen, the lower the premium that needs to be paid to the insurance company. Most of the policyholders choose a waiting time of one month, but the period of fourteen days and the periods of three and six months can also be chosen. The waiting period for the policyholders is not given in the dataset and thus no statistics can be given. However, it is assumed that the waiting period has expired by the time the generating of the paths started. It is then known that none of the policyholders are in their waiting time and it is known for sure that the benefits will be paid out to the policyholder.

6.1.1.2 Missing data

As mentioned in Chapter 3, 9,257 policyholders out of the dataset with 29,756 policyholders have some missing data regarding the first month(s) of disability. This implies that they were not given a disability percentage in (at least) the first month they received a benefit. It is assumed in the transition probabilities as approached by the different models, that the unknown disability percentages in the first month(s) have the same disability percentage as the first known month. There is no clear evidence that this is the best approach, but also none to discard this approach.

6.1.1.3 Indexation and net present value

In this study no indexation is taken into account. Furthermore the benefits to be paid to a policyholder in a specific month is given as a percentage of the insured amount of a policyholder (the percentage equals the average disability percentage of the state the policyholder is in). Therefore the net present value of the benefits is not calculated. These assumptions will be discussed in Chapter 7.3, since it would be of additional value to limit the assumptions and consider both indexation and taking the net present value.

6.1.1.4 Included cash flows

Premiums that need to be paid to insurance companies (by the policyholder) are not taken into account in this study. Also the IBNR claims are disregarded, as already explained in Section 3.1.2. Only the dataset as received is taken into account. Hence no inflow of new claimants is considered.

6.1.2 Comparing of significant variables

The significant covariates in the different models for the different transitions can be compared regarding their signs. The conclusion can be drawn that in none of the different transitions, the same explanatory variables are included (except for transition fully disabled to active) in the different models. The same conclusion can be drawn regarding the sign of the covariates. In none of the different transitions, the same signs are allocated to the significant variables of different models.

Even though the covariates may not match, this does not have to imply that the transition probabilities of the different models differ completely. The transition probabilities, as given in Figure 6.1, are more important to

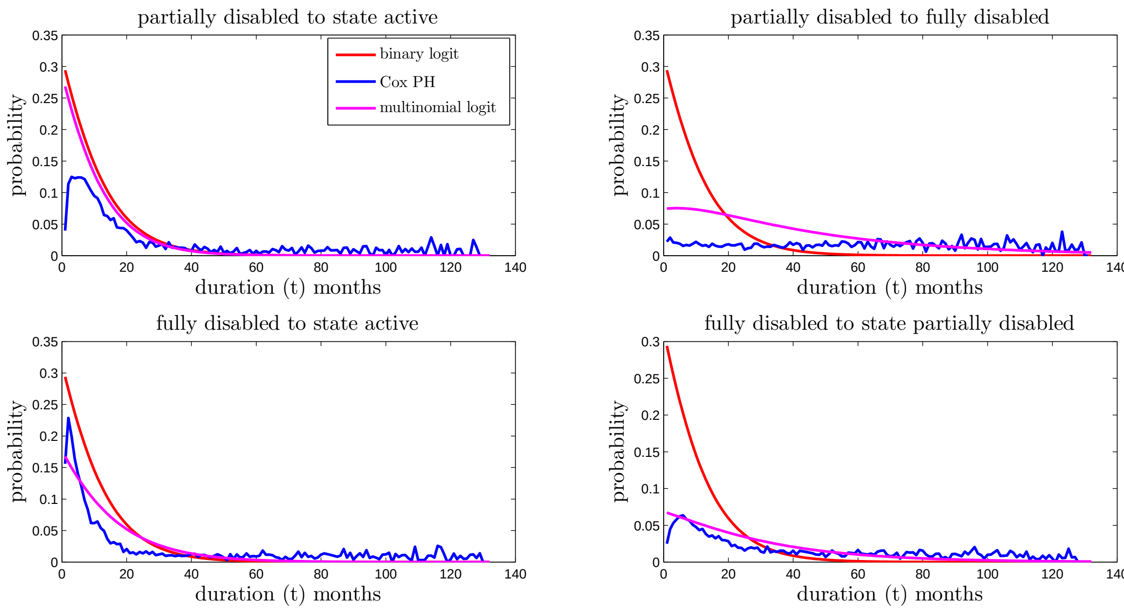

Figure 6.1: Transition probabilities for male policyholders of 25 years old (at the start of the disability), with an insured amount of €25,000, class of profession three and previous state 0

investigate. In this figure, it can be seen that when duration is low, the dynamic binary logit model appears to have a high probability. For the transition probability of the state of being partially disabled to the state of being active, the dynamic multinomial logit model goes with the line of the dynamic binary logit model. The Cox PH model shows a lot of fluctuation for all transition probabilities due to the non-parametric baseline hazard.

6.1.3 Benefits

The models explained in the previous chapters are used to model the transition probabilities. Based on the transition probabilities, it is possible to estimate the distribution of the benefits that need to be paid to the policyholders within the subsequent twelve months to the measurement date. The benefits regarding twelve months only are considered since this is of interest in case of the Solvency II regulation (SCR is over one-year horizon).

The expected value of the benefits is based on generated paths. For each model, 10,000 paths are generated of the benefits to be paid to each of the policyholders. The sum of the benefits to be paid to the separate policyholders is the expected value of the benefits to be paid by the insurance company. As is clear intuitively, this expected value may vary in case of generating 10,000 paths multiple times. This is why a 95% confidence interval for the expected value of the benefits is given by

$$\left(\text{expected value of the benefits} \pm 1.96 \cdot \sqrt{\frac{\widehat{\text{var}}(\text{benefits})}{\text{n}}}\right)$$

where n equals the number of observations. In this study there are 10,000 of observations.

Next to the expected value, the 95% quantile, the 99.5% quantile and the 95% coverage interval of the benefits to be paid are given in this section. The quantiles and the coverage interval refer to the model uncertainty. It is given that the probability to transfer between two states is a specific percentage, however this does not imply that precisely that percentage transfers. The 99.5% quantile is given since this is of importance regarding the Solvency II requirement, which states that the insurance company needs to be able to pay the benefits of the event that occurs only once each two hundred years. The 95% quantile is less volatile and is given for the full overview. The 95% coverage interval is given by

$$(q_{2.5}, q_{97.5}),$$

where q_x equals the x% quantile of the distribution. The coverage interval is therefor the interval which contains 95% of the estimated distribution of the benefits to be paid to the policyholders within the susbsequent twelve months to the measurement date. Next to model uncertainty there is parameter uncertainty. Parameter uncertainty will be discussed in Section 6.2.

Benefits to be paid in case of the dynamic binary logit model

In Figure 6.2, the benefits that need to be paid to the 5,501 policyholders for which the data are right censored at the measurement date of the dataset, are shown. Figure 6.2 shows a histogram which is based on 10,000 generated paths, with the transition probabilities based on the dynamic binary logit model. The generated paths represent the benefits that the insurance company need to pay to the policyholders within the subsequent twelve months to the measurement date.

From the estimated distribution it is possible to subtract the expected value of the benefits, the 95% confidence interval of this estimate, the 99.5% quantile and the 95% quantile, and the 95% coverage interval. These results are given in Table 6.1.

dynamic binary logit model	
expected benefits	54,805,000
95% confidence interval	
of expected benefits	(54,701,000; 54,909,000)
95% quantile	64,861,00
99.5% quantile	71,019,000
95% coverage interval of benefits	(46,268,000; 66,845,000)

Table 6.1: Expected value of the benefits, 95% confidence interval of expected value, 95% quantile, 99.5% quantile and 95% coverage interval of the benefits to be paid (in Euros)

The distribution of the benefits to be paid is based on the generated paths of 5,501 policyholders. Those 5,501 policyholders have an average insured amount of €24,035. Based on the expected benefits of €54,805,000, which is an expected benefit of ± €9,963 per policyholder, it is expected that the policyholders stays disabled for about five months subsequent to the measurement data.

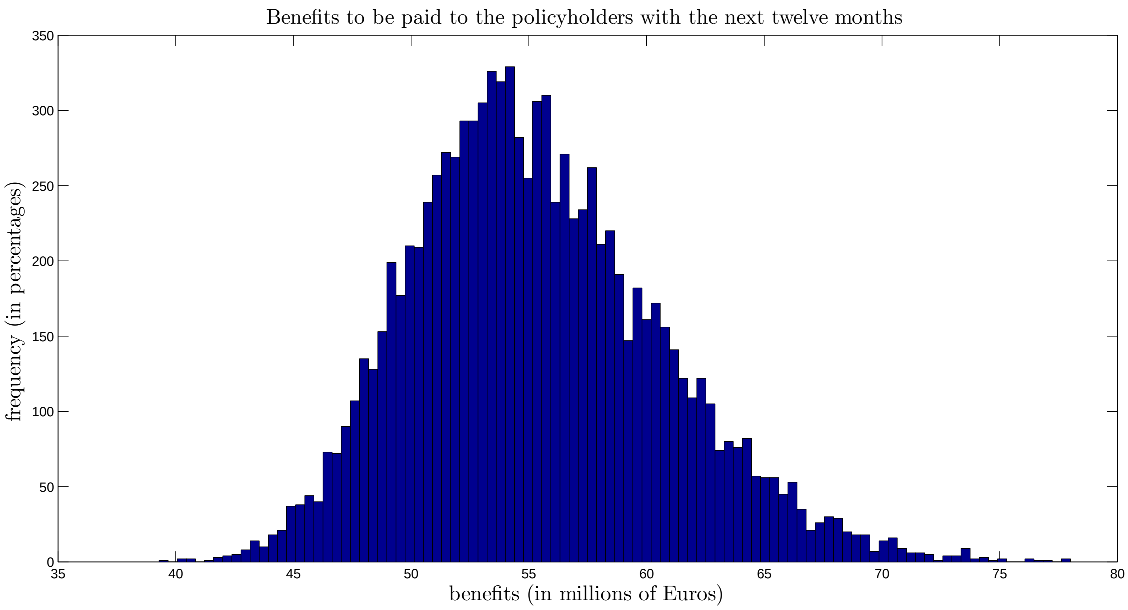

Figure 6.2: Benefits to be paid out to the policyholders in case of a dynamic logit model (in millions of Euros)

Benefits to be paid in case of the dynamic multinomial logit model

The expected value, the 95% confidence interval of this expected value, the 95% quantile and the 99.5% quantile, and a 95% coverage interval of the benefits to be paid are calculated and given in Table 6.2. The results as given in Table 6.2 are based on the transition probabilities as modeled by the dynamic multinomial logit model. In Figure 6.3, a histogram can be seen of the benefits to be paid to the policyholder within the subsequent twelve months to the measurement date.

dynamic multinomial logit model	
expected benefits	63,935,000
95% confidence interval	
of expected benefits	(63,803,000; 64,067,000)
95% quantile	76,294,000
99.5% quantile	81,729,000
95% coverage interval of benefits	(52,622,000; 78,303,000)

Table 6.2: Expected value of the benefits, 95% confidence interval of expected value, 95% quantile, 99.5% quantile and 95% coverage interval of the benefits to be paid (in Euros)

Based on the expected benefits of €63,935,000, which is an expected benefit of ± €11,622 per policyholder, it is expected that the policyholders stay disabled for almost six months subsequent to the measurement data.

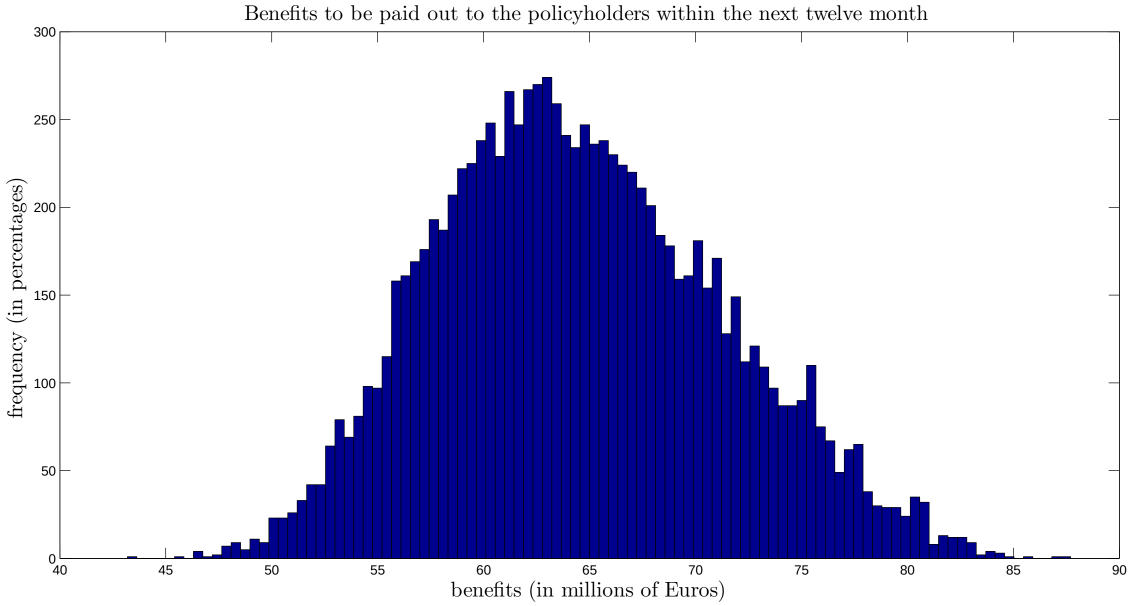

Figure 6.3: Benefits to be paid out to the policyholders in case of a multinomial logit model (in millions of Euros)

Benefits to be paid in case of the Cox PH model

In Chapter 5, the transition probabilities which belong to the Cox PH model have already been discussed. The results as given in Table 6.3 are based on the transition probabilities which belong to the Cox PH model. In Figure 6.4 the outcome is shown based on 10,000 generated paths for the 5,501 policyholders who have right censored data at the measurement date of the dataset. The path generations are about the benefits that the insurance company needs to pay to the policyholders.

Cox PH model	
expected benefits	79,460,000
95% confidence interval	
of expected benefits	(79,358,000; 79,562,000)
95% quantile	84,689,000
99.5% quantile	86,751,000
95% coverage interval of benefits	(65,986,000; 85,387,000)

Table 6.3: Expected value of the benefits, 95% confidence interval of expected value, 95% quantile, 99.5% quantile and 95% coverage interval of the benefits to be paid (in Euros)

Based on the expected benefits of €79,460,000, which is an expected benefit of ± €14,445 per policyholder, it is expected that the policyholders stay disabled for more than seven months subsequent to the measurement data.

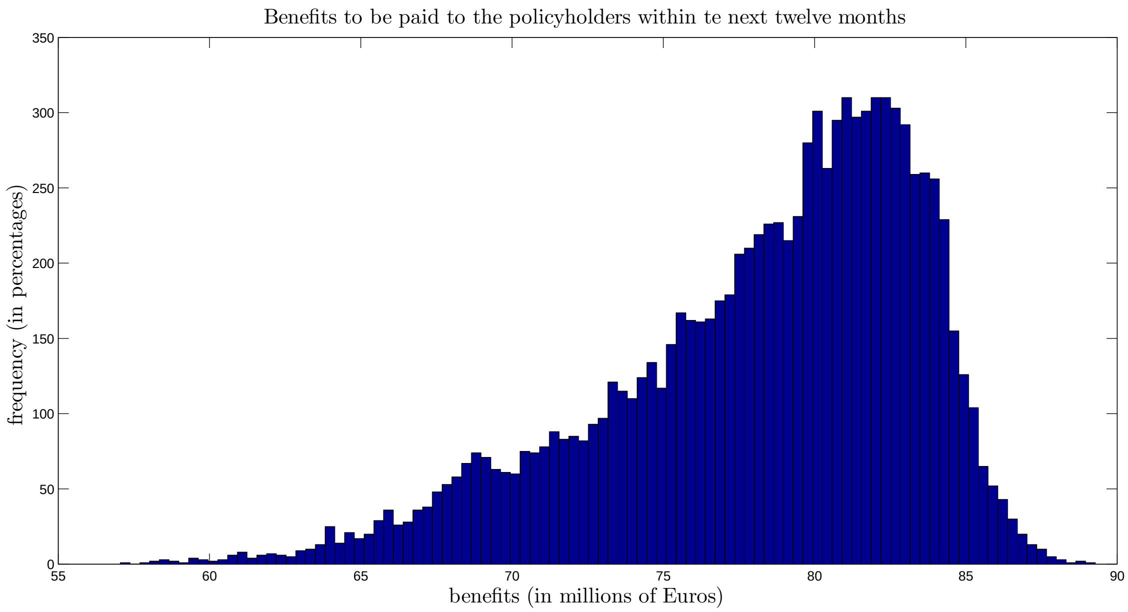

Figure 6.4: Benefits to be paid out to the policyholders in case of a Cox PH model (in millions of Euros)

6.2 Sensitivity analysis

The results as given in Section 6 are based on a disability model with three different disability states: in case a policyholder survives, the policyholder could be either in the active state, the partially disabled state or the fully disabled state. A policyholder is active in case the disability percentage is less than 25%, partially disabled in case the disability percentage is between 25% and 50% and fully disabled in case the disability percentage is more than 50%. Furthermore the estimated coefficients as given in Chapter 4 and Chapter 5 are considered. The results in Section 6 are dependent on the number of disability states, the division of the disability percentages between the different states and the coefficients taken.

In this chapter the effect is examined of a change in the number of disability states, a change in the division of the disability percentages of the different states or a change in the value of the coefficients. In Section 6.2.1 the effect of a change in the coefficients taken is discussed, where in Section 6.2.2 both the possibility of adding a disability state and changing the division of the disability percentages of the different states are considered and discussed.

6.2.1 Parameter uncertainty

The models as explained in Chapter 4 and Chapter 5 estimates the coefficients for the covariates as used in the models. However, the coefficients are estimated and therefore these coefficients may differ. This is why a standard error is given around these estimates. If it is assumed that the coefficient β follows a normal distribution, given both the estimate of β as the standard error of β, it is possible to do sensitivity analyses regarding uncertainty of the parameters. This sensitivity analyses can be performed since the 95% confidence interval for β_i is given by

$$(\widehat{\beta}_i \pm 1.96 * SE(\beta_i)) \tag{6.1}$$

The 1.96 in equation (6.1) is derived from the normal distribution. $\widehat{\beta}_i$ is the estimated coefficient for covariate x_i.

In case a confidence interval needs to be estimated based on the parameter uncertainty, equation (6.1) is used. Consider that the interest is in the value of the benefits at the lower bound of a confidence interval. In this case the transition probabilities p^{10}, p^{20} and p^{21} need to be high ($\widehat{\beta}_i + 1.96 * SE(\beta_i)$) and the transition probability p^{12} need to be low ($\widehat{\beta}_i - 1.96 * SE(\beta_i)$). In case the interest is in the value of the benefits at the higher bound of a confidence interval, the opposite reasoning count. Using equation (6.1), it is possible to calculate a confidence interval for the parameter uncertainty. The confidence interval is referred to as *a confidence interval* due to the fact that the estimated coefficients are taken based on their 95% confidence intervals. This leads to a confidence interval which is higher then 95%. In Table 6.4, an overview is given for the confidence interval for the different models.

	binary logit model	**multinomial logit model**	**Cox PH model**
expected benefits	54,805,000	63,935,000	79,460,000
confidence interval of expected benefits	(43,807,000; 67,026,000)	(52,747,000; 73,277,000)	(75,208,000; 83,363,000)

Table 6.4: Parameter uncertainty: expected value of the benefits (as given in Section 6) and the 95% confidence interval of these benefits (in Euros)

6.2.2 Disability percentages

From Figure 3.3 it can be seen that the disability percentages policyholders have, are not equally distributed over all probabilities. For policyholder who are partially disabled, a lot of times the percentages of 30% and 40% are shown. For policyholders who are fully disabled, the percentages of 50% and 100% occur most of the time. However, 70% and 60% are also shown in the figure. Given Figure 3.3, it is possible to investigate whether the division of the disability percentages of these three disability states is optimal and whether the distribution of three disability states is enough?

In this study the state a policyholders is in could be either active (disability percentage of 0% to 25%), partially disabled (disability percentage of 25% to 50%) or fully disabled (disability percentage of 50% up and to 100%). In this section it is investigated whether the options of another division of the disability percentages of the three disability states and of adding a fourth disability state are of additional value.

6.2.2.1 Different division of the disability states

From information received from the insurance company which provided the data, it is given that the division of the disability states is different than assumed in this study. The state of actives is the same (0%-25%), however the state of partially disabled is larger (25% - 90%) and thus the state of fully disabled is smaller (90% - 100%). The transition probabilities are modeled again and with these transition probabilities it is possible to estimate the distribution of the benefits.

In Table 6.5 the results of the expected value of the benefits, the 99.5% confidence interval of the expected value, the 95% quantile and the 99.5% quantile are given for each of the different models.

Since this model assumes a different division of the disability percentages in the partially disabled state and in the fully disabled state, the average disability percentages of the partially disabled state and of the fully

	binary logit model	multinomial logit model	Cox PH model
expected benefits	58,766,000	62,166,000	77,682,000
95% confidence interval of expected benefits	(58,647,000; 58,885,000)	(62,042,000; 62,291,000)	(77,571,000; 77,792,000)
95% quantile	70,044,000	73,771,000	86,421,000
99.5% quantile	75,925,000	80,826,000	91,421,000

Table 6.5: Results of the benefits of the different models in case of a different division of the disability states (in Euros)

disabled state needed to be calculated again. In this model the average disability percentage for the partially disabled state is 49.81% and the average disability percentage for the fully disabled state is 99.76%.

6.2.2.2 Adding a disability state

Instead of changing the disability percentages of the states, another disability state could be added. In the state of fully disabled policyholders, the focus is on policyholders with a disability percentage of 50% or more. This implies that both policyholder with a disability percentage of 100% and policyholders with a disability percentage of 50% are represented in the fully disabled state. The difference between these two probabilities is big and it could be the case that policyholders who have a disability percentage of only 55%, behave totally different than policyholders who have a disability percentage of 95%. In fact, there is a big probability that they do behave differently. This is why it is considered to add a disability state. With this new disability state, the following four states are present:

- Active, policyholders who have a disability percentage of 0% to 25%;
- Partially disabled, policyholders who have a disability percentage of 25% to 50%;
- Heavily disabled, policyholders who have a disability percentage of 50% to 75%;
- Fully disabled, policyholders who have a disability percentage of 75% up and to 100%;

In Table 6.6 the expected value of the benefits, the 99.5% confidence interval of the expected value, the 95% quantile and the 99.5% quantile are given for each of the different models.

	binary logit model	multinomial logit model	Cox PH model
expected benefits	60,448,000	63,288,000	78,352,000
95% confidence interval of expected benefits	(60,401,000; 60,495,000)	(63,180,000; 63,396,000)	(78,287,000; 78,417,000)
95% quantile	64,587,000	73,343,000	82,897,000
99.5% quantile	67,174.000	78,794,000	84,867,000

Table 6.6: Results of the benefits of the different models in case of adding a different disability state (in Euros)

Because the change in number of disability states and thereby the change of disability percentages in the different disability states, new average disability percentages had to be calculated for the heavily disabled state and the fully disabled state. The average disability percentage for the partially disabled state is equal as in the original model, namely 33.86%. The average disability percentage for the heavily disabled state is 53.28% and the average disability percentage for the fully disabled state is 92.75%.

6.3 Area under the ROC curve

As introduced in Section 4.2, the area under the ROC curve can be used to test the accuracy of a model. The higher the area under the ROC curve, the better the model. In Table 6.7 all the values for the areas under the ROC curves are given. From this table the areas under the ROC curves for the “original” model and for the model with a different division of the disability percentages can be compared. It can be concluded that the areas under the ROC curves for the “original” model have piecewise higher values than the areas under the ROC curves for the model with a different division of the disability percentages. The model with four different disability states is more difficult to compare with the models with only three different disability states, since the transition probabilities to do not perfectly overlay. The areas under the ROC curve are discussed in more detail in Chapter 7.

“original” disability model		
	binary logit model	**multinomial logit model**
p_{10}	0.8797	0.8802
p_{11}	-	0.8285
p_{12}	0.6632	0.6512
p_{20}	0.8001	0.7985
p_{21}	0.6945	0.6847
p_{22}	-	0.7638
Disability model with different division of disability percentage		
	binary logit model	**multinomial logit model**
p_{10}	0.8412	0.8409
p_{11}	-	0.8031
p_{12}	0.6457	0.6447
p_{20}	0.7689	0.7664
p_{21}	0.6915	0.6842
p_{22}	-	0.7300
Disability model with four disability states		
	binary logit model	**multinomial logit model**
p_{10}	0.8782	0.8792
p_{11}	-	0.8281
p_{12}	0.7082	0.7006
p_{13}	0.6370	0.6200
p_{20}	0.8496	0.8494
p_{21}	0.7464	0.7468
p_{22}	-	0.7973
p_{23}	0.6416	0.6348
p_{30}	0.7826	0.7799
p_{31}	0.6639	0.6612
p_{32}	0.7033	0.6996
p_{33}	-	0.7453

Table 6.7: The areas under the ROC curves for the different disability models for both the dynamic binary logit model and for the dynamic multinomial logit model

6.4 Backtesting of the models

One way to compare the different models as explained in both Chapter 4 and Chapter 5 is through backtesting. Backtesting implies that at a specific date within the dataset, it is checked which policyholders are in the partially disabled or fully disabled state. In case the policyholder is in either the partially disabled state or in the fully disabled state, it is estimated what the benefits to this policyholder will be in the twelve months subsequent to the date of backtest. Again 10,000 paths of the benefits to be paid are generated considering the modeled transition probabilities, the mortality rates according to the "Prognosetafel AG2012-2062" and the end age of the policyholder. In case a date of backtest lies before June 2011, data are available on the disability percentages of the policyholders in the twelve months subsequent to this date of backtest. This implies that it is possible to calculate the benefits that are paid to these policyholders in the twelve months subsequent the date of backtest. Comparing these benefits paid out with the expected benefits paid out by the different models, insight is given into the accuracy of the models.

The expected benefits to be paid out to the policyholders are estimated for multiple dates of backtesting for the different models. In Table 6.8 an overview of the expected benefits and the exact benefits, for all dates of backtesting, is shown.

It is assumed that all the policyholders (differentiating between male and female) follow the mortality table of "Prognosetafel AG2012-2062". The mortality rates of the year 2012 are applied (even if the data of backtest is before the year 2012). This assumption needs to be made in order to make it possible to calculate the expected benefits.

date of backtest	**exact benefits**	**benefits binary logit model**	**benefits multinomial logit model**	**benefits Cox PH model**
July 2003	13,976,000	11,834,000	13,743,000	17,271,000
January 2004	14,958,000	12,636,000	14,618,000	18,091,000
July 2004	15,518,000	13,329,000	15,270,000	18,535,000
January 2005	19,156,000	16,612,000	19,091,000	23,253,000
July 2005	21,568,000	18,441,000	21,158,000	25,388,000
January 2006	22,925,000	20,375,000	23,358,000	28,128,000
July 2006	24,344,000	21,061,000	23,983,000	28,013,000
January 2007	27,093,000	23,436,000	26,631,000	31,141,000
July 2007	28,470,000	24,558,000	27,917,000	32,245,000
January 2008	30,479,000	26,409,000	29,883,000	34,135,000
July 2008	33,023,000	28,026,000	31,576,000	35,696,000
January 2009	36,141,000	30,520,000	34,448,000	38,829,000
July 2009	39,664,000	33,182,000	37,420,000	41,948,000
January 2010	47,904,000	40,505,000	45,868,000	50,923,000
July 2010	49,713,000	41,230,000	46,617,000	51,420,000
January 2011	53,512,000	45,358,000	51,496,000	57,069,000
July 2011	55,828,000	45,910,000	51,916,000	56,570,000
mean	31,427,817	26,671,868	30,293,803	34,626,741
MSE		$28.4533 \cdot 10^{12}$	$2.6285 \cdot 10^{12}$	$11.2059 \cdot 10^{12}$

Table 6.8: Outcomes of backtesting the different models for different dates of backtest (in Euros)

The results as presented in Table 6.8 are also shown in Figure 6.5. In this figure the results of Table 6.8 are extended with 95% coverage intervals of the estimated benefits. The results as given in Table 6.8 as well as

Figure 6.5, are discussed in Chapter 7.

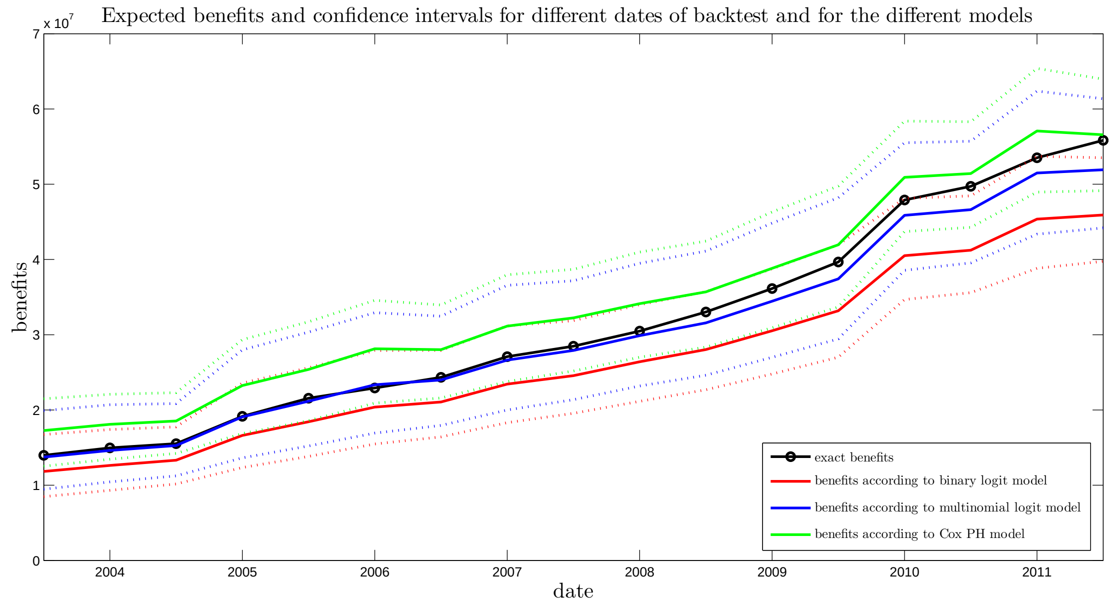

Figure 6.5: The expected benefits and the 95% coverage intervals of the expected benefits given on different dates of backtest, for different models. Including the exact benefits paid out

Chapter 7

Conclusion

7.1 Summary

In this study transition probabilities are approached based on different models. For this, in Chapter 4, discrete choice models are investigated. Both the dynamic binary logit model as well as the dynamic multinomial logit model are used to model the transition probabilities. Next to the discrete choice models, a survival analysis model is discussed in Chapter 5. The Cox PH model is a semi-parametric model which is used to approach transition probabilities. Because the baseline hazard of the Cox PH model is non-parametric, only the transition probabilities up to and including duration $t = 132$ can be modeled. The transition probabilities for a duration higher than $t = 132$ cannot be approached due to the fact that the dataset only includes information up to and including duration $t = 132$. The transition probabilities of the duration $t = 133, \cdots, 144$ are set equal to the transition probabilities of duration $t = 132$.

In both Chapter 4 and Chapter 5 the outcomes of modeling the transition probabilities, regarding the different models, are given. In Chapter 4 the estimates of the coefficient are given (Section 4.3.3 and Section 4.4.3) and by means of equation (4.3) and equation (4.7) the transition probabilities can be calculated respectively for the dynamic binary logit model as for the dynamic multinomial logit model. In Chapter 5, the estimates of the coefficients plus the estimates of the baseline hazard need to be combined with equation (5.11) in order to get the transition probabilities of both p^{11} as p^{22}, in equation (5.13) in order to get the transition probabilities of the transition from the initial state to the event of interest (p^{10} and p^{20}) and in equation (5.14) in order to get the transition probabilities of the transition from the initial state to the competing event, both p^{12} and p^{21}.

The models, as discussed during this study, approach transition probabilities in a different way. The dynamic binary logit model takes each transition probability apart, and assumes the transition of the state of being partially disabled to the state of being active independent of the transition probability of the state of being partially disabled to the state of being fully disabled. The same assumption holds for the transition probability of the state of being fully disabled to the state of being active and the transition probability of the state of being fully disabled to the state of being partially disabled. These assumptions does not seem to be right, and that is why it can be discussed whether the dynamic binary logit model is a good model to use in the first place. This model is included in this study since it gives us more insight and since this model led to the dynamic multinomial logit model. As already mentioned in Section 4.4, the dynamic multinomial logit model takes the transition probabilities p^{10}, p^{11} and p^{12} together, just as the transition probabilities p^{20}, p^{21} and p^{22}. The Cox PH model again approaches the transition probabilities separately, just as the case in the dynamic binary logit model. However, due to the competing risks analysis, there are no assumptions concerning the independence of the transition probabilities p^{10} and p^{12} and of the transition probabilities p^{20} and p^{21}.

In Section 5.17 it is discussed that the proportionality assumption of the Cox PH model can be tested based on plotting the logarithm of different survival functions (different regarding the covariates) and check whether the different plots are parallel. This test of proportionality is applied to the different transition probabilities, with different groups. For each transition probability, the significant covariates are considered and the survival function is applied to different groups regarding the significant covariate. For sex, in case the covariate sex is significant, the survival function is plotted which is based on the dataset of male policyholders only and the survival function is plotted which is based on the dataset of female policyholders only. The same reasoning is applied to the other significant variables. As discussed in Section 5.17, in case the survival functions cross it is known for sure that the proportionality assumption should be rejected. As can be seen in Appendix D, in each of the figures there is at least one subplot which shows two (or more) survival functions which cross. The fact that the survival functions cross would imply that the proportionality assumption as made in the Cox PH model is not valid, and again it can be discussed whether the Cox PH model is a good model to use in the first place.

The results as given in Chapter 6, Section 6.1.3, show that the expected benefits in case of the different models differ widely. The expected benefits in case of the dynamic binary logit model are €54,805,000, the expected benefits in case of the dynamic multinomial logit model are €63,935,000 and the expected benefits in case of the Cox PH model equals €79,460,000. Comparing the 95% coverage intervals of the benefits for the different models, it can be seen that all coverage intervals overlay. The coverage interval of the benefits according to the dynamic binary logit model and the coverage interval of the benefits according to the Cox PH model only have an overlay of €858,431. The overlay of the coverage interval of the benefits according to dynamic multinomial logit model and the coverage interval of the benefits according to both the dynamic binary logit model as the Cox PH model is larger. However, it has to be noticed that the coverage interval of the benefits according to the dynamic multinomial logit model is larger on its own, and so the probability of overlay is higher as well. Based on the coverage intervals only, it is not possible to draw a conclusion about which model is preferred.

The area under the ROC curve, as explained in Section 4.2, is a tool to compare different models. As can be seen throughout this study, the ROC curves, and specifically the areas under the ROC curves, are only calculated for the dynamic binary logit model and for the dynamic multinomial logit model. This makes the area under the ROC curve not applicable to draw conclusions concerning all models. However, the area under the ROC curve did gave a lot of information regarding the sensitivity analysis. As discussed in Section 6.2.2.1, based on the areas under the ROC curves, the models (both the dynamic binary logit model as well as the dynamic multinomial logit model) with a different division of the disability states turn out to be less accurate than the models with the "original" division of the disability states. Therefore, it is recommended to approach the transition probability based on the division of 0% - 25% (active), 25% - 50% (partially disabled) and 50% - 100% (fully disabled). Comparing the areas under the ROC curves of the original discrete choice models (three different states of disability) with the discrete choice models with four different states of disability, it is more difficult to draw conclusions. This since the transition probabilities do not interfere anymore (the transition probabilities p^{12} in one model is not the same as the transition probability p^{12} in the other model) and some areas need to be taken together to compare the areas under the ROC curves. It is assumed that the model with four different disability states is better in approaching transition probabilities due to the fact the coefficients can be estimated more precisely. It need to be taken into account that the more states a policyholders can be in, the more transition probabilities need to be estimated and the more difficult the model will be. Furthermore, the more transition probabilities present in a model, the lower the amount of observations for each transition probability and therefore the less precise the estimates of the coefficients will be. In this study, with only three or four different disability states and with the dataset used, this is not a problem.

Next to the areas under the ROC curves and the confidence intervals of the benefits for the different models, a test of backtest is applied on multiple dates. Based on this backtest, it is given how high the estimate of the benefits to be paid to the policyholder for the subsequent twelve months to the date of backtest is for the different models. This backtest is performed on seventeen different dates, ranging from July 2003 till July 2011 with an interval of six months between each date. In Table 6.4 the Mean Squared Error for the different models is given. The MSE of the dynamic multinomial logit model is smallest. Furthermore it can be seen from Figure 6.5 that the exact benefits always lie inside the 95% coverage interval given by the dynamic multinomial logit model. This also counts for the 95% coverage interval of the benefits according to the Cox PH model.

Overall it can be concluded that the dynamic multinomial logit model is preferred over the dynamic binary logit model as well as over the Cox PH model. Not only the doubts about the dynamic binary logit model and the Cox PH model are taken into account, also the results from the backtests performed lead to this conclusion. Based on this study and the particular dataset as used in this study, it is recommended to insurance companies to use the dynamic multinomial logit model, over the dynamic binary logit model as well as over the Cox PH model, for modeling the transition probabilities.

7.2 Discussion

As mentioned in Section 4.3.2 and in the introduction of Chapter 5.3, throughout this study it is assumed that the clock forward way is applied. This assumption is made for the discrete choice models as well as for the Cox PH model. The clock forward definition of duration implies that the duration of the disability is defined as the duration of the disability from the start of the disability onward. It can be doubted whether this assumption is applicable in the Cox PH model. Because of the doubts of this assumption on case of the Cox PH model, the Cox PH model is also investigated regarding the clock reset definition of duration. The clock reset definition implies that the duration is defined as the duration from the time the policyholder entered the specific state. If the clock reset way is assumed in the Cox PH model, and the distribution of the benefits to be paid to the policyholders within the subsequent twelve months to the measurement is estimated, the expected value for the benefits is €39,827,000. Compared to the values as discussed in Chapter 6, this value is low. Also the backtest is applied for the Cox PH model with the assumption of the clock reset definition of duration, which lead to a MSE of $142.5153 \cdot 10^{12}$. This value of the MSE is much higher than the value for the Cox PH model with the assumption of the clock forward definition of duration. Because of the results find both in case of the clock reset definition of duration and in case of the clock forward definition of duration, it is decided to focus on the Cox PH model which assumes the clock forward definition of duration.

Although this study could be interesting for insurance companies, it is not complete. It is certain that transition probabilities (as discussed in this study) are of importance for insurance companies, however the disability probabilities might be more important. Next to the disability probabilities, as stated in Section 6.1.1.4, IBNR claims are not taken into account and neither are the inflow of new claimants. It is recommended to further research to include IBNR claims and the inflow of new claimants into the modeling of transition probabilities and investigate the disability probabilities as well. This would lead to a complete research on disability insurance and transition probabilities.

One of the assumptions made in this study is that the indexation rate of the insured amount of the policyholders equals the interest rate taken into account when calculating the net present value. This implies that the insured amount at the moment of disability is the insured amount which need to be paid to the policyholder in each year subsequent to the start of disability, when taking the net present value of the insured amount which is indexed. In this case there is assumed that there is no such thing as an interest yield which implies that the interest rate increases in case a longer period is taken into account. Investigating whether the assumption of an interest yield makes an influence on the estimated benefits to be paid and thus on the accuracy of the models would be of additional value to this study.

7.3 Recommendations for further research

Considering Figure 6.5, it can be seen that the benefits to be paid according to the multinomial logit model (the expected benefits to be paid) are lower than the exact benefits at each date of backtest. This makes the statement that the dynamic multinomial logit model is preferred over the other models less strong. It would be better if the insurance company assumes a model which sometimes under performs the exact outcome and sometimes over performs the exact outcome. Therefore it could be a solution if the dynamic multinomial logit model is taken together with the Cox PH model. However, including the Cox PH model would make the model a lot more complicated and it is not known whether this would lead to a better estimate. It could be investigated whether the combination of multiple models will even lead to a better model to estimate the transition probabilities and therefore lead to a better expected value.

In this study the transition probabilities are modeled based on different models. In Chapter 4 the discrete choice models used are explained and in Chapter 5 the Cox PH model is explained. The discrete choice models and the semi-parametric survival analysis model are used for modeling the transition probabilities. However, it would be of additional value to investigate a parametric survival analysis model. A parametric survival analysis model could be for example the parametric survival model which assumes the Weibull distribution. This model, as explained in e.g. Klein and Moeschberger (1997), makes an assumption on the underlying distribution, namely that it is distributed according to the Weibull distribution. As seen in Figure 6.1, the semi-parametric Cox PH model has a fluctuating path, whereas the dynamic binary logit model as the dynamic multinomial logit model follow a smooth path. These fluctuations occur due to the non-parametric form of the baseline hazard, and this is why extending this study with a parametric survival analysis model could be of additional value. Due to limited amount of time, this model is excluded in this study.

Bibliography

J.D. Angrist and J. Pischke. *Mostly harmless econometrics: an empiricist's companion.* Princeton: Princeton University Press, 2009.

L.J. Bain and M. Engelhardt. *Introduction to probability and mathematical statistics.* Pacific Grove, CA : Duxbury, 1992.

G. Bakoyannis and G. Touloumi. A practical guide on modeling competing risk data. (unpublished).

F. Bartolucci. *Analysis of binary panel data by static and dynamic logit models.*, December 2009.

J. Beyersmann, A. Buchholz, A. Latouche, and M. Schumacher. Simulating competing risks data in survival analysis. *Statistics in Medicine*, 28(6):956–971, March 2009.

P.W. Bultena. Master's thesis actuarial studies: Application of claim reserving in a multiple state model for disability insurance. Master's thesis, University of Groningen, 2009.

M.L. Calle, G. Gmez, N. Malats, and N. Porta. Competing risks methods. December 2007. (unpublished).

A.C. Cameron and P.K. Trivedi. *Microeconometrics: methods and applications.* New York, NY : Cambridge University Press, 2005.

Y.Q. Chen and M. Wang. Analysis of accelerated hazards models. *Journal of the American Statistical Association*, 95(450):608–618, June 2000.

D. Faraggi and B. Reiser. Estimation of the area under the roc curve. *Statistics in Medicine*, 21(20): 3093–3106, October 2002.

T. Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874, June 2006.

F.K. Gregorius. Disability insurance in the netherlands. *Insurance: Mathematics and Economics*, 13(2): 101–116, November 1993.

S.P. Jenkins. Survival analysis. (unpublished), 2004.

J.D. Kalbfleish and R.L. Prentice. *The statistical analysis of failure time data.* New York [etc.] : Wiley, 1980.

J.P. Klein and M.L. Moeschberger. *Survival Analysis Techniques for Censored and Truncated Data.* New York [etc.] : Springer, 1997.

T.J. Klein. *Lecture Notes in Microeconometrics.*, October 2011. (Tilburg University).

R.H. Koning and L. Spierdijk. Claim reserving in a multi-state model for income insurance. (unpublished), May 2011.

M. Salm. *Lecture notes Microeconometrics course, Lecture 5 & 6.*, 2011. (Tilburg University).

D. Van Waarden. Master's thesis mathematics: Disability income insurance. Master's thesis, University of Amsterdam, 2012.

M. Verbeek. *A guide to Modern Econometrics.* Southern Gate, Chichester, West Sussex, England ; Hoboken, NJ : John Wiley & Sons, 2004.

G.G. Yin and Q. Zhang. *Discrete-Time Markov Chains. Two-Time-Scale Methods and Applications.* New York, NY : Springer, 2005.

Appendix A

Binary logit model results

p^{10}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	0.2906 (0.1561)	1.3372
duration	-0.0992 (0.0014)*	0.9056
sex	-0.0972 (0.0466)**	0.9074
age	-0.0236 (0.0016)*	0.9767
insured amount	$-0.0174 \cdot 10^{-4}$ $(0.0131 \cdot 10^{-4})$	1.0000
class of profession two	-0.2574 (0.1370)	0.7731
class of profession three	-0.3227 (0.1362)**	0.7242
class of profession four	-0.3420 (0.1313)*	0.7104
previous state	0.4840 (0.0156)*	1.6226
p^{12}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-2.7169 (0.2611)*	0.0661
duration	-0.0198 (0.0008)*	0.9804
sex	-0.1376 (0.0677)**	0.8714
age	-0.0058 (0.0023)**	0.9942
insured amount	$-0.0049 \cdot 10^{-4}$ $(0.0194 \cdot 10^{-4})$	1.0000
class of profession two	0.3536 (0.2377)	1.4242
class of profession three	0.3796 (0.2358)	1.4618
class of profession four	0.4397 (0.2298)	1.5522
previous state	-0.0715 (0.0193)*	0.9310
p^{20}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-1.5527 (0.1064)*	0.2117
duration	-0.0717 (0.0010)*	0.9308
sex	0.2963 (0.0312)*	1.3448
age	-0.0166 (0.0009)*	0.9835
insured amount	$-0.0922 \cdot 10^{-4}$ $(0.0086 \cdot 10^{-4})$*	1.0000
class of profession two	0.3070 (0.0980)*	1.3594
class of profession three	0.3081 (0.0968)*	1.3608
class of profession four	0.4952 (0.0933)*	1.6409
previous state	-0.1919 (0.0384)*	0.8254

* significant at $\alpha = 0.01$

** significant at $\alpha = 0.05$

continued on next page

continued from previous page

p^{21}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-2.6103 (0.1323)*	0.0735
duration	-0.0346 (0.0007)*	0.9659
sex	0.0612 (0.0376)	1.0631
age	-0.0174 (0.0012)*	0.9827
insured amount	$0.0551 \cdot 10^{-4}$ ($0.0105 \cdot 10^{-4}$)*	1.0000
class of profession two	0.2997 (0.1199)**	1.3495
class of profession three	0.3297 (0.1189)*	1.3905
class of profession four	0.3761 (0.1148)*	1.4566
previous state	1.0223 (0.0292)*	2.7795

* significant at $\alpha = 0.01$

** significant at $\alpha = 0.05$

Table A.1: Estimated coefficients and standard errors for the different covariates for the different transition probabilities, according to the binary logit model

Appendix B

Multinomial logit model results

p^{10}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	2.4139 (0.2959)*	11.1775
duration	-0.0783 (0.0016)*	0.9246
sex	0.0386 (0.0789)	1.0394
age	-0.0141 (0.0027)*	0.9860
insured amount	$-0.0099 \cdot 10^{-4}$ $(0.0225 \cdot 10^{-4})$	1.0000
class of profession two	-0.5848 (0.2662)**	0.5572
class of profession three	-0.6479 (0.2643)**	0.5231
class of profession four	-0.7342 (0.2570)*	0.4799
previous state	0.4497 (0.0191)*	1.5678
p^{11}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	2.1905 (0.2631)*	8.9396
duration	0.0238 (0.0008)*	1.0240
sex	0.1606 (0.0680)**	1.1742
age	0.0098 (0.0024)*	1.0099
insured amount	$-0.0070 \cdot 10^{-4}$ $(0.0194 \cdot 10^{-4})$	1.0000
class of profession two	-0.3408 (0.2385)	0.7112
class of profession three	-0.3801 (0.2365)	0.6838
class of profession four	-0.4125 (0.2305)	0.6620
previous state	0.0755 (0.0158)*	1.0784
p^{20}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-1.3910 (0.1068)*	0.2488
duration	-0.0704 (0.0010)*	0.9320
sex	0.2989 (0.0313)*	1.3483
age	-0.0176 (0.0010)*	0.9825
insured amount	$-0.0905 \cdot 10^{-4}$ $(0.0086 \cdot 10^{-4})$*	1.0000
class of profession two	0.3016 (0.0982)*	1.3520
class of profession three	0.3117 (0.0970)*	1.3658
class of profession four	0.4780 (0.0935)*	1.6129
previous state	-0.1632 (0.0154)*	0.8494

* significant at $\alpha = 0.01$

continued on next page

** significant at $\alpha = 0.05$

continued from previous page

p^{21}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
constant	-2.6107 (0.1327)*	0.0735
duration	-0.0361 (0.0007)*	0.9645
sex	0.0971 (0.0375)*	1.1020
age	-0.0198 (0.0012)*	0.9804
insured amount	$0.0721 \cdot 10^{-4}$ ($0.0104 \cdot 10^{-4}$)*	1.0000
class of profession two	0.3663 (0.1199)*	1.4424
class of profession three	0.4336 (0.1189)*	1.5427
class of profession four	0.4917 (0.1149)*	1.6351
previous state	0.3352 (0.0161)*	1.3982

* significant at $\alpha = 0.01$

** significant at $\alpha = 0.05$

Table B.1: Estimated coefficients and standard errors for the different covariates for the different transition probabilities, according to the multinomial logit model

Appendix C

Cox PH model results

p^{10}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
sex	-0.0435 (0.0407)	0.9574
age	-0.0129 (0.0014)*	0.9872
insured amount	$0.1758 \cdot 10^{-4}$ $(0.0109 \cdot 10^{-4})$*	1.0000
class of profession two	0.1098 (0.1192)	1.1161
class of profession three	0.0895 (0.1187)	1.0936
class of profession four	0.1432 (0.1146)	1.1539
previous state	-0.1829 (0.0138)*	0.8329
p^{12}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
sex	-0.1418 (0.0661)**	0.8678
age	0.0060 (0.0024)**	1.0061
insured amount	$0.1670 \cdot 10^{-4}$ $(0.0185 \cdot 10^{-4})$*	1.0000
class of profession two	0.5591 (0.2348)**	1.7490
class of profession three	0.4740 (0.2330)**	1.6065
class of profession four	0.6964 (0.2273)**	2.0064
previous state	-0.4012 (0.0191)*	0.6689
p^{20}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
sex	0.1649 (0.0297)*	1.1792
age	-0.0165 (0.0009)*	0.9836
insured amount	$0.0301 \cdot 10^{-4}$ $(0.0078 \cdot 10^{-4})$*	1.0000
class of profession two	0.4727 (0.0950)*	1.6044
class of profession three	0.4523 (0.0939)*	1.5719
class of profession four	0.6999 (0.0906)*	2.0136
previous state	-1.1082 (0.0367)*	0.3302

* significant at $\alpha = 0.01$

continued on next page

** significant at $\alpha = 0.05$

continued from previous page

p^{21}	$\widehat{\beta}$ (SE($\widehat{\beta}$))	exp($\widehat{\beta}$)
sex	0.0180 (0.0366)	1.0181
age	-0.0159 (0.0012)*	0.9842
insured amount	$0.1653 \cdot 10^{-4}$ $(0.0098 \cdot 10^{-4})$*	1.0000
class of profession two	0.4474 (0.1180)*	1.5642
class of profession three	0.4998 (0.1171)*	1.6485
class of profession four	0.5861 (0.1133)*	1.7970
previous state	-0.3853 (0.0272)*	0.6803

* significant at $\alpha = 0.01$

** significant at $\alpha = 0.05$

Table C.1: Estimated coefficients and standard errors for the different covariates for the different transition probabilities, according to the Cox PH model

Appendix D

Proportionality test

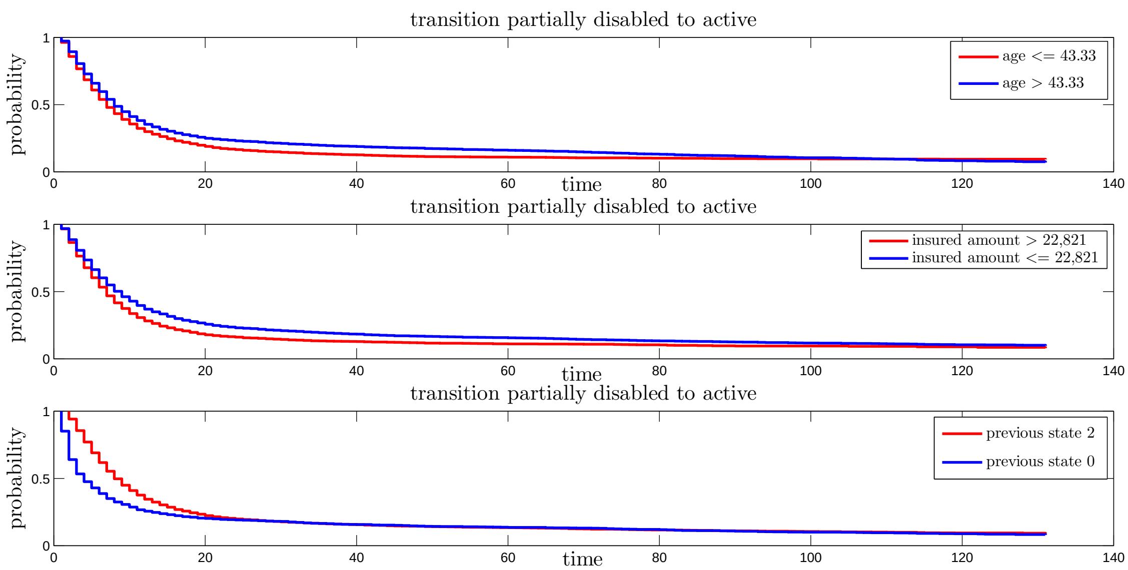

Figure D.1: Survival functions for different groups of policyholders (different characteristics), for the transition of the partially disabled state to the active state

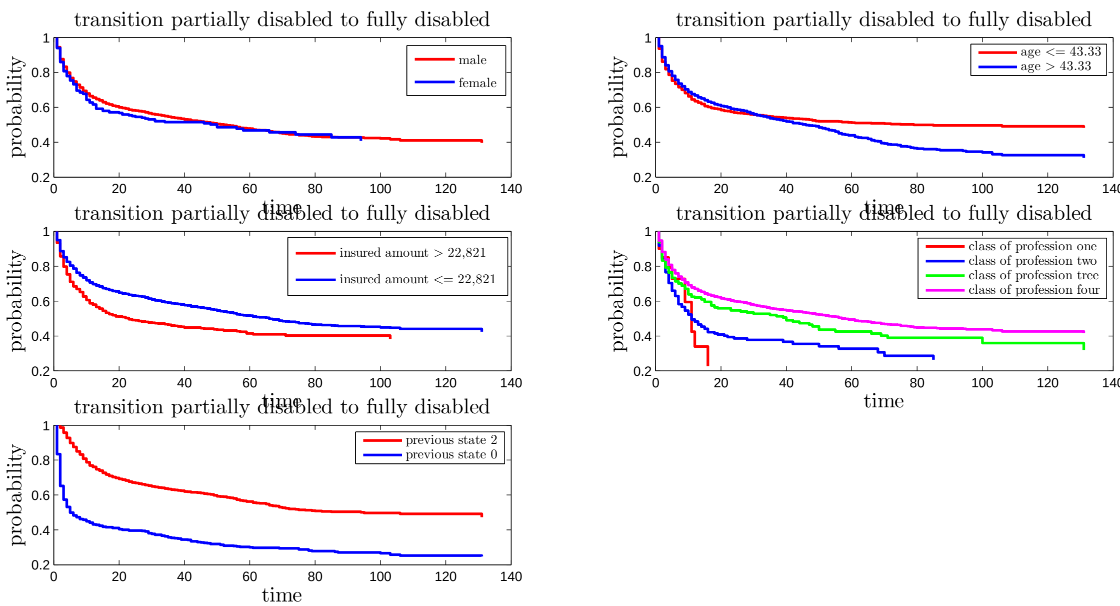

Figure D.2: Survival functions for different groups of policyholders (different characteristics), for the transition of the partially disabled state to the fully disabled state

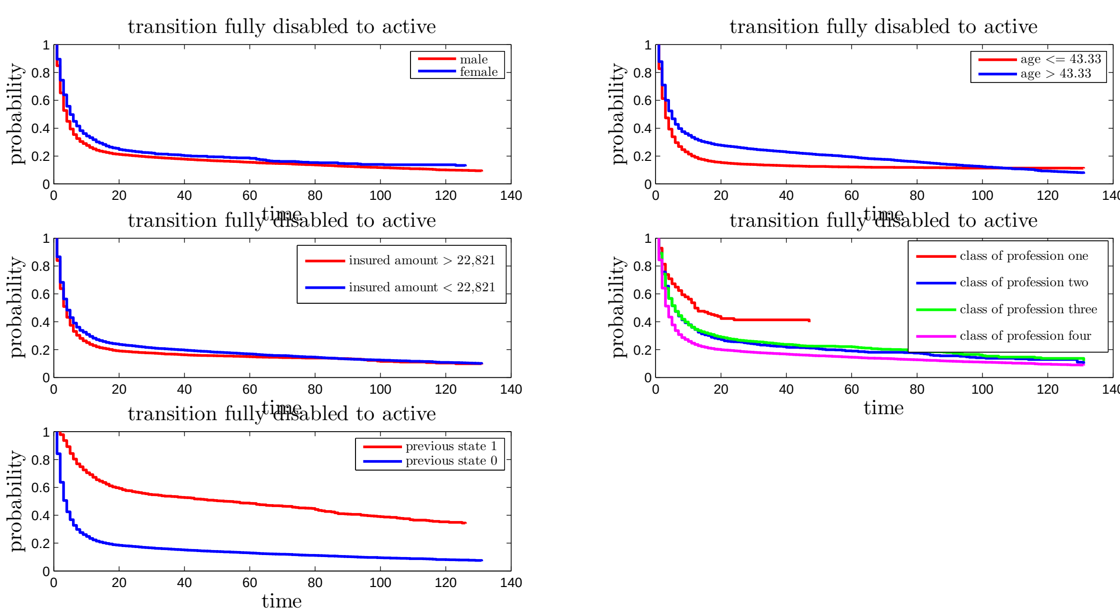

Figure D.3: Survival functions for different groups of policyholders (different characteristics), for the transition of the fully disabled state to the active state

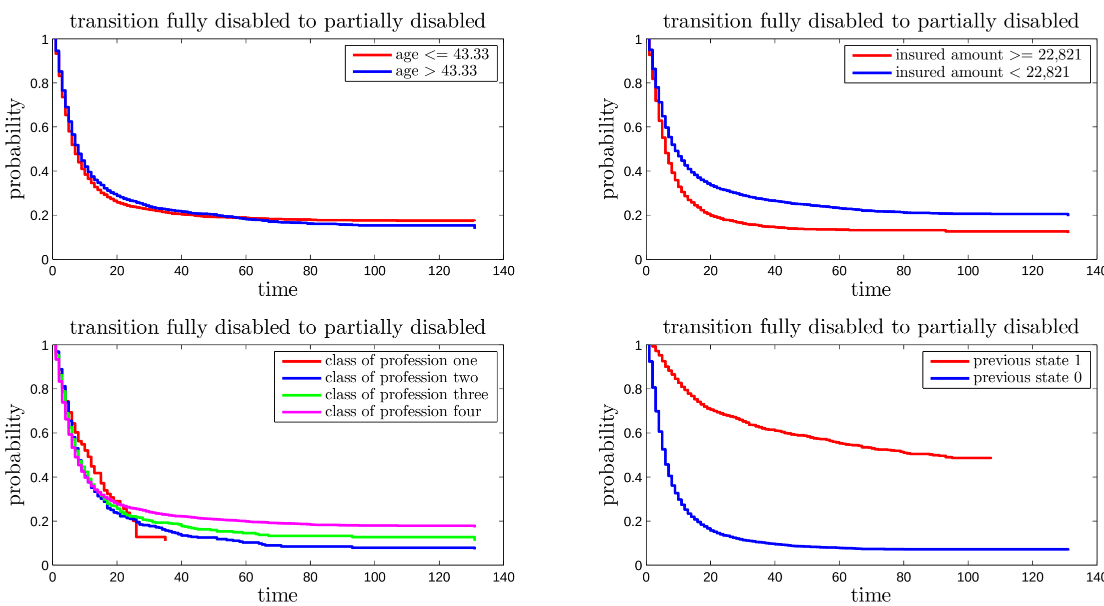

Figure D.4: Survival functions for different groups of policyholders (different characteristics), for the transition of the fully disabled state to the partially disabled state

Appendix E

Elaborated equations

$$L_i(t) = Pr(t = T_0)^{d_i^0} \cdot Pr(t = T_{CE})^{d_i^{CE}} \cdot Pr(t = T_C)^{(1-d_i^0-d_i^{CE})} \tag{E.1}$$

$Pr(t = T_0)$ is the probability that the event of interest will happen at duration t. This implies that the policyholder transfers to the active state at duration t. This is given by the probability that the competing event will not occur until duration t, so survival $S^{aCE}(t)$, multiplied by the probability that the event of interest happens at duration t, $f^{a0}(t)$. With the same reasoning it is given that $Pr(t = T_{CE}) = f^{aCE}(t) \cdot S^{a0}(t)$. The probability that a policyholder is censored at duration t implies that the event of interest did not occur up to duration t and neither did the competing event occured up to duration t. This implies that $Pr(t = T_C) = S^{a0}(t) * S^{aCE}(t)$. This gives:

$$\begin{aligned} L_i(t) &= \left(f^{a0}(t) \cdot S^{aCE}(t)\right)^{d_i^0} \cdot \left(f^{aCE}(t) \cdot S^{a0}(t)\right)^{d_i^{CE}} \cdot \left(S^{a0}(t) * S^{aCE}(t)\right)^{(1-d_i^0-d_i^{CE})} \\ &= \left(\frac{f^{a0}(t)}{S^{a0}(t)}\right)^{d_i^0} \cdot S^{a0}(t) \cdot \left(\frac{f^{aCE}(t)}{S^{aCE}(t)}\right)^{d_i^{CE}} \cdot S^{aCE}(t) \end{aligned} \tag{E.2}$$

Taking the logarithm of this equation, the equation is given by

$$\ell_i(t) = \log L_i(t) = d_i^0 \cdot \log\left(\frac{f^{a0}(t)}{S^{a0}(t)}\right) + \log\left(S^{a0}(t)\right) + d_i^{CE} \cdot \log\left(\frac{f^{aCE}(t)}{S^{aCE}(t)}\right) + \log\left(S^{aCE}(t)\right) \tag{E.3}$$

$$\ell_i(t) = \log L_i(t) = d_i^0 \cdot \log\left(\lambda^{a0}(t)\right) + \log\left(S^{a0}(t)\right) + d_i^{CE} \cdot \log\left(\lambda^{aCE}(t)\right) + \log\left(S^{aCE}(t)\right) \tag{E.4}$$

Appendix F

Mortality rates "Prognosetafel AG2012-2062"

	male		female	
	2012	**2013**	**2012**	**2013**
18	0.0003155	0.0002992	0.0001285	0.0001214
19	0.0003703	0.0003541	0.0001547	0.0001492
20	0.0004102	0.0003940	0.0001576	0.0001521
21	0.0004158	0.0003991	0.0001545	0.0001488
22	0.0004151	0.0003999	0.0001483	0.0001412
23	0.0004034	0.0003886	0.0001633	0.0001561
24	0.0003873	0.0003723	0.0001811	0.0001734
25	0.0003743	0.0003580	0.0001923	0.0001838
26	0.0003750	0.0003587	0.0002016	0.0001935
27	0.0003851	0.0003690	0.0002345	0.0002274
28	0.0004018	0.0003851	0.0002482	0.0002415
29	0.0004363	0.0004194	0.0002668	0.0002597
30	0.0004624	0.0004455	0.0002898	0.0002818
31	0.0005065	0.0004896	0.0003167	0.0003079
32	0.0005286	0.0005107	0.0003422	0.0003318
33	0.0005496	0.0005315	0.0003659	0.0003541
34	0.0005667	0.0005477	0.0003898	0.0003763
35	0.0006190	0.0005994	0.0004295	0.0004152
36	0.0006755	0.0006538	0.0004745	0.0004593
37	0.0007340	0.0007115	0.0005286	0.0005124
38	0.0008095	0.0007867	0.0005897	0.0005717
39	0.0009113	0.0008882	0.0006680	0.0006496
40	0.0009877	0.0009619	0.0007346	0.0007149
41	0.0010451	0.0010142	0.0008132	0.0007908
42	0.0010874	0.0010492	0.0009017	0.0008751
43	0.0011755	0.0011304	0.0010099	0.0009804
44	0.0013186	0.0012655	0.0011309	0.0010951

continued on next page

continued from previous page

	male		**female**	
	2012	**2013**	**2012**	**2013**
45	0.0014938	0.0014345	0.0012758	0.0012361
46	0.0016718	0.0016079	0.0013994	0.0013518
47	0.0019363	0.0018702	0.0015820	0.0015310
48	0.0021878	0.0021166	0.0017621	0.0017093
49	0.0024306	0.0023556	0.0019782	0.0019283
50	0.0026961	0.0026176	0.0022232	0.0021751
51	0.0029407	0.0028543	0.0024756	0.0024306
52	0.0032862	0.0031934	0.0028152	0.0027786
53	0.0036702	0.0035681	0.0031278	0.0030871
54	0.0040645	0.0039533	0.0034106	0.0033693
55	0.0045800	0.0044592	0.0036357	0.0035845
56	0.0051743	0.0050399	0.0039578	0.0039064
57	0.0057365	0.0055876	0.0042616	0.0042066
58	0.0062149	0.0060436	0.0045902	0.0045312
59	0.0067391	0.0065540	0.0049583	0.0048926
60	0.0073151	0.0071005	0.0053362	0.0052664
61	0.0079935	0.0077485	0.0057555	0.0056774
62	0.0089390	0.0086750	0.0061638	0.0060689
63	0.0099445	0.0096458	0.0067213	0.0066180
64	0.0111232	0.0107879	0.0072705	0.0071546
65	0.0123690	0.0119796	0.0079056	0.0077758

Table F.1: "Prognosetafel AG2012-2062"